# Intergenerational income mobility in New Zealand

**Leon Iusitini**

**A thesis submitted to**

**Auckland University of Technology**

**in fulfilment of the requirements for the degree of**

**Doctor of Philosophy (PhD)**

**2022**

**AUT Business School, Faculty of Business, Economics, and Law**

## Abstract

The transmission of economic advantage and disadvantage within families is a topic of considerable interest to policymakers and the general public in many countries including New Zealand, yet there has been very little research attention directed to understanding the extent to which income persists across generations in New Zealand or the mechanisms by which this occurs. This thesis aims to extend knowledge and understanding of this issue by quantifying and explaining intergenerational income persistence and mobility in New Zealand using two longitudinal data sources. Different analytical approaches are taken to these data sources reflecting differences in the nature of the samples and the income data collected in each. The first data source is the New Zealand Longitudinal Census, a dataset that links census microdata at the individual level over a 32-year timespan. The analysis of this data uses a 'permanent income' approach to estimate the intergenerational elasticity of individual total income between son-father, son-mother, daughter-mother, and daughter-father pairs, where sons are born between 1967 and 1979, daughters born between 1967 and 1974, and a proxy for the lifetime average income of offspring and parents is constructed, taking into account well-known sources of measurement error including attenuation and lifecycle biases. The results find that, by international comparison, New Zealand has a relatively low to middling degree of intergenerational income persistence (mobility is average to high), except for persistence between mothers and daughters which is comparatively high. The second data source is the Christchurch Health and Development Study, a longitudinal study tracking a birth cohort of children born in Christchurch, New Zealand in 1977 and followed for 40 years. The analysis of this data adopts a 'childhood resources' approach to estimate the intergenerational elasticity and rank correlation of sons' and daughters' lifetime average earnings and family income with respect to their parents' family income over the offspring's childhood. The results find a high degree of intergenerational income persistence (low mobility) by comparative standards. Both analyses are subjected to a range of robustness checks and sources of bias and limitations are described. Finally, both data sources are used to investigate mechanisms or pathways that mediate the intergenerational association in income, including the role of educational and occupational attainment in adulthood (in both datasets) as well as childhood behavioural problems and cognitive ability (in the Christchurch Health and Development Study). In general, educational attainment and occupation play important roles in intergenerational income persistence, but cognitive skills in childhood also stand out as a key early life mechanism linking parents' and offspring's incomes.

**Table of contents**

**List of figures**

## List of tables

**Attestation of authorship**

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

______________________________ 13th December 2021

Leon Iusitini

## Acknowledgements

Foremost gratitude goes to my primary supervisor, Professor Gail Pacheco, who championed my doctoral cause from the outset and stuck with me through changes in supervisors, changes in faculties, and other bureaucratic obstacles. Gail, you charted the course, navigated me through, and trusted in me to steer the ship. While I faced some headwinds and sometimes had to trim my sails, the understanding and respect you showed me made the voyage a less choppy, much smoother sail and prevented me from abandoning ship. Thank you.

My thanks to my secondary supervisors Dr Michael Fletcher and Dr Alexander Plum. Michael, your wisdom and eye for detail shone through in your thoughtful, considered comments on my thesis chapters. Alex, your econometric expertise and Stata prowess sped up the voyage and improved my own knowledge and skills, and your comments (including persistent efforts to rein in my often dense writing style) were helpful. Thank you both.

I am extremely grateful to Associate Professor El-Shadan Tautolo and Professor Janis Paterson, my current and former line managers at AUT, respectively, for their support and advice over the period of my doctoral study and many years prior. Dan, I simply could not have completed this thesis without the time allowance and other support you provided me. Janis, your encouragement, support, empathy, and humility over the years has been appreciated. Thank you.

Thank you to my partner Liz Pollard and our two children, Julius and George, for your love and support and for keeping me anchored. Liz, this would not have been possible without your help at home, keeping the ship on an even keel, which you skippered with me the whole way, through thick and thin. Thank you to my parents, Sue and Tafa Iusitini, and Liz's mother Dawn Rigby, for your love and practical support – especially caring for your grandchildren – over the years of my study.

Thank you to Professor John Horwood (Otago University) for supplying data for this thesis and for your recurrent help and patience in fielding my queries. Thank you to Dr Robert Didham (Stats NZ) for helpfully and promptly answering my queries about census data, and to Dr Simon Chapple (Victoria University of Wellington) and Professor Erling Rasmussen (Auckland University of Technology) for helpful advice and support at the outset.

Thank you to the following academic colleagues who encouraged me to undertake doctoral study: Emeritus Professor Charles Crothers (Auckland University of Technology), Associate Professor Barry Milne (University of Auckland), Professor Jennifer Curtin (University of Auckland). For inspiration, thank you to Dr Ross McDonald (University of Auckland) who, 17 years ago, unbeknownst to him, changed my life.

## Disclaimer

Access to the data used in this study was provided by Stats NZ under conditions designed to give effect to the security and confidentiality provisions of the Statistics Act 1975. The results presented in this study are the work of the author, not Stats NZ or individual data suppliers.

# **Chapter 1: Introduction**

The aim of this thesis is to quantify and explain the intergenerational transmission of income in New Zealand. As its first research question, this thesis asks, '*how closely associated are people's incomes to those of their parents in New Zealand?*'. Having estimated intergenerational income transmission, the thesis then seeks to explain the observed income transmission, asking as a second research question, '*what are the mechanisms that produce the intergenerational transmission of income in New Zealand?*'. There is a lack of research on intergenerational income dynamics in New Zealand (Easton, 2013a; Rashbrooke, 2013), despite an increasing focus by New Zealand policymakers on intergenerational wellbeing and inequalities in its distribution (Karacaoglu, 2015; Low, Gluckman, & Poulton, 2021; New Zealand Treasury, 2018). This thesis addresses this gap in knowledge.

The study of intergenerational transmission within families seeks to understand how people's outcomes in life – their educational attainment, income, occupation, or some other characteristic or trait – relate to their family background. Economists tend to focus on the intergenerational transmission of education, earnings, income, and wealth, while sociologists tend to focus on socioeconomic status and social class based on people's occupations which place them in a status or class hierarchy. The focus may be on continuity in the outcome across generations, termed *intergenerational persistence*, or change in the outcome across generations, termed *intergenerational mobility*. This thesis focuses primarily (but not exclusively) on persistence of income across generations.

Björklund and Jäntti (2020) outline four general approaches to the study of intergenerational transmission of income, one of which is to simply describe how closely associated people's incomes are with those of their parents. Such an approach is a 'reduced form' exercise that attempts to summarise intergenerational transmission in a single parameter or set of parameters and focuses solely on the joint distribution of offspring's and parents' incomes, ignoring other family background factors. Under this approach, estimates of these intergenerational parameters are descriptive statistical associations which are not given causal interpretations. As discussed in Björklund and Jäntti (2020), this descriptive approach contrasts with three other approaches that could be taken to the study of intergenerational income transmission, namely, identifying the causal effect of parental income and/or other family background characteristics on offspring's income (the intergenerational effects approach), estimating the association between the incomes of siblings (the sibling correlation approach, which captures not just the influence of parents but also school effects, neighbourhood effects, and any other background factors shared by siblings), and estimating the extent to which people's incomes are determined by circumstances outside of their control as compared with factors within their control, termed 'effort' (the equality of opportunity approach). This thesis adopts the first descriptive approach outlined above to the estimation of intergenerational income persistence in New Zealand.

This introductory chapter provides a contextual background for the analyses presented in the subsequent chapters. Section 1.1 explains why intergenerational income mobility and persistence is a topic worthy of investigation, with a focus on its normative foundations. Section 1.2 précises the reality and the myth of New Zealand's reputation as an egalitarian country, providing the historical context for the current analysis. Section 1.3 provides an overview of changes in the labour market over the period sampled in this thesis, including trends in individual incomes, these being the empirical basis for the subsequent analyses of intergenerational mobility and persistence.

**1.1 The significance of intergenerational income mobility**

Why should we care about intergenerational income persistence or mobility? Studies of intergenerational mobility are almost always, implicitly or explicitly, underpinned by a normative concern about equality of opportunity. The underlying presumption is that it is unfair if children from different family income backgrounds systematically fare differently in life (Roemer, 2012). If there is only a loose relation between parents' incomes and offspring's incomes, "then in some sense the playing field might be thought of as level, children's position in the income distribution being the result of their own efforts rather than accidents of birth" (Roemer, 2004, p. 48). Thus, higher mobility is considered to indicate greater equality of opportunity and fairness, based on "the meritocratic idea that someone's life chances should depend on their own abilities and efforts rather than on who their parents were" and other circumstances beyond their control (Jäntti & Jenkins, 2015, pp. 814-815). In addition, there will be strong economic reasons to be concerned about mobility and equality of opportunity if children from poorer families are not able to fully develop their productive potential, the wastage or underutilisation of which can reduce economic efficiency. Equality of opportunity "implies that all talents will be fully realised, and in this sense it means labour markets will be more efficient, that overall productivity will grow at a faster rate" (Corak, 2020b, p. 230). Inequality of opportunity and lack of mobility may also weaken social cohesion and trust in social institutions and the political system. As Nash (1993, p. 47) puts it, "economic and administrative efficiency is best served (and a potential source of social discontent alleviated) by systems which nurture talent and allow it to flourish wherever it might emerge". A study of intergenerational income mobility is therefore important on both efficiency and equity grounds.

However, intergenerational mobility is an *imperfect* indicator of equality of opportunity. The one need not imply the other. From a normative perspective, it is necessary to understand what gave rise to any intergenerational association in incomes – the reasons why higher-income parents tend to have higher-income offspring. Parents influence their children's life outcomes in a variety of ways. They act, intentionally or reflexively, in ways that enhance their children's wellbeing, and do things to, with, and for their children that further their interests (Swift, 2005). In addition to passing on their genes, parents may, for example, provide loving, emotionally-secure familial relationships, read to their children and engage them in intellectually-stimulating conversations, instil in their children traits of curiosity and perseverance, buy them enriching educational goods and experiences, buy a home in an affluent neighbourhood surrounded by similarly-advantaged families, purchase them a private education, use social connections to secure them jobs, or bequeath them an inheritance. Thus, whether the degree of intergenerational income persistence found in any given society at any given time is considered just or unjust (and judgements about whether it is 'too high' or 'too low') depend crucially on (1) which mechanisms produced the persistence and (2) whether the members of that society view those mechanisms as morally acceptable or not.

Some mechanisms "may be morally suspect…in a way that would justify attempts to deter or prevent them" while "[o]thers may be legitimate, perhaps even morally required" and "would exist even in an altogether just society" (Swift, 2004, p. 1; 2005, p. 256). An example of the former might be "the use of connections to get jobs by the children of the wealthy when other candidates are better qualified", which is "manifestly inefficient" so "the component of the intergenerational correlation due to nepotism would be considered by most to best be eradicated" (Black & Devereux, 2011, p. 1489). An example of the latter might be the intergenerational transmission of certain preferences, such as between work and leisure, or occupational choices. As Roemer (2012, p. 486) notes, "parents who are teachers may instil in their children a desire to be teachers, and bankers may transmit to their children the desire to earn high incomes". To the extent such legitimate preferences and choices are transmitted from parents to children, an intergenerational income association of zero will not necessarily be ethically desirable (Roemer, 2012).[1] Conventionally, it is also considered normatively legitimate and meritocratic that people are rewarded in the labour market for their productive capacities, including their innate abilities (Swift, 2004).[2] Consequently, "differences in ability and human capital will tend to lead to an intergenerational correlation of greater than zero in any well-functioning market economy" and "[p]olicies that compel employers to favor less qualified applicants in terms of employment or pay may reduce the intergenerational correlation but at a high cost to society in terms of efficiency and incentives for human capital accumulation" which "are not necessarily desirable" (Black & Devereux, 2011, p. 1489).[3]

More broadly, there are a myriad of parent-child interactions, of the kind that most would consider are constitutive of a valuable loving family relationship, such as reading bedtime stories to children, which are likely to perpetuate family advantage (Swift, 2005). Consequently, the egalitarian objective to equalise opportunity eventually butts up against the widely-held desire to keep the private sphere of the family immune from state interference. To this extent, the family is ultimately an obstacle to equality of opportunity, as "[t]hose who regard the family as a valuable institution are likely to see it as properly limiting the pursuit of equality of opportunity", highlighting "the difficulty in insulating children's equality of opportunity from parents' inequalities in outcome" (Swift, 2004, p. 7; 2005, p. 268). For all the reasons outlined above, "reasonable and well-informed observers are likely to disagree about the optimal level of intergenerational mobility" due

[1] However, an egalitarian will want to know whether the intergenerational transmission of such preferences is itself just, in that some may be 'adaptive preferences' formed under unfavourable circumstances which limit aspirations in an unduly pessimistic way (Swift, 2004). At what point, then, can we hold adults responsible for their preferences? Roemer (2004, p. 51) mentions one answer proposed in the philosophical literature, namely that "even though one's preferences may have been instilled to a large extent in childhood, one acquires responsibility for them if one comes to identify with them" and, citing Dworkin (1981), "a person should be held responsible for his preferences if and when he is glad he has them".

[2] As natural abilities are subject to the luck of the genetic lottery, 'luck egalitarians' (egalitarians who hold a more-radical conception of equality of opportunity) deem rewards to these abilities as 'undeserved', in the sense that attainment of these abilities do not reflect individual effort but rather circumstances beyond the individual's control, so individuals are not responsible for, and should not be rewarded on the basis of, innate abilities that are, from a moral point of view, arbitrary (Swift, 2004).

[3] Such incentive effects are also a relevant normative consideration, even if only on Rawlsian moral grounds, as Swift (2004, p. 9) explains: "If we care about productivity – even if only because we want to use the product to maximize the advantage of the worst off – then we have reason to permit some intergenerational transmission, despite its unequalizing impact and even if we regard that transmission as fundamentally unjust…that reason is the moral urgency of helping the worst off" as "[n]ot permitting it will worsen their position".

to "different value judgments about what constitutes a fair earnings distribution and about the extent to which efficiency losses should be suffered to achieve it" (Solon, 1999, p. 1795; 2008, p. 410).

Yet despite the conceptual discrepancies between intergenerational income mobility and equality of opportunity, in practice the two concepts have been found to be empirically highly correlated. Using Australian income tax data on over one million offspring-parent pairs, Deutscher and Mazumder (2021, p. 7) find that the intergenerational income elasticity – the main measure of intergenerational persistence used in this thesis – is "highly correlated with an index of inequality of economic opportunity", which suggests that "the very widely used and empirically driven descriptive measures of intergenerational persistence may do a reasonably good job of capturing the same concept of inequality of opportunity as a more theoretically grounded measure".

### 1.2 The reality and myth of egalitarianism in New Zealand

Since the onset of British colonial settlement, New Zealand has forged a self-image as an egalitarian and harmonious society, and a "myth of classlessness" has become central to the national identity of New Zealanders (Belich, 2001, p. 126). According to this narrative, during the late nineteenth and much of the twentieth century, New Zealand was characterised by high living standards, social equality, and political consensus or lack of class conflict, a view held by New Zealanders of themselves and by others commentating on New Zealand from the outside (Nolan, 2007, 2009).[4] This egalitarian ethos can be traced back to the early phase of organised British colonisation. The abandonment of the entrenched class distinctions of the Old World and the building of a more equal society was a deliberate objective of many settlers (Olssen, Griffen, & Jones, 2011). Furthermore, the experience of migration to New Zealand, "often in steerage for around three months, made the colony a fluid society from the outset" and "[t]he very process of moving to new and often crude settlements-in-progress eroded former differences in status and income", settlements where "land was freely available, there was little industrialisation and small workplaces produced intimate social relations" (Nolan, 2009, p. 369; Olssen et al., 2011, p. 194). The colony "did not possess a modern society's sophisticated division of labour" and there was "a premium on [skill] versatility, breaking down advantages based on previous occupational positions" and limiting consciousness of occupational distinctions (Fairburn, 1979, p. 47; Olssen et al., 2011, pp. 194, 195). These egalitarian beginnings were reinforced by the establishment in 1877 of the modern national education system which "helped to produce a society with narrow social differentiation" (Nolan, 2009, p. 357).[5] Social mixing – exogamous marriages and mixed schools and neighbourhoods where people from different classes, nationalities, and religions intermingled – characterised the frontier settler society. This intermingling further fostered the development of an egalitarian culture characterised by a belief in the dignity of manual work, a high regard for skill, dexterity, and physical strength, no marked differences between classes in

[4] Similarly, New Zealand's reputation as a 'social laboratory' – a place where social policy innovation, particularly in the 1890s and 1930s, led to the implementation of experimental progressive policies (presumed to increase social mobility) that attracted international attention – has been described by James (2020, pp. 27, 34) as "mostly self-ascribed but at times attested by foreign observers" and as being a claim that, by the Second World War, "had merit".

[5] The egalitarian foundations may have been facilitated by the fact that, firstly, members of England's aristocracy and gentry did not, in the main, migrate to New Zealand, and secondly, that the effect of wealth inheritance (hence inequality) was limited due to the brief history of the new settler colony (Olssen et al., 2011).

speech, type of education, or leisure activities, and a desire to live in a society where class and status were not central to social organisation (Olssen et al., 2011).

But as historians have pointed out (Easton, 2020; Nolan, 2007, p. 127; Wilkes, 1994), the narrative drew upon “a rich amalgam of truth and myth”, for while New Zealand may have been relatively egalitarian compared to other countries over periods of the late nineteenth and twentieth centuries, the egalitarian ideals were never extended to women, Māori, and other groups. Easton (2020) argues there was a closer correspondence between the myth and the reality in the three decades after the Second World War – when unemployment was low, women’s participation in the labour market was increasing, income inequality was relatively subdued, and social networks “reached far enough to provide most New Zealanders with knowledge about and some empathy for many other social groups” – than is likely to be the case for cohorts born after 1990, when reductions in social welfare benefits resulted in increases in poverty concentrated among households with children, reflecting the ‘anti-egalitarian’ effects of the Rogernomics neoliberal policy agenda (Easton, 2020, p. 616).[6]

### 1.3 The labour market context in New Zealand

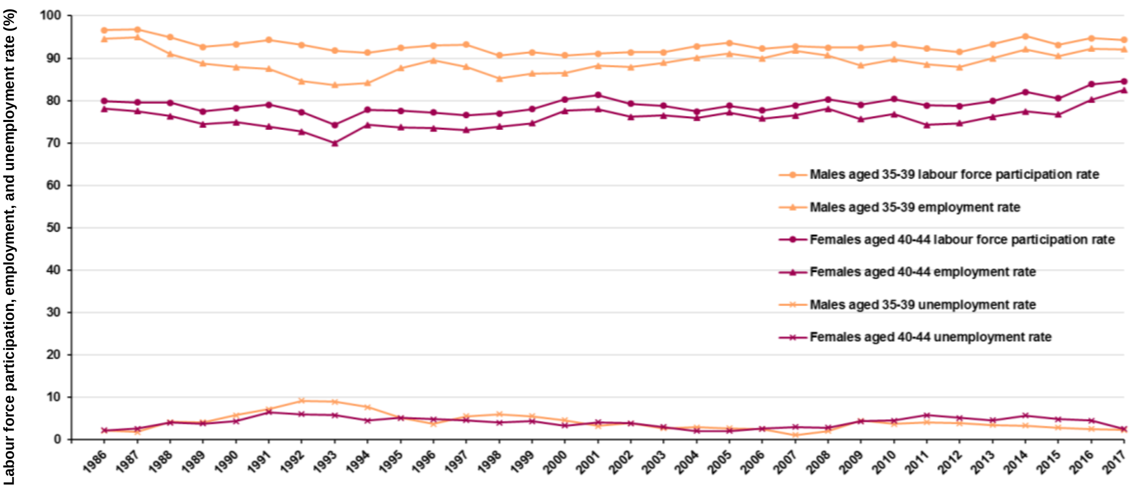
The data used in this thesis span the 40-year period from 1977 to 2017 for the Christchurch Health and Development Study and, encompassed within that, the 32-year period from 1981 to 2013 for the New Zealand Longitudinal Census. Significant changes in New Zealand’s labour market have occurred over this time. The beginning of these two periods, 1977 to 1981, marks a transition in New Zealand’s economic development from policies of ‘insulationism’ and industrialisation to de-industrialisation – declines in the manufacturing, construction, and agricultural sectors – and a rise in the services sector (Bertram, 2009). Reflecting this shift in the economy, the labour market over the period 1977 to 2017 is characterised by a changing occupational mix of increasing employment shares in professional and community and personal service occupations and declining shares in labouring and machine operation (Maré, 2019). Other changes in the labour market over the period include increasing labour force participation of women and older workers (Hyslop, Rice, & Skilling, 2019), increases in part-time employment and associated decreases in average weekly hours worked (Hyslop et al., 2019), increasing levels of educational attainment (Hyslop & Maré, 2005; Hyslop, Maré, & Timmins, 2003; Maré, 2018), and increasing earnings premia to tertiary qualifications over the period 1981 to 1991, but stabilisation and then declines through to 2017 (Maani, 1999; Maani & Maloney, 2004; Maré, 2018).

Labour force participation and employment rates among working-age males declined from 1986 to 1991, but have been fairly stable through to 2017, albeit lower in 2017 than in 1986. Among working-aged females, labour force participation and employment rates declined from 1986 to 1989, but have steadily increased through to 2017 (Hyslop et al., 2019). Trends in unemployment rates over the period were similar for men and women, increasing steeply from 1986 to 1991, then an overall decline through to 2017 interrupted by increases over the 1997 to 1999 recession linked to the Asian financial crisis and over 2008 to 2010 amidst the global financial crisis (Hyslop et al., 2019). Figure 1 plots rates of labour force participation, employment, and unemployment

[6] See also Cotterell (2017) and Rashbrooke (2021) who make similar arguments.

for males aged 35 to 39 years and females aged 40 to 44 years over the period 1986 to 2017. These age ranges are chosen because they are the ages of most relevance to this thesis; for reasons outlined in chapter 2, the analyses in this thesis generally compare the incomes of parents and offspring when males (fathers and sons) are around the age of 35 and females (mothers and daughters) are around the age of 40. Among males aged 35 to 39, labour force participation and employment rates were higher in 1986 (97 percent and 95 percent, respectively) than in 2017 (94 percent and 92 percent, respectively), while the unemployment rate was about the same (two percent in both years). Among females aged 40 to 44, labour force participation and employment rates were lower in 1986 (80 percent and 78 percent, respectively) than in 2017 (85 percent and 83 percent, respectively), while the unemployment rate was about the same (two percent in both years). The trend over time is similar for men and women, with generally declining participation and employment, and rising unemployment, from 1986 to 1993, followed by steadily increasing participation and employment and fairly stable unemployment through to 2017.

**Figure 1***. Labour market outcomes of men aged 35 to 39 and women aged 40 to 44 in New Zealand, 1986 to 2017.*




**Source:** The data source is a table of September quarter figures for each year from the Household Labour Force Survey available on Stats NZ's (2021a) Infoshare website.

What happened to individual incomes over this period?[7] While the precise details of the trend can differ depending on the data source, the general picture from quinquennial census data is as follows. Among working-age males, the median individual annual total income, in real (inflation-adjusted) terms, increased between 1976 and 1981, had fallen by 1986 and fell again by 1991, had increased modestly by 1996, then had increased more strongly by 2006 to end slightly below the 1981 peak (Callister, 2006; Callister, Von Randow, & Cotterell, 2011; Coleman, 2006; Coleman & McDonald, 2010; Martin, 1997a, 1997b, 1999, 2002).[8] Thus, the median working-age

[7] The following discussion of income trends focuses on the distribution of *individual* incomes. For analyses of changes in household incomes over the period sampled in this thesis, see for example Perry (2019), Hyslop and Maré (2005), Easton (1996), and O'Dea (2000).
[8] The census pattern over 1981 to 1996 is also observed for median and mean annual *market* income of working-age males, based on Household Economic Survey data (O'Dea, 2000; Statistics New Zealand, 1999). See also Hyslop and Yahanpath (2006) who analyse changes in individual total income (annualised weekly gross income from the Household Labour Force Survey) for the working-age population (males and females pooled) over the period 1998 to 2004, finding that median individual total income increased by 12 percent over the period, that income gains were spread broadly across the distribution, and that the inequality in individual annual total income (measured by the Gini coefficient) was flat over the period.

male was worse off in real income terms in 2006 than his same-aged counterpart in 1981.[9] In contrast, the median total income of working-age females increased at every census except for a dip in 1986, partly reflecting increasing participation rates.[10] Thus, the median working-age female was better off in real income terms in 2006 than in 1981 (by between 35 and 69 percent for prime-age females; Coleman & McDonald, 2010).

Thus, Callister (2006, p. 91) concludes that “women have dramatically increased their real average [individual] incomes, from a low base, but this increase is not evident for men” and Coleman and McDonald (2010, p. 11) and Coleman (2006, p. 13) conclude that “the deterioration of New Zealand incomes has been most marked for middle-aged New Zealand males” for whom “the period between 1976 and 1981 was a golden era, with high wages, and low unemployment” but “the years between 1981 and 2001 were an earnings nightmare” as “[t]he farming downturn, and the shake-outs within many industrial sectors post 1984…directly contributed to declining incomes”. In his review of the literature on New Zealand’s income distribution, O’Dea (2000, p. 11) concludes that “[t]he mid 1980s represent an historical break in real income growth, especially for men”, who, up until the 1980s, “could expect to earn on average a higher real income than a person older than him…had earned at the same age” but “[t]his ceased to hold…from 1981, because of a substantial fall in labour force participation by males, which has yet to be fully recovered [by 1996]”.[11]

Using census data on ‘total personal income’ – individual annual gross income from all sources – Figure 2 plots mean total personal income, in real (inflation-adjusted) terms, for males aged 35 to 39, and females aged 40 to 44, over each census between 1976 and 2018.[12] Incomes are deflated using the consumer price index and expressed in September 2012 dollars.

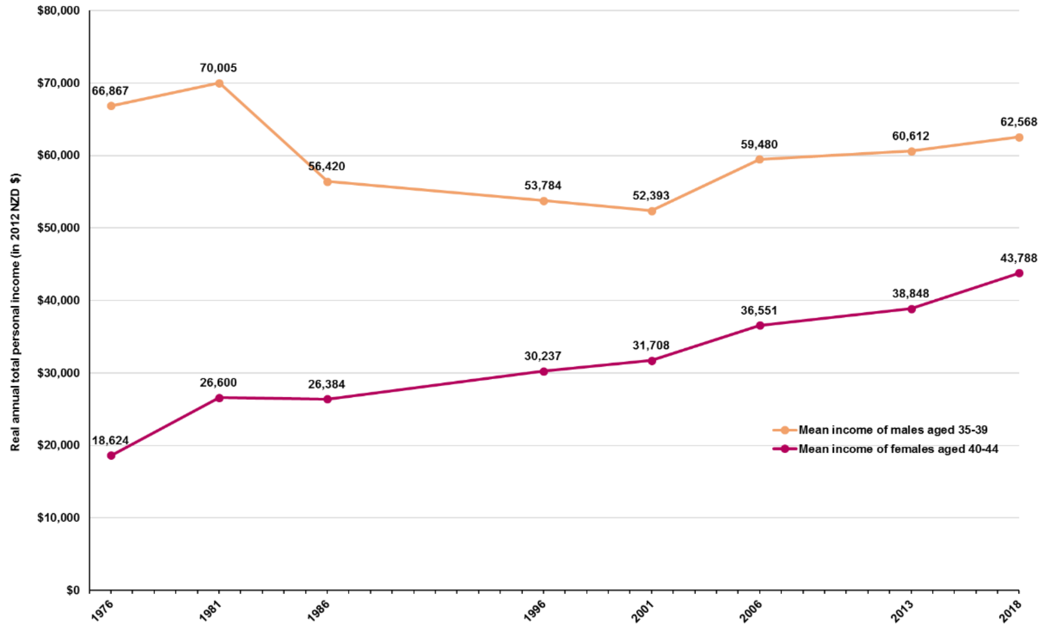
On average among males aged 35 to 39, real individual total incomes increased from 1976 to 1981, declined substantially in 1986 and continued to decline through to 2001, then increased sharply to 2006 and more modestly to 2018, when they still had not reached the levels of 1976 or 1981. On average among females aged 40 to 44, real incomes increased consistently across the full period except for 1986 when they were flat.

Thus, in 2018, the average 35 to 39-year-old man was about 10 percent worse off in real income terms than his counterpart 37 years earlier in 1981, when average male incomes peaked, while the average 40 to 44-year-old woman was about 65 percent better off in real terms than her counterpart in 1981.

---

[9] When Coleman and McDonald (2010, p. 8) disaggregate their results by income level, the real incomes of prime-age males at the 25th percentile were approximately 15 to 20 percent lower in 2006 than in 1981, while the real incomes of prime-age males at the 75th percentile were between two and 11 percent higher, so “low income males were the hardest hit group” over this 25-year period.
[10] Martin (1997b) finds the dip in female median income occurred in 1991, not 1986. All census analyses show a dip in female *mean* income occurring in 1991. Median and mean annual *market* income among females generally increased over 1982 to 1996 (mostly with no dip at any measured timepoint; O’Dea, 2000; Statistics New Zealand, 1999).
[11] In addition to rising unemployment from 1987, the declines in men’s real incomes after 1981 have been attributed to high inflation and the government’s wage and price freeze from 1982 to 1984 (Easton, 1996; Martin, 1997b).
[12] Except for the 1991 census, for which the relevant data have not been published. See the notes to Figure 2.

**Figure 2**. *Mean real total personal income of men aged 35 to 39 and women aged 40 to 44 in New Zealand, 1976 to 2018.*




**Notes:** For each census, mean total personal income is calculated using income band counts (including those with zero and negative incomes) for the usually-resident population by age and sex in conjunction with band medians estimated for the 1986 to 2013 censuses by Statistics New Zealand using more granular survey data (including estimates of 'representative values' for the top open-ended categories) which were supplied to the author by Robert Didham from Stats NZ. Band mid-points are used for the 1976 and 1981 censuses. Income counts by age and sex at the 1991 census, based on the full, original set of income bands used at that census, have not been published.

**Source**: Data for the 1976, 1981, and 1986 censuses sourced from Department of Statistics (1980b, p. 16; 1983, pp. 105-109; 1987, p. 30), for the 2001 census from Statistics New Zealand (2002, pp. 231-236), and for the 2006, 2013, and 2018 censuses from tables from Stats NZ's (2021b) NZ.Stat website.

Inequality in individual annual total income, as measured by the Gini coefficient for men and women pooled, declined from 1981 to 1986, increased substantially between 1986 and 1996,[13] and was then fairly stable through to 2013 (Easton, 2013a, 2014).[14] Mirroring this pattern, the income shares of groups at the top of the personal pre-tax income distribution (e.g., the top one percent) are estimated to have generally declined from 1977 to 1986, but then increased rapidly between 1986 and 1994, and by 2005 were about the same as they had been in 1994 (A. B. Atkinson & Leigh, 2008).[15]

Of course, these trends in aggregate income growth across generations are not informative about intergenerational mobility or persistence, since the former tells us nothing about 'which offspring go with which parents' – there may be income growth *in aggregate* over generations, but is there a tendency for high-income fathers to have high-income sons, or low-income mothers to have low-income daughters, for instance? It is to this question that this thesis turns. Chapter 2 reviews the relevant literature, with a focus on the concepts, theories, measurement, and evidence of intergenerational persistence of most relevance to the analyses conducted here. Chapter 3

[13] Martin (1997b) finds the same pattern in the Gini coefficient over 1981 to 1996 using census data on working-age males.
[14] A similar trend in the Gini coefficient is observed when estimated with tax data from Inland Revenue (Creedy, Gemmell, & Nguyen, 2018), with one exception: based on tax data, inequality in individual annual total income was relatively stable from 1977 to 1986, whereas it is decreasing from 1981 in the census data analysed by Easton (2013a). After 1986, the Gini coefficients estimated from the two series follow broadly the same pattern. See Creedy et al. (2018) for a discussion of the differences between the tax and census data series and Easton (2013a) for the change that occurred between the 1976 and 1981 censuses in the income concept collected, denoting a break in comparability pre- and post-1981.
[15] Easton (2013b, 2015a) finds no increase in top (pre-tax) income shares in the 1980s, instead finding that the increase was concentrated in the 1990 to 1994 period and remained flat thereafter to 2012. However, Easton's (2015b) estimates of *after-tax* income shares at the top follow the general trajectory over 1977 to 2005 shown in Atkinson and Leigh (2008), remaining flat between 2005 and 2010 with an uptick to 2012.

estimates intergenerational income persistence in New Zealand using data from the New Zealand Longitudinal Census, a dataset that links census microdata at the individual level over a 32-year timespan. The analysis of this data adopts a ‘permanent income’ approach to estimate the intergenerational elasticity of individual total income between son-father, son-mother, daughter-mother, and daughter-father pairs, where sons are born between 1967 and 1979, daughters born between 1967 and 1974, and a proxy for the lifetime average income of offspring and parents is constructed, taking into account well-known sources of measurement error including attenuation and lifecycle biases. Group-specific (e.g., ethnic group) elasticities are also estimated. Chapter 4 estimates intergenerational persistence using a second data source, the Christchurch Health and Development Study, which is a longitudinal study tracking a birth cohort of children born in Christchurch, New Zealand in 1977 and followed for 40 years. Analysis of this data adopts a ‘childhood resources’ approach to estimate the intergenerational elasticity and rank correlation of sons’ and daughters’ lifetime average earnings and family income with respect to their parents’ family income over the offspring’s childhood. The analyses in both chapters 3 and 4 are subjected to a range of robustness checks and results are compared cross-nationally. Chapter 5 uses both data sources to investigate mechanisms or pathways that mediate the intergenerational association in income, including the role of educational attainment and occupation (in both datasets) as well as childhood behavioural problems and cognitive ability (in the Christchurch Health and Development Study). Throughout the thesis, sources of bias and data limitations are described and results are discussed in the context of the relevant literature.

## Chapter 2: Literature review

This chapter presents a selective review of the literature on intergenerational income mobility and persistence with a focus on the aspects of the literature that are of most relevance to the analyses presented in this thesis. Section 2.1 discusses concepts of intergenerational mobility, section 2.2 discusses theories of intergenerational mobility, section 2.3 discusses the measurement of intergenerational mobility including forms of measurement error affecting estimates of mobility, and section 2.4 discusses evidence of intergenerational mobility both across countries and within New Zealand.

### 2.1 Concepts of intergenerational income mobility

Intergenerational income mobility is the movement of people up and down the income distribution relative to their parents. The study of intergenerational income *persistence* focuses on how closely people's incomes are related to those of their parents. There are many different concepts of, and dimensions to, intergenerational income mobility. The discussion below focuses on three conceptual distinctions: structural versus exchange mobility, absolute versus relative mobility, and a 'permanent income' approach versus a 'childhood resources' approach to the study of intergenerational mobility.[16]

#### *2.1.1 Structural versus exchange mobility*

The joint distribution of parents' and offspring's incomes can be separated into two components: the joint distribution of parents' and offspring's ranks, formally known as the *copula* of the distribution, and the *marginal distributions* of parent's and offspring's incomes. *Structural mobility* is a scalar measure of changes in the marginal distributions across generations (such as income growth and changes in cross-sectional inequality) while *exchange mobility* is a scale-invariant measure of changes in the copula (re-ranking of individuals across generations). Structural mobility is derived by aggregating across every individual in a society the change in income amount experienced by each individual relative to their parents' income, where the individual-level changes might be gains (upward structural mobility) or losses (downward structural mobility) (Jäntti & Jenkins, 2015). As improvements in productivity over time generate economic growth and wage rises, children tend, on average, to be better off than their parents, even if there is little or no movement between the ranks of the rich and poor in society. On the other hand, e*xchange mobility* focuses on *positional change* abstracting from changes in the marginal distributions. Differences in the marginal distributions in each generation are standardised (to a common uniform distribution) and mobility is measured as the change in fractional rank that occurs between parent position and offspring position. Whereas with structural mobility it is possible for everyone to be upwardly (or downwardly) mobile, this is not so with exchange mobility; if one person moves up a rank, another person must, by definition, move down (so in aggregate,

[16] There are other dimensions to intergenerational income mobility not discussed here. For example, Deutscher and Mazumder (2021) classify 19 different measures of intergenerational mobility used in the literature by four key properties: (1) whether they focus on the joint distribution of income or consider other family background factors as well; (2) whether they are absolute or relative measures of mobility; (3) the units in which income is measured (income level, log income, ranks, or fixed ranks), and; (4) whether they are 'global' measures summarising mobility over the entire distribution, or whether they are 'local' measures focused on specific points of the distribution.

changes in ranks cancel each other out). For example, if every offspring is $10,000 better-off in real terms than their parents across the entire parental income distribution, then there has been structural mobility, but no exchange mobility since every offspring occupies the same position in the distribution as their parent. In other words, individual income growth has shifted the marginal distribution of income to the right, but ranks are preserved across generations.

### *2.1.2 Absolute versus relative mobility*

Mobility estimators can be classified into those measuring *absolute mobility* (for these measures, any increase in incomes in the offspring generation unambiguously increases mobility, regardless of who receives it) and those measuring *relative mobility* (for these measures, an increase in incomes in the offspring generation has an ambiguous effect on mobility – whether mobility is increased or decreased depends on which particular offspring received the income increase). In practice this means that absolute measures are insensitive, and relative measures are sensitive, to *who benefits* from income gains across generations. The primary measure of intergenerational persistence estimated in this thesis is the intergenerational income elasticity (IGE), defined in section 2.3.1 below, which is considered a measure of relative mobility because an income increase in the offspring generation could increase or decrease the elasticity, depending upon the recipient; if individual income growth accrues to particular offspring in such a way that offspring incomes become more closely related to parents' incomes, in aggregate, then the elasticity will increase. On the other hand, a doubling of income of *each and every individual* in the offspring generation will leave the elasticity unchanged (Deutscher & Mazumder, 2021).

### *2.1.3 Association of permanent incomes versus the effect of childhood resources*

The literature on intergenerational income persistence has traditionally focused on the association between the permanent incomes of parents and offspring, but recent studies have taken a family-based focus on the association between parental income during the offspring's childhood and offspring's earnings in adulthood. As Gregg, Macmillan and Vittori (2017, p. 80) explain:

> There are two slightly different focuses here. One is to consider intergenerational mobility as the differences between lifetime incomes or earnings across two complete generations of parents and children. The other assesses the extent to which a child's adult outcomes mirror their childhood circumstances. In this second setting, ideally the degree of intergenerational mobility would be measured as the association between the socio-economic status of parents throughout a person's childhood and their lifetime earnings as an adult. Both approaches are very data-intensive but the second less so.

Thus, the 'permanent income approach' and the 'childhood resources approach' differ in the 'who, what, and when' of intergenerational mobility. The 'who' of mobility refers to the definition of the income-receiving unit being analysed, the 'what' refers to the definition of income (the income concept) and is closely related to the definition of 'who', and the 'when' of mobility refers to the length of time, ages, and periods over which incomes are studied (Jäntti & Jenkins, 2015).[17] The

[17] In the literature, definitions of income (mobility of what) vary based on whether they are market, gross, or disposable income (pre-tax pre-transfer, pre-tax post-transfer, and post-tax post-transfer, respectively) and whether they are measured at the individual, family or household level (and if household income is used, whether to apply an equivalence scale). The income-receiving unit (mobility among whom) is often defined as workers, but in principle all individuals, including stay-at-home parents and retirees, can be included in an analysis of intergenerational mobility if they are assumed to receive the equivalised total income of the family or household to which they belong. The 'whom' has also traditionally been defined as men (sons and fathers), but mothers and daughters are increasingly being included in analyses (e.g., Chadwick & Solon, 2002; Chetty, Hendren, Kline, & Saez, 2014; Hirvonen, 2008). In the literature it is

permanent income approach is concerned with the income of *an individual parent*, so changes in father and mother figures (e.g., due to re-partnering) are relevant since permanent income is an individual-specific concept that cannot be computed over different parental figures. The childhood resources approach is concerned with the total financial resources available to the offspring during their childhood, so the incomes of *both parents* matter (or indeed all family members) and changes in parents/caregivers are not relevant per se since the interest is in the resources available to the child *regardless of who earned/received them*.

The permanent income approach is interested in the *parent's lifetime income* whereas the childhood resources approach is only interested in *parental income received during the offspring's childhood*. Consequently, in the absence of lifetime income data, the *age of parents at observation* is relevant to the former approach since the computation of permanent income is age-dependent. In contrast, the age of parents is not relevant per se to the second approach because, again, the interest is in what financial resources were available to the offspring *regardless of how old their parents were at the time*.

The permanent income approach is partly an outgrowth of the standard theoretical models of intergenerational transmission which adopt a single-parent/single-child framework (see section 2.2 below).[18] The approach may be particularly suitable when the interest is in the intergenerational transmission of latent earnings potential. On the other hand, Mazumder (2005, p. 250) adopts the childhood resources approach because "children's earnings as an outcome…provides a valid reading of children's economic opportunity as opposed to directly inherited economic status" since "earning capacity (skills, effort, and the like) cannot be transferred from parents to children, say, the way a house or a financial asset can", while parental family income is "a potentially less error-prone measure of parent permanent status", so although "the elasticity of family income on children's earnings is not a measure that comes out of the standard theoretical models, it may still shed light on the magnitude of the effect of family resources on children's future earning opportunities, which is presumably a question of great interest".[19]

As previously discussed in chapter 1, Björklund and Jäntti (2020) outline a taxonomy of approaches to the study of intergenerational transmission. They distinguish between descriptive accounts of intergenerational associations which aim to estimate the degree of mobility/persistence versus an 'intergenerational effects' approach which aims to identify causal

common to define parents as those with whom the offspring was living at a particular age in childhood or to birth parents (Jäntti & Jenkins, 2015), but analysts frequently face the issue of how to define the parent generation when offspring come from single-parent families or have step-, adoptive, or foster parents. Using Swedish administrative data that allows the identification of both biological and nonbiological parents, Björklund and Chadwick (2003, p. 245) show that "the definition of the father used in measuring intergenerational income elasticities has the potential to substantially affect the results". In answering the 'mobility when' question, the main issues arising are transitory income fluctuations and age-income profiles that differ by parental-income background which produce attenuation and lifecycle biases, discussed in section 2.3.4 below. In answering the questions of mobility what, among whom, and when, researchers' choices ultimately depend on the goals of their analysis; however, in practice these choices are often constrained by the sources of longitudinal data that are available to the researcher (Jäntti & Jenkins, 2015).

[18] Carneiro, García, Salvanes and Tominey (2021, p. 6 of appendix) also note that it is "standard…to construct lifetime permanent income across the life cycle of an individual and not across the age of children" as per the permanent income approach.

[19] The childhood resources approach is also adopted by Gregg, Macmillan et al. (2017) and Gregg, Jonsson, Macmillan and Mood (2017). Mazumder (2005) and Mitnik, Bryant, Weber and Grusky (2015) argue that using parental family income, rather than an individual parent's earnings as is common in the literature, has a number of other conceptual and empirical advantages. These are discussed in section 4.1.2 of chapter 4.

effects of family background on offspring outcomes. While the 'permanent income' and 'childhood resources' approaches outlined above do not map directly to this distinction (as both are associational), the childhood resources approach is implicitly motivated by an interest in what *effect* resources in childhood have on offspring's adult outcomes, and would therefore fall within the 'intergenerational effects' approach under Björklund and Jäntti's (2020) taxonomy.[20] Ultimately, the choice of approach depends on the analyst's research question and mobility concept of interest. However, in practice the choice is often driven by data availability; it is frequently the case that the available measures of income are family income in the parent generation and individual earnings in the offspring generation, as in some of the British birth cohort studies (Jäntti & Jenkins, 2015). Such practical data-driven considerations have guided the choice of approach taken with the two datasets used in this thesis. The permanent income approach is used in the analysis of data from the NZLC in chapter 3, while the childhood resources approach is used in the analysis of data from the CHDS in chapter 4. The reasons for these choices are discussed in the relevant chapters.

## 2.2 Theories of intergenerational income mobility

### *2.2.1 The Becker-Tomes model*

Why should we expect to observe any link at all between parental income and offspring income? Becker and Tomes (1979, 1986) developed a theoretical model of the transmission of earnings, assets, and consumption from parents to children. Grounded in human capital theory, the model is based on utility-maximising behaviour by parents who are concerned about the welfare of their children which they influence through 'endowments' passed on to them and investments made in them.[21] The model makes the simplifying assumptions of one period of childhood, one period of adulthood, one child per family, and one parent per family. The parent transmits endowments to the child, subject to their degree of heritability, in an exogenous or 'mechanical' way via both genetic and environmental transmission. Endowments are defined broadly as both genetically-determined ability and the knowledge, skills, goals, social connections, and other cultural endowments acquired through the family environment. But "economic forces are at work as well", as the parent decides how much of their resources to invest in the human capital of their child, how much to allocate to financial transfers to their child, and how much to allocate to their own consumption, subject to the constraints of their budget and their allocation preferences or degree of altruism towards their child (Mulligan, 1999, p. S188). By assumption, parental investment in the child's human capital is positively related to the child's inherited ability, that is, children with higher innate ability receive larger investments. When the child grows up, their own earnings will depend on that investment, their inherited endowments, and their luck in the labour market.

[20] The childhood resources approach is a sort of 'half-way' amalgam of the intergenerational association and intergenerational effects literatures, since it is implicitly motivated by an interest in what effect family resources have on offspring earnings, yet it does not try to estimate causal effects, instead being simply another approach to describing or quantifying mobility.

[21] It is this explicit focus on optimising behaviour which sets the Becker-Tomes model apart from contemporaneous empirical and statistical work which omitted this as a central tenet and from a Galton-type heredity model which is non-economic (Mogstad, 2017; Mulligan, 1999).

With perfect credit markets, parents are not constrained from borrowing against the earnings potential of their child, so parental income and preferences play no role in determining investment in children's human capital. As Mogstad (2017, p. 1864) puts it, "[n]o matter their income, parents can borrow freely in the market to finance the optimal investment level", so parents invest in their child's human capital – up to the limits set by the child's endowments – until the marginal return diminishes to the opportunity cost of investments (foregone interest on financial investments). This is the privately (Pareto) efficient amount.[22] As a result, equally able children receive equal investments regardless of their parent's income or human capital, and intergenerational earnings persistence is driven by the heritability of endowments and efficient human capital investments. But with imperfect credit markets, parents are prevented from borrowing against their child's earnings potential to fund the optimal level of investment. Hence, children of credit-constrained parents acquire less human capital than children with similar innate abilities but wealthier parents. In this case, intergenerational earnings persistence is driven by both the persistence of inherited endowments and borrowing constraints that prevent profitable investments in human capital from being made.[23]

Solon (2004) modifies the Becker-Tomes model to understand variation in intergenerational mobility over time and place. He augments the Becker-Tomes model by allowing for government investment in the child's human capital, which can vary in its progressivity (the extent to which there is a relatively greater benefit to the least well-off families, measured as the ratio of government investment to parental investment which decreases in parental income). The model shows that intergenerational persistence increases with the heritability of endowments, the productivity of human capital investment, and the earnings return to human capital, and decreases with the progressivity of government investment.[24] Regarding applications of the model to cross-country comparisons, Black and Devereux (2011, p. 1500) observe that "[g]iven that the heritability of endowments is unlikely to differ significantly across developed countries or over time (and differences in the efficiency of the education system are hard to measure), explanations for differences across countries tend to focus on differences in the returns to skills (primarily education) and differences in government investments". Hence, countries with lower estimated intergenerational persistence may best be explained "either by their compressed earnings distributions (low returns to skills) or by social and educational policies regarding childcare and

[22] However, as Caucutt, Lochner and Park (2017, p. 104) point out, "it might be socially desirable to encourage investment beyond the privately optimal amount due to human capital externalities in production…or related to crime…or citizenship".
[23] This intergenerational borrowing constraint, introduced in the 1986 version of the Becker-Tomes model, implies a concave relationship between parents' and children's earnings, since lack of access to credit for lower-income families acts as a brake on the mobility of their children, so the correlation between parents' and children's earnings is steeper in the lower part of the parental income distribution but flatter in the upper part of the distribution as credit becomes more available thus weakening the influence of parents' earnings on children's earnings. However, higher-income families may also be constrained if their optimal level of investment is high due, for example, to having a gifted child, so being credit-constrained is not synonymous with low family income or low investment levels (Heckman & Mosso, 2014).
[24] In Solon's (2004) model, the earnings return to human capital is an indicator of inequality in the labour market and therefore countries with labour markets characterised by more cross-sectional inequality – that is, a higher rate of return to education – will exhibit less intergenerational mobility, driven by higher-income families not only having more resources to invest in their child's human capital but greater incentives (a higher payoff) to do so. More generally, in Solon's (2004) model, "the factors that drive intergenerational persistence, such as the heritability of human capital endowments, the returns to education, and the progressivity of public education expenditure, affect cross-sectional inequality with the same signs" (Jäntti & Jenkins, 2015, p. 889). The model therefore predicts a positive association at the country level between income inequality and intergenerational income persistence, a relationship later termed 'the Great Gatsby curve' (see section 2.4.1).

education that tend to equalize educational opportunities for children" (Black & Devereux, 2011, p. 1500).

### *2.2.2 Critiques of the Becker-Tomes model*

More recent theoretical developments have examined some of the assumptions underlying the Becker-Tomes model, in particular that investments at any stage of childhood are equally effective, that there is a sharp distinction between innate ability and acquired human capital and these coalesce into a single cognitive skill that determines earnings, and that parental investments are of a single monetary kind (Heckman & Mosso, 2014).[25] The Becker-Tomes model collapses childhood into a single period thereby assuming that investments at all child ages are perfect substitutes and that parental investments are made on the basis of permanent income so the timing of receipt of income over the lifecycle is unimportant. But Cunha, Heckman, Lochner and Masterov (2006) argue that the timing of income is crucial because the technology of skill formation is a multistage one encompassing sensitive and critical periods (skills can be formed more productively in some periods than others or their formation is uniquely suited to a particular period) and is characterised by *dynamic complementarity* (skills acquired in one period build upon skills attained earlier and also enhance the effectiveness of later learning).[26] These features result in a 'multiplier effect' in which skills beget skills, so that returns to investing early in the life cycle are high because they raise the payoff to future investments, while remediation of inadequate early investments with compensatory investments at later ages is difficult and costly and full remediation may not be possible.

Parents may not only be restricted from borrowing against the earnings potential of their children (the intergenerational credit constraint identified in Becker-Tomes) but also prevented from

[25] Becker, Kominers, Murphy and Spenkuch (2018) also extend the Becker-Tomes model, notably by abandoning the implicit assumption that parents are equally good at investing in their children. Instead, Becker et al. (2018) assume that parents' human capital makes them not only more productive in the labour market (thus having higher earnings) but also in the household with respect to developing their child's human capital (by investing more time, choosing more effective inputs, helping with school work, and other inputs into skill development). As a consequence of this complementarity between parents' human capital and investments in their children, intergenerational earnings persistence will occur even in the absence of credit constraints and differences in genetic endowments. As Becker et al. (2018, p. S16) state, "the earnings of parents and children are directly related through the intergenerational transmission of human capital, even when capital markets are perfect and there are no differences in innate ability", a result which "differs notably from that of Becker and Tomes (1986), who ignore complementarities in production". Heckman and Mosso (2014, p. 702, italics in original) also make the point that the productivity of investments is higher for children with higher-skilled parents, which is "a type of market failure due to the *accident of birth* that includes a correlation of human capital and earnings across generations, even in the absence of financial market imperfections" and, furthermore, the child's genetically-endowed abilities also affects investments, creating "a second channel of intergenerational dependence due to the *accident of birth* if it is genetically related to parental endowments, as considerable evidence suggests".

[26] Relatedly, some economists have questioned the near exclusive focus on permanent income as the most appropriate variable for estimating intergenerational income mobility, a focus which rests on two assumptions: firstly, that long-run income is the most relevant measure of economic well-being, and secondly, that deviations of short-run from long-run income are treated as classical errors, being uncorrelated with long-run income and having a constant variance (Jäntti & Jenkins, 2015; Jäntti & Lindahl, 2012). In the permanent income model, if transitory shocks are largely classical, and "if capital markets are well functioning and individuals have a fair idea of what their permanent income is (so they know if they have been hit by a negative or positive shock)", then people smooth their consumption by saving and borrowing (Jäntti & Jenkins, 2015, p. 896). According to Jäntti and Jenkins (2015, p. 896), a focus on permanent income is justified only under these "demanding conditions". Jäntti and Lindahl (2012, p. 165) question all these premises, stating "[l]ong-run income is most welfare relevant only if individuals have access to well-functioning capital markets and can distinguish between transitory and permanent shocks to income, which is not very realistic" and "[t]he assumption that short-run deviations are classical has been falsified" so "[t]he focus on long-run economic status is unwarranted and obscures other transmission mechanisms". On the assumption that individuals prefer a stable income to a fluctuating one, then, other things being equal, welfare is reduced by short-run income fluctuations (if the fluctuations are known in advance, but the welfare losses are even greater in the more realistic scenario in which individuals cannot predict these fluctuations, at least for risk-averse individuals; see Jenkins and Van Kerm (2009)). Thus, it is not just permanent income but also *income volatility* that matters for economic well-being and its transmission across generations, and there is evidence that income volatility is associated with the level of long-run income and is correlated across generations (Jäntti & Jenkins, 2015).

borrowing fully against their own future earnings (an intragenerational credit constraint). Owing to dynamic complementarity in skill formation, credit constraints that bind during the child's early years will result in lower investments in later years, even if the parent is not constrained in later periods, and may be particularly harmful to children if they bind during a critical period of development (Caucutt et al., 2017; Heckman & Mosso, 2014).[27] In the models of intergenerational economic transmission by both Lee and Seshadri (2019) and Caucutt and Lochner (2020), it is this coupling of intragenerational borrowing constraints with dynamic complementarity (alongside inheritance of innate ability) that drives intergenerational earnings persistence. Lee and Seshadri's model (2019, pp. 901, 855) shows that "the earliest stage of investment is the most important when the childhood human capital accumulation displays dynamic complementarity" but "it is also when parents are the most financially constrained" thereby preventing optimal investment in their children, so "borrowing constraints faced by young parents are important for understanding the persistence of economic status across generations".[28] Caucutt and Lochner's (2020, p. 1068) model shows that the timing of borrowing constraints are crucial when human capital investments display dynamic complementarity, because "[w]hen investments are sufficiently complementary, a policy that encourages investment at one stage of development will tend to increase investment at other stages", which "can present a challenge for families that are severely constrained when their children are young" who "may be unable to take advantage of college-age subsidies or loans, because the inability to invest early may render late investments unproductive". In Caucutt and Lochner's model, it is the presence of such borrowing constraints across all stages of childhood, in conjunction with other market frictions such as uninsured labour market risk as well as intergenerational transmission of ability, that generate gaps in child investment by parental income and ultimately persistence of earnings across generations.

Recent research has found that noncognitive skills such as perseverance, dependability, patience, and conscientiousness are just as important as cognitive skills in determining what people earn, that noncognitive skills enhance cognitive skills, and that cognitive ability, while it has a genetic component, is not fixed from birth but is "susceptible to environmental influences, including in utero experiences", and the same applies to noncognitive skills (Cunha et al., 2006; Heckman & Mosso, 2014; Heckman, Stixrud, & Urzua, 2006). These results suggest, *contra* Becker-Tomes, that parental investments of time and resources can enhance the cognitive skills and noncognitive skills of their children and that both play important roles in intergenerational persistence.

In the Becker-Tomes model, parental investments are conceptualised as "educational goods analogous to firm investments in capital equipment", ignoring parental time investments which have been shown to be "a prime factor influencing child skill formation" reflecting the notion that "[q]uality parenting is a time-intensive process" involving interactions which guide or scaffold the

[27] Adding uncertainty in income growth to the model compounds the suboptimality of investments, as parents (even currently unconstrained ones) reduce current investments in their children in order to accumulate a buffer against future income shocks (Heckman & Mosso, 2014).

[28] The Becker-Tomes model also assumes exogenous fertility, yet parents with a high preference for optimal child development may delay having children until such a time as they can make greater investments of time and money, so credit constraints may be less important to human capital investments than implied by models that take fertility as exogenously determined (Heckman & Mosso, 2014). The Becker-Tomes model also assumes exogenous mating decisions, so the role of marital sorting in intergenerational persistence is ignored.

child's learning within an optimal zone tailored to their individual development, enabling parents to "more accurately assess the capacities of their children and to make more precisely targeted investment decisions" (Heckman & Mosso, 2014, pp. 690, 704). For this reason, Lee and Seshadri (2019) include investments of both goods and time in their theoretical model.[29]

In sum, the Becker-Tomes model "remains the main building block of economic research on intergenerational mobility" but researchers continue to refine the theory in light of new evidence and test revised theories with contemporary datasets (Mogstad, 2017, p. 1863).

### 2.3 Measurement of intergenerational income mobility

This section discusses the main methods used to measure or estimate intergenerational mobility and the effects that different forms of measurement error can have on such estimates.

#### *2.3.1 The intergenerational income elasticity*

The standard approach to measuring intergenerational income persistence is to apply ordinary least squares to equation (1):

$$Y_{ij}^{\text{offspring}} = \alpha + \beta Y_i^{\text{parent}} + \varepsilon_i \tag{1}$$

where $Y_{ij}^{\text{offspring}}$ is the natural logarithm of lifetime average income[30] of offspring $j$ in family $i$, $Y_i^{\text{parent}}$ is the natural logarithm of lifetime average income of the parent in family $i$, and $\varepsilon_i$ is a random error term capturing factors uncorrelated with $Y_i^{\text{parent}}$. The slope coefficient $\beta$ is the intergenerational income elasticity (the 'IGE'), a commonly-used descriptive statistic summarising the association between offspring's and parents' incomes. The IGE measures intergenerational income *persistence*, which is the inverse of mobility (thus, 1 minus the IGE measures mobility or the rate of regression to the mean). If the IGE takes a positive value, higher parental incomes are associated with higher offspring incomes, and there is a degree of intergenerational persistence.[31] A high IGE implies high persistence of incomes, or low mobility across generations, while a low IGE implies low persistence of incomes, or high mobility across generations. For example, an IGE of 0.4 means that, on average, for an income difference between two families of 1 log point in the parent generation, we would expect the income gap in the offspring's generation to be 0.4 log points. Because incomes are measured in logs, the difference in log income approximates the percentage difference in income. Thus, the IGE denotes the percentage change in offspring's long-run income associated with a marginal percentage change in long-run parental income and can be depicted visually as the slope of the line that best fits the data when plotting offspring's log income against parents' log income (Deutscher & Mazumder, 2021).

[29] See Piketty (2000) for a discussion of other theories of intergenerational persistence related to taste differences between status-based reference groups, residential segregation by socioeconomic status, employer discrimination, and social class differences in self-fulfilling beliefs.
[30] The term 'lifetime average income', used extensively in the literature, refers to *annuitised* lifetime income. Alternative concepts of lifetime income, such as the present discounted value of the sum of inflation-adjusted income over the lifecourse, are briefly discussed in Mitnik and Grusky (2020, p. 88).
[31] If the IGE takes a negative value – as some estimates in this thesis do – then increases in parental income are associated with *decreases* in offspring's income.

Empirically, the IGE typically lies between zero and 1, where zero means there is no relationship between parents' and offspring's incomes, and 1 means there is a perfect dependence between the two. Thus, in a society where parents' incomes had no influence on their offspring's incomes, the IGE would be at or near zero. Equation (1) is a reduced form statistical model that describes the total relationship between parents' and offspring's long-run incomes. The IGE is not given a causal interpretation but rather reflects the influence on offspring income of all factors correlated with parent income. The IGE is thus "a 'catch-all' measure of the intergenerational association that encompasses a variety of transmission mechanisms, including ability and genetic transmission, socialization and preference formation, as well as economic resource constraints" (Lefranc, 2018, p. 811).

However, covariates for the age of offspring and parents are commonly added to equation (1). This is done in the analysis of NZLC data in chapter 3 but is unnecessary in the analysis of CHDS data in chapter 4 as all offspring were born in a four-month window in 1977 and parents' age is not relevant to the childhood resources approach.[32]

As explained above, the joint distribution of parent's and offspring's incomes can be separated into the copula (the joint distribution of ranks) and the marginal distributions. The IGE does not involve standardising the marginal distributions of income in each generation, so it does not abstract from changes in these distributions across generations (e.g., income growth, changes in inequality). Thus, the IGE depends on both the copula and the marginal distributions, and therefore captures both the extent of re-ranking across generations and the spread of the income distributions (Gregg, Macmillan, et al., 2017). In other words, it makes no distinction between structural and exchange mobility.

*2.3.2 Mitnik's intergenerational elasticity of expected income*

Mitnik and Grusky (2020) point out that the IGE has long been assumed to refer to the expectation of offspring's income when in fact it pertains to the geometric mean of offspring's income, resulting in the IGE being widely misinterpreted. They propose an alternative estimand – the IGE of the expectation – which, like the conventional IGE, also measures intergenerational income persistence. They argue that this new IGE restores the conventional interpretations wrongly attached to, and overcomes the conceptual and methodological weaknesses of, the IGE of the geometric mean. The conventional IGE is:

$$\beta_{\text{IGE of geometric mean}} = \frac{d\,E\left(\ln Y^{\text{offspring}} \middle| Y^{\text{parent}} = y^{\text{parent}}\right)}{d \ln y^{\text{parent}}} \qquad (2)$$

But this refers to the elasticity of the conditional geometric mean of offspring's income (the percentage differential in the geometric mean of offspring's long-run income with respect to a marginal percentage differential in parents' long-run income), not the elasticity of the conditional

[32] Mitnik et al. (2015) prefer to omit controls for parental age on the grounds that the age at which parents have their children is not exogenous to their income (as high-income parents are more likely to delay childrearing) and parental age may affect children's life chances, and therefore controlling for parental age is inconsistent with the objective of measuring the gross association between parents' and offspring's incomes. Consequently, a robustness check omits the age variables from the estimation of the IGE with NZLC data.

expectation of offspring's income, which is what mobility analysts had assumed they were measuring (see Mitnik and Grusky (2020, pp. 54-55) for evidence). Mitnik and Grusky (2020, p. 56) can find "no statement to the effect that the conventional approach to estimating the IGE has been undertaken because of some genuine interest in recovering the elasticity of the conditional geometric mean of long-run income or earnings". In short, the conventional IGE is "the wrong estimand" given the interpretations usually applied to it (Mitnik et al., 2015, p. 22). The IGE of the expectation is:

$$\beta_{\text{IGE of expectation}} = \frac{d \ln E\left(Y^{\text{offspring}} \middle| Y^{\text{parent}} = y^{\text{parent}}\right)}{d \ln y^{\text{parent}}} \quad (3)$$

That is, the percentage differential in the expectation of offspring's long-run income with respect to a marginal percentage differential in parents' long-run income,[33] which is consistent with the intended definition of the IGE.

Furthermore, Mitnik and Grusky (2020) and Mitnik et al. (2015) show that their alternative IGE has several advantages over the conventional IGE. Most notably, because offspring's income, rather than their log income, is used as the dependent variable, the IGE of the expectation allows for the retention of offspring with zero income in the sample. As the logarithm of zero is undefined, these offspring are often excluded when estimating the conventional IGE. Thus, the IGE of the expectation avoids the selection bias that arises from dropping offspring with zero income who often have weaker labour force attachment.[34] Mitnik and Grusky (2020, p. 49) argue that this "zeros problem" is not a negligible issue because the short-run income measures commonly available in surveys "typically have a substantial probability mass at zero", particularly if the analysis is of individual earnings rather than family income. Even long-run income measures may equal zero in the case of women's earnings (since "for a great many women, lifetime earnings *are* zero") and in the case of offspring who are "institutionalized (e.g., incarcerated, hospitalized) over their entire life or are disabled and dependent on their parents or other relatives over their entire life" (Mitnik & Grusky, 2020, pp. 66, 52).

Does imputing zero incomes with a small positive value side-step any selection bias? Mitnik and Grusky (2020) contend that this is an unsatisfactory solution because estimates of the conventional IGE have been shown empirically to be highly sensitive to the particular imputed value used (e.g., $1 or $1,000). Estimates "vary wildly across different imputations" because "the IGE of the geometric mean is dominated by small absolute differences at the lowest quantiles of

[33] Or more succinctly, "the share of percent differences in parents' income found between the expected incomes of their children" (Mitnik & Grusky, 2020, p. 71).

[34] There is selection bias because "the probability of having zero earnings or income [in] any year is negatively correlated with parental income" and "those who have zero earnings or income in the measurement year tend to have lower long-run income or earnings than do others with the same parental income" (Mitnik & Grusky, 2020, p. 63). Mitnik and Grusky (2020) demonstrate that intergenerational persistence is substantially higher when estimated with their new estimand than with the conventional estimand, implying that the exclusion of offspring with zero incomes does indeed generate a selection bias. They estimate the elasticity of offspring's earnings with respect to parents' family disposable income for son-parent and daughter-parent pairs in the US using three different data sources, the Statistics of Income Mobility (SOI-M) Panel, the National Longitudinal Survey of Youth (NLSY79), and the Panel Study of Income Dynamics (PSID). For each data source, the elasticity is estimated with both estimands. A comparison of the estimates reveals that those of the IGE of the expectation are between 11 and 42 percent larger among sons, and between 12 and 63 percent larger among daughters, compared to estimates of the IGE of the geometric mean. In other words, mobility is overstated when offspring with zero earnings or income are excluded from the analysis or included with incomes imputed at low values.

the child's conditional distributions" (Mitnik & Grusky, 2020, p. 79).[35] What about using multiyear averages of offspring income to reduce or eliminate the number of zero incomes? Mitnik and Grusky (2020, p. 65) concede that this is "attractive in principle" but argue that it is problematic in practice because many countries lack such multiyear data and because it would impose "a much longer "waiting time" before the IGE for any given cohort can be estimated".

Given these concerns with the conventional IGE, all analyses in this thesis estimate Mitnik and Grusky's IGE of expected income alongside the conventional IGE, using their preferred Poisson pseudo-maximum likelihood (PPML) estimator (Mitnik & Grusky, 2020). Multiyear averages of offspring's income are also used wherever possible.

### *2.3.3 Other measures of intergenerational mobility*

Other measures of intergenerational mobility include the Pearson linear correlation *r*, which is equal to the IGE multiplied by the ratio of the standard deviations of parent and offspring permanent incomes, thereby abstracting from changes in cross-sectional inequality in the marginal distributions:

$$r = \mathrm{Corr}_{Y_i^{\text{parent}}, Y_i^{\text{offspring}}} = \beta \left( \frac{\mathrm{SD}^{Y_i^{\text{parent}}}}{\mathrm{SD}^{Y_i^{\text{offspring}}}} \right) \tag{4}$$

The Pearson intergenerational correlation only partially standardises the marginal distributions: changes in only one distributional feature (inequality) are controlled for, and then only using one particular measure of inequality (Jäntti & Jenkins, 2015). Hence, as with IGE, the intergenerational correlation depends on both the copula and the marginal income distributions and thus does not (fully) distinguish between structural and exchange mobility. Alternatively, rank-based measures of mobility abstract from *any* distributional differences between generations. Rank-based measures also quantify aspects of mobility not captured by a single summary statistic like the IGE, such as differentiating upward and downward movements and identifying nonlinearities in the relationship between parents' and offspring's incomes (Corak, Lindquist, & Mazumder, 2014). One such measure is Spearman's rank correlation, which measures the association between parents' *position* and offspring's *position* in their respective income distributions with equation (5):

$$R_{ij}^{\text{offspring}} = \alpha + \rho R_i^{\text{parent}} + \varepsilon_i \tag{5}$$

where $R_{ij}^{\text{offspring}}$ and $R_i^{\text{parent}}$ represent offspring and parent ranks (usually percentiles) in their respective income distributions. The slope coefficient *ρ* is the rank correlation, also known as the 'rank-rank slope'. The regression coefficient equals the correlation coefficient because both offspring and parent ranks follow a uniform distribution by construction. A high rank correlation implies high persistence of position in the income distribution across generations, while a low rank correlation implies low persistence of income positions. For example, a rank correlation of 0.4

[35] Mitnik and Grusky (2020, p. 10 of appendix) explain with this example: "For a low quantile, if the child's incomes corresponding to two different parental-income values are $100 and $200, the proportional difference is 100 percent. By contrast, the corresponding child's incomes for a high quantile might be $90,000 and $120,000 (for the same parental values), which counts as a proportional difference of only 30 percent even though the absolute difference is 300 times larger".

means that a 10-percentile difference in ranks in the parent generation is associated with an expected four percentile difference in the offspring generation. Because the rank correlation standardises the marginal distributions across generations, it depends only on the copula of the joint distribution and therefore is a measure of pure positional or exchange mobility.

Another rank-based approach to measuring intergenerational mobility is with a transition matrix. This divides each generation's income distribution into equal-sized quantiles or bins and cross-tabulates them, generating a set of transition probabilities of intergenerational movement across these quantiles. Often researchers have an interest in specific elements of the transition matrix such as the probability of rising from the bottom to the top of the distribution ('rags to riches' upward mobility) or persistence at the very top or bottom ('inheritance of top incomes' and 'poverty traps'). Transition matrices can also be collapsed into a single summary index such as the immobility ratio which quantifies how much clustering is on the leading diagonal (that is, the degree of intergenerational persistence). In the case of complete immobility, the immobility ratio equals 100 percent; if there is complete 'origin independence' (offspring are equally likely to end up in each quantile regardless of their parents' quantile), the immobility ratio equals 20 percent.[36]

*2.3.4 Measurement error affecting estimation of intergenerational mobility*

Datasets containing lifetime or long-run income histories of both parents and their offspring are rare. Commonly, only short-run income data such as annual income observed in a single year or over a few years is available to the researcher to proxy for unobserved lifetime average income. This makes accurately estimating intergenerational income mobility a challenging task. Using current snapshots of income to approximate lifetime values introduces measurement error that can lead to severe biases in estimating intergenerational mobility. In the case of the IGE, estimates are sensitive to the *length of time* and the *specific ages* over which income is observed. Following from the classical linear regression mismeasurement model (Hausman, 2001), the combination of transitory fluctuations and measurement error (collectively, 'noise') in short-run measures of parental income produces an errors-in-variables *attenuation bias*, which biases the IGE estimate downwards (Mazumder, 2005; Solon, 1992). Averaging over more parental income observations smooths out transitory variation and measurement error, decreasing the noise component and increasing estimates of the IGE.[37] Attenuation bias will be further aggravated by the use of homogenous or unrepresentative samples in which the sample variance of parental income is less than the population variance. In such a sample, the 'signal' – variance in long-run parental income – is reduced, which will lead to a downward bias in the IGE if there is no corresponding reduction in noise (Solon, 1989, 1992). Consequently, intergenerational mobility will be overstated if parental incomes are measured over too short a period or among an unrepresentative sample (and likely severely attenuated if both apply).

[36] For a discussion of other summary statistics of mobility that can be derived from transition matrices, such as the normalized trace index, see Jäntti and Jenkins (2015).

[37] An alternative approach to overcoming attenuation bias is to use instrumental variables (IV) to predict parents' lifetime average income based on other parental characteristics. However, this tends to lead to an upward bias in IGE estimates since almost every variable that is correlated with parents' lifetime average income (hence a candidate for an instrument, like parents' education or occupation) might also have an independent effect on offspring's income. IV estimates therefore "provide an upper bound on the true extent of intergenerational transmission" (Blanden, 2013, p. 41).

Over what length of time does parental income need to be averaged? Mazumder (2005) performs simulations with US social security data and finds that even a five-year average produces substantial bias in estimates of the IGE because transitory fluctuations persist (are serially correlated) over subsequent years. Gregg, Jonsson, et al. (2017, p. 133) use the Swedish population register to estimate the elasticity of offspring's earnings with respect to family income averaged across near-complete childhood. The resulting elasticity estimate – which is considered the 'true' elasticity as it utilises parental income data over near-complete childhood – is then compared to estimates utilising fewer observations of parental income. Gregg, Jonsson, et al. (2017, p. 133) find that "averaging over adjacent years removes less bias in elasticities than averaging over years further apart". When parental family income was averaged over two *consecutive* years of childhood (ages 16 and 17), the estimated elasticity was approximately 70 percent of the true elasticity, but when it was averaged over two observations measured *six years apart* (ages 10 and 16) – hence breaking the serial correlation in transitory variation across consecutive income observations – the estimated elasticity increased to 80 percent of the true elasticity (Gregg, Jonsson, et al., 2017, p. 132).

In the classical model, noise in the left-hand-side dependent variable (offspring's income) does not lead to a biased estimate of the IGE, only to less statistical precision in its estimation (Hausman, 2001). However, contrary to the classical model, Nybom and Stuhler (2016) and Mitnik (2019) find that averaging incomes on the left-hand-side also reduces bias in estimates of the IGE. Using Swedish population register data on son-father pairs, Nybom and Stuhler (2016) estimate the intergenerational elasticity using almost lifelong income histories as a benchmark estimate. They then compare this benchmark elasticity to estimates using (logs of) three-, five-, and seven-year averages of sons' income (with no measurement error in father's incomes) to quantify the bias from using averages of different lengths. They find that the bias falls in the number of sons' income observations (but is not eliminated) because using logs of multiyear averages for offspring "counteract[s] the disproportionate influence of occasional low-income episodes" (Nybom & Stuhler, 2016, p. 266). They conclude with a recommendation to "averag[e] over multiple income observations from midlife (if available) for *both* fathers and sons" (Nybom & Stuhler, 2016, p. 266, italics in original).

As mentioned, the IGE is also sensitive to the specific ages over which short-run income is observed, due to offspring from different parental-income backgrounds having different age-income profiles, which produces *lifecycle bias* (Grawe, 2006; Haider & Solon, 2006; Jenkins, 1987). This bias arises because there is heterogeneity among the population in income growth over the lifecycle – the association between current and lifetime income changes over the lifecycle and this trajectory varies among individuals – so differentials in current income across individuals at a particular age will generally not be representative of differentials in lifetime average income. In particular, individuals with relatively high lifetime incomes tend to have relatively low incomes at young ages and steeper income trajectories than individuals with relatively low lifetime incomes, due to differences in human capital accumulation at early ages. As a result, when comparing the current incomes of those who will eventually have high and low lifetime income, an early-career comparison tends to understate the gap in their lifetime incomes, while a late-career comparison may overstate it (Haider & Solon, 2006, p. 1310).[38] Because these differences in age-income profiles correlate with parental characteristics, this introduces an asymmetric bias in the estimation of the IGE. Using current income early (late) in the lifecycle as a proxy for *offspring's* lifetime income will lead to an underestimation (overestimation) of the IGE, while using current income early (late) in the lifecycle as a proxy for *parent's* lifetime income will lead to an overestimation (underestimation) of the IGE (Lefranc, 2018). This is precisely what is observed in the literature: IGE estimates tend to increase as the age of sampled sons increases towards the thirties (Grawe, 2006; Solon, 1999) and as the age of sampled fathers decreases towards the thirties (Grawe, 2006). Haider and Solon (2006, p. 1317) find among their US sample of men that differences in current earnings measured "between the early thirties and mid-forties" closely mimic differences in lifetime annuitised earnings, so IGE estimates derived from earnings observed around mid-life are immune from lifecycle bias and subject only to classical attenuation bias.[39] Adding age control variables adjusts for the average income at each age but does not control for variance around the average (that is, different income profiles at each age), hence does not eliminate lifecycle bias.

Rank-based measures such as Spearman's rank correlation and transition matrices are also susceptible to these measurement error biases, but much less so than the IGE. Using a Swedish administrative series of near-lifetime income, Nybom and Stuhler (2017) compare the magnitude and direction of attenuation and lifecycle biases derived from estimation of the IGE, the Pearson linear correlation, the Spearman rank correlation, and transition probabilities. They find that the rank correlation is strongly underestimated at young ages because individuals with high lifetime income tend to have relatively low income (hence relatively low ranks) at the beginning of their career (Nybom & Stuhler, 2017, p. 811). They also find that rank-based measures are particularly prone to measurement error in the tails of the distribution. Here estimates are likely to be most sensitive, since top-income individuals will occasionally have low-income episodes, which will tend to overstate long-distance upward mobility (in the case of top-income parents with temporarily low incomes) and long-distance downward mobility (in the case of top-income offspring with temporarily low incomes) and understate persistence at the bottom.[40] But they also conclude that, of the four measures they evaluate, the rank correlation and transition probabilities "are both the least attenuated and the most stable over age [produce the smallest lifecycle bias, once measured beyond young ages]" (Nybom & Stuhler, 2017, p. 801). Similarly, Gregg, Jonsson,

---

[38] The measurement error that generates lifecycle bias is not of the classical type since the error in current income is correlated with true lifetime average income and also affects the dependent variable.

[39] Using Swedish administrative data, Nybom and Stuhler (2016) corroborate Haider and Solon's main finding that measuring incomes around mid-life substantially reduces bias in IGE estimates, but find that, even at this age, there is lingering residual lifecycle bias owing to heterogeneity in income trajectories beyond that induced by age which is likely to be correlated with parental characteristics such as income. As Nybom and Stuhler (2016, p. 241) put it, "the shape of income profiles varies with parental background even for a given level of lifetime income" meaning lifecycle bias "tends to be larger than [Haider and Solon's] model predicts".

[40] Top-income individuals may have low-income episodes for many reasons, including that they "could choose leisure and live off their wealth" or "they could earn their income abroad or for other reasons avoid domestic taxation" but these episodes "are unlikely to extend over long spans" (Nybom & Stuhler, 2017, p. 815). More generally, "anyone may temporarily move into the bottom of the distribution, while the top is more persistently populated by a select few" which has the effect of understating the inheritance of poverty and overstating 'riches to rags' downward mobility in annual data (Nybom & Stuhler, 2017, p. 820). Consequently, "[r]esearchers need to exercise particular caution when studying long-distance mobility, the inheritance of poverty, or the inheritance of top incomes" when using rank-based measures of mobility (Nybom & Stuhler, 2017, p. 819).

et al. (2017) and Mazumder (2016) find that the rank correlation is considerably more robust than the IGE to attenuation and lifecycle biases.

### *2.3.5 Empirical evidence of lifecycle bias*

Thus, the aim of the mobility analyst is to observe incomes at an age – call it the 'proxy age' – when current income differentials are broadly representative of lifetime income differentials. What is this proxy age? This is an empirical question, to which the answer will be country-specific, sex-specific (proxy ages being generally older for women compared to men), and probably specific to other sociodemographic groups as well (Nybom & Stuhler, 2016). To the author's knowledge, there are no studies that have estimated this age with New Zealand populations. However, studies in other countries have found that roughly between the early thirties and early forties for men, and the 40 to 50-year age window for women, is when current income is a reasonable proxy for lifetime average income, as listed in Table 1 below.

While the results outlined in Table 1 bear out the expected heterogeneity across populations, they nevertheless indicate some cross-country overlap in the age windows. Indeed, Brenner's (2010, p. 403) comparison of lifecycle bias in Germany with analogous US results from Haider and Solon (2006) and Swedish results from Böhlmark and Lindquist (2006) lead him to "cautiously conclude that the associations between annual and lifetime earnings of men are very similar across these Western industrialized countries". In the absence of any New Zealand evidence, these studies are the only source available to guide the choice of proxy ages used in this thesis.[41] On the basis of the evidence presented in Table 1, and given data constraints on the ages that can feasibly be used with the NZLC and CHDS datasets, the general rule of thumb adopted in this thesis (whenever *individual* measures of income are used) is to construct proxies for lifetime average income around age 35 for men (sons and fathers)[42] and around age 40 for women (daughters and mothers), using multiyear averages wherever possible.[43] The implementation of this rule of thumb, which differs between the two datasets, is discussed in chapter 3 (NZLC) and chapter 4 (CHDS).

[41] Nybom and Stuhler (2016, p. 261), who quantify the effect of lifecycle bias on son-father IGE estimates using Swedish administrative data, suggest that "[i]f the relationship between annual and lifetime incomes cannot be directly estimated then external evidence on the pattern of [lifecycle bias] may provide a useful approximation". Age-income profiles have been estimated for men and women in New Zealand (Coleman, 2006; Coleman & McDonald, 2010; Creedy, 1996; Hyslop, 2000). For example, Hyslop (2000) finds that women's market incomes display a bimodal lifecycle pattern, rising to a local peak in the mid-twenties, falling through to the early thirties, then rising again to a high in the mid-forties (reflecting fertility patterns), whilst men's market incomes display an inverted U shape, rising through the twenties and thirties, peaking between ages 40 and 45, and then declining steadily from the early fifties through to retirement. However, these studies do not estimate the age at which current income is most representative of lifetime average income in New Zealand.

[42] Corak, Lindquist and Mazumder (2014) also use time averages of income centred on age 35 for their sample of Swedish sons.

[43] It is likely that the age at which current income best represents lifetime average income *has changed over generations in New Zealand*. This is implied by the changing age-income profiles over pseudo-cohorts (proxying generational change) depicted in Coleman (2006) and Coleman and McDonald (2010). On the basis of his investigation into lifecycle bias in earnings IGE estimates for son-father pairs in the US, Canada, and Germany, Grawe (2006, p. 565) cautions that the age at which lifecycle bias is minimised for sons is not necessarily equal to the age at which the bias is minimised for fathers, noting "the tremendous change in educational attainment experienced over the last 40 years nearly guarantees that age earnings profiles have changed between generations". In Sweden, an increase over cohorts or generations in the proxy age has been documented by Böhlmark and Lindquist (2006) and Nybom and Stuhler (2016). But since there is no New Zealand evidence indicating what this proxy age might be for *one* generation, let alone two, the analysis is kept simple by using *the same* (sex-specific) proxy age for both offspring and parents.

**Table 1**. *Empirical studies estimating the age at which current income is most representative of lifetime average income.*

| Sex | Country | Cohort | Age window | Study |
|---|---|---|---|---|
| Men | USA | 1931-1933 | "between early thirties and mid-forties" | Haider and Solon (2006, p. 1318) |
| Men | Sweden | 1948-1950 | 34 | Böhlmark and Lindquist (2006, p. 892) |
| Men | Germany | 1939-1944 | "between the ages of 30 and 40"[1] | Brenner (2010, p. 399) |
| Men | Norway | 1959-1962 | "early thirties" | Nilsen, Vaage, Aakvik and Jacobsen (2012, p. 14) |
| Men | Sweden | 1955-1957 | "early 40s" | Nybom and Stuhler (2016, p. 265) |
| Men | Great Britain | 1970 | 33-34 | Gregg, Macmillan et al. (2017, p. 14 of appendix) |
| Men | Canada | 1963-1966 | "late thirties or early forties" | Chen, Ostrovsky and Piraino (2017, p. 4) |
| Men | USA | 1952-1960 | 37[2] | Mitnik (2019, p. 4) |
| Men | Italy | 1959-1965 | "between 34 and 35" | Bloise and Raitano (2021, p. 105) |
| Women | Canada | 1963-1966 | "late 30s" | Chen et al. (2017, p. 5) |
| Women | Germany | 1939-1944 | "around the age of 50" | Brenner (2010, p. 399) |
| Women | Sweden | 1948-1950 | "after the age of 40" | Böhlmark and Lindquist (2005, p. 26) |
| Women | Norway | 1959-1962 | 40 | Nilsen et al. (2012, p. 10) |

**Notes:**
[1] This refers to Brenner's (2010) estimate for left-side (son's) life-cycle measurement error. For right-side (father's) life-cycle measurement error, Brenner (2010, p. 399) finds bias is minimised in "the age range 35 to 45".
[2] This refers to Mitnik's (2019) estimate for sons' earnings. When he pools sons and daughters and uses offspring's family income rather than earnings, the age window at which lifecycle bias is minimised is "close to 40 years" (Mitnik, 2019, p. 4).

### 2.4 Empirical estimates of intergenerational mobility

This section reviews the evidence of intergenerational mobility or persistence, firstly across countries and secondly, specifically for New Zealand.

#### *2.4.1 Cross-national comparisons of intergenerational income mobility*

Discerning differences in intergenerational persistence across countries is a difficult task. It is frequently unclear whether differences in estimates from single-country studies reflect real cross-country differences in intergenerational persistence or simply data differences (data sources, variable definitions, sample selection criteria, the length of time and ages at which offspring's and parents' incomes are observed, the estimation methods used, and so on). Multi-country studies which attempt to harmonise data from existing sources across countries provide a better evidence base to assess cross-country differences, but even these studies are subject to the limitations of the original data sources and can never achieve perfect comparability. Furthermore, point estimates for different countries often have large standard errors. For this reason, Jäntti and Jenkins' (2015, p. 922) review of the literature concludes that "very little is known about how intergenerational income persistence and mobility vary across countries".

Despite these empirical challenges, the literature has settled on some 'stylised facts' based on comparative reviews of single-country studies and multi-country studies specifically designed to achieve maximal cross-country comparability for valid international comparisons. Comparative reviews of son-father earnings elasticities include a review of eight countries by Solon (2002), nine countries by Corak (2006), 12 countries by OECD (2008), 11 countries by Björklund and Jäntti (2009), and 12 countries by Blanden (2013), all of whom account for methodological differences across country-specific studies (to varying degrees, e.g., Blanden (2013) scales down instrumental variables estimates by an empirically-derived factor of 0.75 to make them more comparable to OLS estimates). Corak (2013) expands his earlier comparison to 22 countries including an estimate for New Zealand from Gibbons (2010), while OECD (2018) report estimates of son-father earnings elasticities for 32 countries (26 OECD members including Gibbons' (2010) estimate for New Zealand plus six emerging economies), of which 14 estimates were calculated by the OECD for their report using existing cross-sectional datasets and a two-sample two-stage least squares estimator. Multi-country studies include Jäntti et al. (2006) who estimate intergenerational elasticities of sons' and daughters' earnings with respect to their fathers' earnings in Denmark, Finland, Norway, Sweden, the UK, and the US, a follow-up paper by Raaum et al. (2007) who estimate the elasticity of sons' and daughters' own earnings and family earnings with respect to their parents' combined earnings (using nonlinear specifications) in the same countries except Sweden, Corak et al. (2014) who estimate intergenerational elasticities of sons' earnings with respect to their fathers' earnings in Sweden, Canada, and the US, and Bratberg et al. (2017) who estimate intergenerational elasticities of offspring's family or household income (pooling sons and daughters) with respect to their parents' family or household income in Norway, Sweden, Germany, and the US (however, the IGE is not the focus of the analysis in Corak et al. (2014) or Bratberg et al. (2017)).

Some patterns emerge from these comparative reviews and multi-country studies. Son-father earnings elasticities in the Nordic countries are generally estimated to be in the 0.15-0.25 range[44] and in the UK and the US to be in the 0.40-0.50 range. Elasticities in Canada are generally estimated to lie in the upper part of the Nordic range, around 0.25; in Australia it is estimated to be around 0.25-0.35;[45] while Italy and France tend to have estimates in the UK/US range.[46] Elasticities in lower-income developing countries (Brazil, China, Colombia, India, Peru) lie in the 0.50-0.75 range. Gibbons' (2010) son-father estimate for New Zealand is 0.290, which ranks New Zealand as seventh-most mobile of the 22 countries in Corak (2013) and sixth-most mobile of the 32 countries in OECD (2018). Estimates of daughter-father elasticities (Jäntti et al., 2006; Raaum et al., 2007), intergenerational Pearson correlations (Corak et al., 2014; Jäntti et al., 2006), and the intergenerational rank correlation (Bratberg et al., 2017; Corak et al., 2014) produce similar cross-country rankings but the gaps between the Nordic countries and the UK/US are reduced and the UK tends to have higher daughter-father elasticities than the US.

Country estimates of intergenerational earnings elasticities have been plotted against estimates of cross-sectional income inequality for each country (the Gini coefficient of some measure of income in the parent generation), revealing a pattern in which countries with higher cross-sectional inequality tend to have higher intergenerational earnings persistence (less mobility), a

[44] Among the Nordic countries, Denmark is consistently found to have the lowest persistence (around 0.15) while Sweden is typically found to have the highest persistence (around 0.25) (see the seven reviews cited).

[45] More recent evidence from Deutscher and Mazumder (2020) – the first based on administrative data – estimate son-father persistence in individual total income in Australia to be 0.157, which is at the lower end of the Nordic range.

[46] An exception is Blanden (2013), whose scaled estimates for Italy and France are around 0.32-0.33, thus ranking below (exhibiting more mobility than) the UK and the US. Also noteworthy is the comparative study of intergenerational income persistence in Denmark and the US by Landersø and Heckman (2017), who argue that the choice of income concept is very important when making cross-country comparisons, as they estimate persistence to be slightly *higher* in Denmark than in the US when the income concept is gross income excluding public transfers (but considerably lower in Denmark on four other income measures).

relationship referred to as 'the Great Gatsby curve'.[47] Plots of the Great Gatsby curve appear in the comparative reviews mentioned above by Björklund and Jäntti (2009), Blanden (2013), Corak (2013), and OECD (2018), the latter two including New Zealand.[48] Corak (2013, p. 114) concludes that "[m]ore inequality at a point in time is associated with less generational mobility", with New Zealand appearing in a cluster in the lower-left quadrant alongside Australia and Canada, each with a relatively low degree of persistence and average level of inequality (but distinct from the Nordic cluster located in the extreme lower-left corner). OECD (2018, pp. 16, 35) concludes that "earnings mobility prospects tend to be usually weaker in countries where income inequality is high, and stronger in countries where inequality is low" but notes that "a few European countries buck this trend" namely "Hungary, France, Germany and Austria [which] combine both lower inequality and lower earnings mobility" while at the same time "there are no countries which combine high inequality with high mobility". As in Corak (2013), OECD (2018) places New Zealand in a cluster with Australia and Canada with relatively low persistence and average inequality. Within-country versions of the Great Gatsby curve, in which geographical areas within a country that have higher income inequality tend to have higher intergenerational persistence, have also been found in the US (Chetty et al., 2014; Nam, 2021) and Canada (Corak, 2020a).[49]

### *2.4.2 Studies of intergenerational mobility or persistence in New Zealand*

There have been several empirical studies estimating intergenerational persistence of income, socioeconomic status, social class, educational attainment, and social welfare benefit receipt in New Zealand. Table 2 lists these studies, their data sources (for the New Zealand data only in the case of comparative studies that use different data sources for different countries), the offspring-parent pairs analysed (the term 'offspring' is used where sons and daughters are pooled, 'parents' is used where fathers' and mothers' information is combined or averaged, and 'parent' is used where the analysis uses the parent with the higher income, socioeconomic status, social class, or educational attainment), the sample size (for New Zealand only where the study is a comparative one, and the largest sample size where more than one specification or offspring-parent pair is used), the main method(s) used to quantify intergenerational persistence, and the main quantitative estimate(s) of intergenerational persistence. Note that for many of these studies, estimating intergenerational persistence is one of several aims, or is not the main aim. This is particularly the case for studies of socioeconomic and class persistence which typically

[47] The Gini coefficient of income inequality is usually taken from when the offspring were children (i.e., income inequality among their parents, plotted against intergenerational persistence when those offspring are adults). However, Blanden (2013) correlates persistence against inequality measured during childhood *and adulthood* (and for a variety of measures of inequality including the Gini coefficient, the Atkinson inequality index, the 90-10 percentile ratio, etc.). She finds a "strong positive relationship between inequality and intergenerational mobility" across all measures of inequality with correlations always over 0.5, although intergenerational persistence is more strongly correlated with income inequality in adulthood (correlations of about 0.85 compared to about 0.6 for childhood), because it is "pick[ing] up the influence of labour market returns…to [adult] characteristics such as education and occupation" (Blanden, 2013, pp. 57, 58).

[48] Björklund, Jäntti and Nybom (2017) and Narayan et al. (2018) also present updated versions of the Great Gatsby curve (excluding New Zealand).

[49] However, Deutscher and Mazumder (2021, p. 60) do not find the Great Gatsby relationship holds within Australia, because when they compare regions across Australia, "the correlations between the relative measures [of intergenerational persistence such as the IGE] and our measure of inequality, the standard deviation of income, is quite weak" indicating that "there is no mechanical relationship between inequality and relative intergenerational mobility", but note that "the cross-country Great Gatsby Curve (and its absence within Australia) neither rules in nor rules out any deeper causal mechanisms linking inequality to mobility in either setting".

**Table 2**. *Studies of intergenerational persistence of income, socioeconomic status, social class, educational attainment, and social welfare benefit receipt in New Zealand.*

| Intergenerational variable | Study | Data source for New Zealand | Offspring-parent pairs | Sample size | Estimation method | Main quantitative estimate of persistence |
|---|---|---|---|---|---|---|
| Income | Andrews and Leigh (2008) | 1999 International Social Survey Programme | Son-father | <300[1] | Regression of son's log hourly wages on father's predicted log hourly wages (predicted from father's occupation) | Son-father IGE of 0.245, 10th most mobile of 16 countries, intergenerational correlation of 0.191 |
| Income | Gibbons (2010) | Dunedin Study (32-year follow-up in 2003-2005) | Offspring-parents<br>Offspring-father<br>Son-father<br>Daughter-father | 926 | Regression of offspring's log individual total income on father's or combined parents' log total income | Offspring-parents IGE of 0.272, offspring-father IGE of 0.264 , son-father IGE of 0.290, daughter-father IGE of 0.215[2] |
| Socioeconomic status | Pearson (1980b)[3] | Marriage certificates 1883-1975 from Wellington | Son-father | 1,437 | Transition matrix (5x5 SES statuses) | 43 percent persistence in SES status[4] |
| Socioeconomic status | Pearson and Thorns (1983) | Marriage certificates 1883-1970 from Wellington | Son-father | 3,116 | Transition matrix (5x5 SES statuses) | 43 percent persistence in SES status[5] |
| Socioeconomic status | Hall, Thorns and Willmott (1983) | Marriage certificates 1866-1981 from Christchurch | Son-father<br>Daughter-father | 1,577 | Transition matrix (7x7 SES statuses) | 32 percent persistence in SES status among son-father pairs[6] |
| Socioeconomic status | Pearson and Thorns (1986) | Marriage certificates 1928-1970 from Wellington and Christchurch | Son-father | 3,986 | Transition matrix (4x4 SES statuses) | 42 percent persistence in SES status[7] |
| Socioeconomic status | Nash (1993) | 1989 Access and Opportunity in Education Survey | Son-father | 1,096 | Transition matrix (6x6 SES statuses) | 31 percent persistence in SES status[8] |
| Socioeconomic status | Gibbons (2010) | 1996 New Zealand Election Study | Offspring-father<br>Son-father<br>Daughter-father | 3,268 | Regression of offspring's SES score on father's SES score | Regression coefficient of 0.20 for offspring-father pairs, 0.23 for son-father pairs, 0.18 for daughter-father pairs |
| Socioeconomic status | Belsky et al. (2018) | Dunedin Study (38-year follow-up in 2010-2012) | Offspring-parents | 831 | Transition matrix (5x5 SES quintiles) | 25 percent persistence in SES status[9] |
| Social class | Davis (1979) | 1976 Survey of Adult Oral Health | Son-father | 2,530 | Transition matrix (6x6 social classes)[10], regression of son's social class rank on father's social class rank | 31 percent persistence in social class, rank correlation coefficient (Somer's *D*) of 0.290 |
| Social class | Jones and Davis (1986, 1988) | 1976 Survey of Adult Oral Health | Son-father | 1,453 | Transition matrix (8x8 social classes), log linear modelling | 26 percent persistence in social class[11], degree of exchange mobility similar in New Zealand and Australia |
| Social class | Lauder and Hughes (1990) | 1982 Christchurch School Leavers Study | Offspring-parents | 2,245 | Transition matrix (3x3 social classes) | 45 percent persistence in social class[12] |
| Social class | Olssen, Griffin and Jones (2011) | Marriage certificates 1881-1940 from Dunedin | Son-father | 3,795 | Transition matrix (9x9 social classes), log linear modelling | 37 percent persistence in social class |

**Table 2 (continued)**. *Studies of intergenerational persistence of income, socioeconomic status, social class, educational attainment, and social welfare benefit receipt in New Zealand.*

| Intergenerational variable | Study | Data source for New Zealand | Offspring-parent pairs | Sample size | Estimation method | Main quantitative estimate of persistence |
|---|---|---|---|---|---|---|
| Educational attainment | De Broucker and Underwood (1998) | 1996 International Adult Literacy Survey | Offspring-parent<br>Son-parent<br>Son-father<br>Son-mother<br>Daughter-parent<br>Daughter-mother<br>Daughter-father | 2,801 | Regression of offspring's highest educational qualification on parent's highest educational qualification | Rank correlation of 0.34 for offspring paired with more-qualified parent, 4th most mobile of 11 countries |
| Educational attainment | Hertz et al. (2007) | 1996 International Adult Literacy Survey | Offspring-parents | 2,675 | Regression of offspring's years of schooling on parents' average years of schooling | Regression coefficient of 0.40, 6th most mobile of 42 countries |
| Educational attainment | Pfeffer (2008) | 1996 International Adult Literacy Survey | Offspring-parent | 2,062 | Log-linear modelling of exchange mobility in educational attainment abstracting from changes in marginal distributions | 3rd most mobile of 20 countries |
| Educational attainment | Narayan et al. (2018) | 1999 International Social Survey Programme[13] | Offspring-parent(s)<br>Son-parent(s)<br>Son-father<br>Son-mother<br>Daughter-parent(s)<br>Daughter-mother<br>Daughter-father | 725 | Regression of offspring's years of schooling on parent's (or parents' average) years of schooling | Regression coefficient of 0.178, 3rd most mobile of 111 countries |
| Educational attainment | OECD (2018) | 2014-2015 Programme for the International Assessment of Adult Competencies | Offspring-parent | Not reported (but ~4,000)[14] | Transition matrix (3x3 educational attainment categories) | 31 percent persistence in educational attainment among offspring whose parents have not attained upper secondary qualification, 66 percent persistence among offspring with a tertiary-qualified parent |
| Educational attainment | Lee and Lee (2021) | 2014-2015 Programme for the International Assessment of Adult Competencies | Offspring-parent | Not reported (but ~4,000)[14] | Regression of offspring's years of schooling on parent's years of schooling (threshold regression with censored covariate) | Regression coefficient of 0.167, 2nd most mobile of 30 countries |
| Educational attainment | Nel (2021) | 2017-2020 World Values Survey | Offspring-parent | 8,003 | Regression of offspring's highest educational qualification on parent's highest educational qualification | Not reported[15] |
| Social welfare benefit receipt | Maloney, Pacheco and Maani (2003), Pacheco and Maloney (2003) | Christchurch Health and Development Study (18-year and 21-year follow-ups in 1995 and 1998) | Offspring-parents<br>Son-parents<br>Daughter-parents | 847 | Regression of offspring's benefit propensity on parents' benefit propensity | Regression coefficient of 0.345 among offspring-parents pairs, 0.236 among son-parents pairs, 0.550 among daughter-parents pairs |

**Notes:**

[1] Andrews and Leigh (2008) do not report their sample size for New Zealand, but inspection of the publicly-available 1999 International Social Survey Programme dataset indicates that fewer than 300 male respondents satisfy their sample selection criteria.
[2] These results from Gibbons (2010) are for cohort members irrespective of country of residence at age 32.
[3] Preliminary results are reported in Pearson (1980b).
[4] This is an approximate estimate derived across all subperiods reported in the transition matrix in Table 6.4 of Pearson (1980a, p. 108).
[5] This is an approximate estimate derived across all subperiods and suburbs reported in the transition matrices in Table 5.1 and Table 5.3 of Pearson and Thorns (1983, pp. 114, 120).
[6] This is an approximate estimate derived across all subperiods and suburbs reported in the transition matrix in Table 8 of Hall, Thorns, and Willmott (1983, p. 133). A similar table for daughters is not reported.
[7] This is an approximate estimate derived across all subperiods and suburbs reported in the transition matrix in Table 2 of Pearson and Thorns (1986, p. 217).
[8] This estimate is derived from the transition matrix in Table 1 of Nash (1993, p. 48). Preliminary results are reported in Nash, Harker and Charters (1990).
[9] This is an approximate estimate derived from the transition matrix presented in Supplemental Figure 1, Panel C in Belsky et al. (2018, p. 33 of online supplemental material).
[10] For comparison with Australian and US results, Davis (1979) collapses the social classes down to a three-strata classification.
[11] Jones and Davis (1986, p. 22) estimate social class persistence to be 25 percent, 30 percent, and 51 percent when using nine-strata, six-strata, and three-strata classifications of class, respectively.
[12] This is an approximate estimate derived from Figure 1 in Lauder and Hughes (1990b, p. 50) and Table 10.2 in Hughes and Lauder (1990, p. 154).
[13] Narayan et al. (2018) use data from the World Bank's Global Database on Intergenerational Mobility, for which the New Zealand data are from the 1999 International Social Survey Programme.
[14] OECD (2018) and Lee and Lee (2021) do not report their sample size for New Zealand, but the 2014-2015 PIAAC surveyed 6,177 New Zealanders, of whom 4,720 were aged 25 to 65, which is approximately the age range analysed in both studies (before further sample restrictions are applied).
[14] Nel (2021) does not report his regression coefficient estimates (for any country) because they are derived only as an explanatory variable in a model attempting to explain attitudes to inequality among New Zealanders, which is the aim of Nel's (2021) paper.

analyse multiple aspects of social and economic stratification. Also note that some studies use a range of methods to quantify and analyse mobility and obtain a set of results from which many conclusions are reached or which present a mixed picture of mobility, the nuance of which will often be lost in a summary like Table 2.

There are only two published studies of intergenerational income persistence in New Zealand. Andrews and Leigh (2008) use data from the 1999 International Social Survey Program (ISSP) to examine the relationship between income inequality and intergenerational earnings persistence across countries. To estimate persistence, they use data on male respondents aged 25 to 54 in the 1999 ISSP (sons, born 1945 to 1974) whose annual earnings are observed in 1999 (which they convert to log hourly wages using data on weekly hours), but there is no data on the incomes of respondents' parents, so they use fathers' occupation – which *is* reported – to impute fathers' earnings. For each country, they first predict earnings at age 40 in different occupations, then all fathers in the same occupation are assigned the same predicted wage.[50] Sons' log hourly wage is then regressed on fathers' predicted log hourly wage for 16 countries including New Zealand. The resulting son-father IGE estimate for New Zealand is 0.245 and the intergenerational correlation (the focus of their analysis) is 0.191.[51] Compared to the other 15 countries, these Zealand estimates lie near the (unweighted) average and rank towards the less-mobile end of the spectrum, indicating a society characterised by an average-to-low degree of mobility; New Zealand's IGE estimate ranks as 10th most-mobile (Russia has the lowest IGE at 0.060, Chile the highest at 0.571, Australia at 0.326) and New Zealand's intergenerational correlation ranks as 11th most-mobile (Russia and Chile again occupy the most and least-mobile ranks). Andrews and Leigh then plot their intergenerational correlation estimates against each country's Gini coefficient in (or near) 1975 to measure income inequality and quantify the resulting association. They find a positive but statistically insignificant relationship for all 16 countries. However, the relationship becomes statistically significant when six former Communist countries are excluded. They find a 10-point rise in the Gini coefficient is associated with a 0.07-0.13 increase in the intergenerational correlation. Andrews and Leigh (2009, pp. 1491-1492) conclude that "sons who grew up in more unequal countries in the 1970s were less likely to have experienced social mobility by 1999" so "[m]oving from rags to riches is harder in more unequal countries".

Gibbons (2010) estimates intergenerational persistence of income and socioeconomic status among sons and daughters in New Zealand (the results for socioeconomic persistence are discussed in chapter 3). For income persistence, Gibbons uses data on a birth cohort of sons and daughters born in 1972-1973 drawn from the longitudinal Dunedin Multidisciplinary Health and Development Study ('the Dunedin Study'[52]). When cohort members (offspring) were aged 13 (in 1985-1986) and 15 (in 1987-1988), the primary parent (usually the mother) reported the father's

[50] Because Andrews and Leigh (2008) use the 1999 ISSP data to observe sons' earnings *and* predict fathers' earnings, their approach does not take into account changes in average earnings of occupations over time.
[51] Andrews and Leigh (2008) – subsequently published in a more condensed form as Andrews and Leigh (2009) – do not state their sample size for New Zealand, but inspection of the publicly-available 1999 International Social Survey Programme dataset indicates that fewer than 300 male respondents satisfy their sample selection criteria (including that they have non-missing responses for age, earnings, and father's occupation).
[52] The Dunedin Study is tracking the health and development of a birth cohort of 1,037 children born in Dunedin between April 1972 and March 1973 who were still living there at the age of three. By age 32, 94 percent of cohort members assessed at age three were still participating.

annual total income and the combined total income of the father and mother ('family income').[53] When offspring were followed-up at age 32 (in 2003-2005), they reported their own annual total income. Offspring's and parents' incomes were collected in bands and they were assigned the mid-point of their band, except for offspring in the open-ended top category who were assigned a value of $135,000 and parents in the top categories in both years who were assigned a value of $80,600 (when deflated to 2008 dollars).[54] An average of the two parental income observations was taken wherever possible. The incomes of the 24 percent of offspring who were living overseas at age 32 were converted to New Zealand dollars using purchasing power parity conversion. Offspring and fathers who report incomes of zero were dropped from the analysis.

Gibbons estimates the IGE of offspring's total income at age 32 with respect to their father's total income (controlling for father's age and age-squared) among the 592 offspring living in New Zealand to be 0.212 (0.253 among sons, 0.167 among daughters) and among all 780 offspring regardless of their country of residence to be 0.264 (0.290 among sons, 0.215 among daughters). Among the latter sample of all offspring irrespective of country of residence, when father's income is replaced with family income (and controls for father's age replaced with controls for offspring's sex), the IGE is virtually unchanged, increasing marginally by three percent from 0.264 to 0.272, indicating that "[c]hanging the explanatory variable…had little effect on the results" which is "not entirely surprising" because "when participants were 13 and 15 on average fathers received 75% of the total income of Dunedin Study families" (Gibbons, 2010, p. 22). The age controls are not statistically significant in any of the models and nor is the difference between the IGEs for sons and daughters. All estimates have wide confidence intervals. When Gibbons (2010, p. 51) plots offspring's log income against father's log income, the resulting graph reveals "a weak [positive] relationship between parents' incomes and the incomes of their children, particularly for female participants".

As a robustness check, Gibbons drops up to three fathers (depending on the model) who have extremely low incomes and up to three offspring (depending on the model) who also have extremely low incomes that are distorted by currency conversions. The resulting estimates of the IGE tend to increase, especially for women, and indicate that "results are sensitive to small changes in the data used" (Gibbons, 2010, p. 25). The IGE excluding low-income outliers among offspring living in New Zealand increases to 0.304 (0.323 among sons, 0.272 among daughters) and among all offspring irrespective of country increases to 0.337 (0.312 among sons, 0.323 among daughters). Among the latter, when the regressor is changed to family income, the IGE is 0.311. Gibbons also experiments with adding a polynomial term (the square of father's income) to test if the relationship between offspring's and parents' income is nonlinear. However, the term was not statistically significant, so he cannot rule out the possibility that the relationship is linear. Gibbons (2011, pp. 54, 41) argues that his results from analysis of the Dunedin Study data "can be cautiously extrapolated to other New Zealanders born in the early 1970s because the study included children from a full range of backgrounds…and because, irrespective of where in New

[53] On average, fathers were aged 42 and mothers were aged 40 across the 13-year and 15-year follow-ups.
[54] At age 13, the income question for parents used 10 bands and specifically excluded social security benefits, while at age 15 the question used 14 bands and included social security benefits.

Zealand they live, all New Zealanders have the same entitlements to social services" but also emphasises that the research is "exploratory and our findings are very tentative".[55]

Gibbons compares these estimates to the most comparable estimates from studies of other developed countries that were available at the time. He finds that the confidence intervals for his preferred estimates for sons and daughters overlap with almost all other countries considered; the exceptions are Denmark, which has significantly lower income persistence than in the Dunedin Study, and one result for the US for which the estimate of persistence among sons is significantly higher than that from the Dunedin Study (although the US data uses a multiyear average of sons' incomes measured at older ages). Gibbons (2010, p. 38) concludes that "rates of intergenerational income mobility for New Zealand men and women are probably within a similar range to rates of intergenerational income mobility in most other developed countries". Gibbons (2011, pp. 58, 59) summarises his conclusions as follows:

> The results indicate that in New Zealand the income or SES of a person's parents when they are a teenager appears to have a modest effect on their subsequent economic outcomes…. Intergenerational income mobility rates for New Zealanders appear to be in a similar range to rates for people born in other developed countries…Although the results improve our knowledge of intergenerational mobility in New Zealand, further research using larger data sets would be desirable.

Studies of intergenerational mobility of socioeconomic status and social class use information on the occupations of offspring and parents to assign them a socioeconomic score or ranking (using indices that rank occupations based on the average income and education levels of persons with those occupations in the New Zealand census, or based on the social prestige of occupations), in which case it is said to be a study of 'intergenerational socioeconomic mobility', or to assign them to a social class or strata in a class hierarchy, in which case it is a study of 'intergenerational social mobility'. Studies analysing intergenerational socioeconomic mobility in New Zealand include Pearson (1980a), Pearson and Thorns (1983, 1986), Hall et al. (1983), Nash (1993), Gibbons (2010), and Belsky et al. (2018).[56] Studies analysing intergenerational social class mobility in New Zealand include Davis (1979), Jones and Davis (1986, 1988), Lauder and Hughes (1990b), and Olssen et al. (2011).[57] In all these studies, the data sources containing the occupational information are of two broad types: survey data, primarily cross-sectional with retrospective reporting of parents' occupation, or historical information gleaned from marriage certificates which record the occupations of bride, groom, and their respective fathers at the time

[55] Gibbons (2010, p. 38) recommends his findings be "cross-validated using income data from the Christchurch Study of 1,265 children born in mid-1977", a recommendation taken up in this thesis.

[56] See also Reuben et al. (2017), who extend the analysis of Dunedin Study data in Belsky et al. (2018) by testing whether lead exposure in childhood is associated with downward intergenerational socioeconomic mobility, and Baldock (1971) who compared the occupational aspirations of 3,773 14-year-old sons in New Zealand with the occupations held by their fathers, where occupations were ranked by social prestige and income, thus measuring the intergenerational transmission of occupational preferences.

[57] See also Loomis (1985, p. 22) who replicates Davis (1979) with a small survey of 208 male Cook Islanders residing in Auckland in 1981 and finds that, compared to the general population in Davis's (1979) sample, Cook Islands men experienced greater structural mobility but were "much less mobile than the average New Zealander" in terms of exchange mobility. Another small, early study by Robb and Cloud (1970) – perhaps the first empirical study of intergenerational mobility in New Zealand – used data on 293 New Zealand European sons drawn from a survey of a random sample of Rotorua's adult population in 1957 who reported their occupation and their father's occupation, which they code to six social classes and cross-tabulate in a transition matrix. They find that 23 percent of sons were in the same social class as their father (intergenerational persistence), 37 percent had moved upwards, and 40 percent had moved downwards, concluding there is "more mobility, and considerably less inheritance, in this New Zealand town than in the United States samples" with which they compare their results (Robb & Cloud, 1970, p. 50).

of marriage.[58] The main method typically used in these studies to estimate intergenerational mobility/persistence is to compute a transition matrix by cross-tabulating offspring's (usually son's) socioeconomic status or social class with parent's (usually father's) socioeconomic status or social class, and then report probabilities of intergenerational persistence (the proportion of offspring clustered on the leading diagonal of the transition matrix) and of upward and downward mobility. Because the number of socioeconomic status groups or social classes differs between studies due to different socioeconomic indices or social class schema used to code occupations, the resulting transition probabilities, which are mechanically affected by the number of categories used in the matrix, are not directly comparable.[59] In almost all of these studies, there is no abstraction from changes over time in the marginal distributions of occupation, so the transition probabilities capture both changes in the marginal distributions (that is, offspring experiences of 'structural mobility' relative to their parents as the proportion of lower-skilled occupations shrinks and higher-skilled occupations expands in the economy over time) and changes in ranks (offspring experiences of 'exchange mobility' as offspring move to a higher or lower position in the evolved occupational structure compared to their parents' rank in the earlier occupational structure). These studies have generated a range of quantitative estimates and qualitative conclusions about the degree of intergenerational persistence in New Zealand and whether New Zealand is best characterised as an 'open', mobile society with equality of opportunity or an immobile society in which family background strongly influences people's outcomes in life. Since in practice any study of intergenerational transmission will find some degree of persistence and some degree of mobility in the population, the conclusions drawn can depend on which of the two the analyst wishes to emphasise or, as Mazumder (2003, p. 3) puts it, whether the analyst views "the glass as "half-full" or "half-empty"". This highlights the utility of cross-national comparisons which enable estimates to be benchmarked against those from studies of other countries. The socioeconomic and class persistence studies listed in Table 2 are discussed in more detail in chapters 3 and 4.

There have also been several empirical studies estimating intergenerational educational persistence in New Zealand using survey data capturing the educational attainment of respondents and their parents, by De Broucker and Underwood (1998), Hertz et al. (2007), Pfeffer (2008), Narayan et al. (2018), OECD (2018), and Lee and Lee (2021).[60] The first three of these studies obtain their New Zealand data from the 1996 International Adult Literacy Survey. Narayan et al. (2018) make use of the World Bank's Global Database on Intergenerational Mobility, for which the New Zealand data are from the 1999 International Social Survey Programme. OECD

[58] Social historians recognise that a major limitation of marriage certificates for studying intergenerational mobility is that "a snapshot of the father's and the son's occupations on the day when the marriage record was filled out [generally] compares a father at the height of his occupational achievement with a son who has not long been in the labour market" (Olssen et al., 2011, p. 22). Other limitations of marriage certificates include that reported occupations are often ambiguous, the permanently single are ignored, the occupations of brides and mothers were rarely recorded, and informal marriages are not captured.

[59] For any given study, the smaller the number of categories or bins used in the transition matrix, the greater the estimate of intergenerational persistence, *ceteris paribus*.

[60] See also Nel (2021) who uses data from the 2017-2020 World Values Survey to estimate intergenerational educational persistence in Australia, Germany, Greece, Japan, New Zealand, and the US, measured by the slope coefficient obtained from regressing offspring's educational attainment on parents' educational attainment. However, he does not report these estimates (for any country) because they are derived only as an explanatory variable in a model attempting to explain why New Zealanders are, on average, less supportive of income redistribution and less averse to inequality compared to their counterparts in other high-income OECD countries, which is the aim of Nel's (2021) paper.

(2018) and Lee and Lee (2021) obtain their New Zealand data from the 2014-2015 Programme for the International Assessment of Adult Competencies (PIAAC). Except for Pfeffer (2008) and OECD (2018), all these studies estimate intergenerational persistence in educational attainment either as the correlation between offspring's and parents' highest qualification measured in categorical terms, or the coefficient from a regression of offspring's years of schooling on parents' years of schooling measured continuously, where parents' attainment is usually either the attainment of the more-qualified parent or an average of mothers' and fathers' attainment (sometimes fathers' and/or mothers' attainment are used separately). These studies, discussed further in chapter 3, are almost unanimous in finding that New Zealand has very low intergenerational persistence of educational attainment. Finally, Maloney, Maani and Pacheco (2003) estimate the intergenerational correlation in social welfare benefit receipt between parents and their offspring in New Zealand and Pacheco and Maloney (2003) estimate it separately by gender, the results of which are discussed in chapters 3 and 4.

## Chapter 3: The New Zealand Longitudinal Census

In this chapter, census microdata linked over time is used to estimate intergenerational income persistence in New Zealand for sons and daughters born in New Zealand in the late 1960s and 1970s. Section 3.1 describes the data source and income variables used in the analysis. Section 3.2 describes the methods used to construct the long-run income variables and the selection of the samples analysed. Section 4.3 presents descriptive statistics of the sociodemographic and income characteristics of the samples and assesses their representativeness. Section 4.4 presents results of estimates of the intergenerational income elasticity and estimates of persistence and mobility in income rank. Section 4.5 compares these results with those from other countries, and section 4.6 discussed the limitations of the analyses.

### 3.1 Data source and income variables

#### *3.1.1 Description of data source*

Created by Statistics New Zealand, the New Zealand Longitudinal Census (NZLC) is a dataset that links individuals' records from the New Zealand Census of Population and Dwellings over time (Didham, Nissen, & Dobson, 2014).[61] Each census is linked backwards through time to the previous census, creating six pairs (2013-2006, 2006-2001, 2001-1996, 1996-1991, 1991-1986, 1986-1981).[62] Because each pair has a common boundary with one or more other pairs, individuals can be tracked longitudinally over varying lengths of time. At most, individuals can be linked across all seven censuses, that is, from 2013 right back to 1981, a 32-year timespan.

In each census there is a subset of the usually resident population who are eligible for linking to the earlier census, known as the 'theoretical population', made up of those individuals who, at the previous census, had been born *and* had filled out a census form *and* resided in New Zealand and stated their address (since address information is used to link individuals across censuses). All individuals who meet these criteria should, in theory, be linkable to the earlier census.

A two-stage linking process was applied to the theoretical population: deterministic linking in the first stage, followed by probabilistic linking of records not linked in the first stage. The deterministic stage linked records based on several 'blocking variables' (that is, partitioning records on a key variable to reduce the number of comparisons needing to be made to match on other key variables), namely sex, exact date of birth, area unit of usual residence[63] (to distinguish between individuals with the same sex and date of birth), country of birth, and Māori descent. Names and exact addresses were not available for linking census records. Two records were treated as linked if (1) they matched exactly on the blocking variables; (2) the link was unique (no other record was

[61] The census is usually conducted every five years, however there is a gap of seven years between the 2006 and 2013 censuses due to a major earthquake in Christchurch in 2011 which resulted in the scheduled 2011 census being cancelled.
[62] Initially, linking was completed from the 2006 census backwards in time through pairs of censuses to 1981. Statistics New Zealand then published the results of this work (in Didham et al., 2014). Only subsequent to the completion of this work and publication of the aforementioned report was the 2013 census then added to the NZLC. Consequently, much of the description of the NZLC dataset detailed in this chapter relates to the 2006-1981 linking work documented in Didham et al. (2014) (that is, excluding the linking of the 2013 census).
[63] Although the census questionnaires ask respondents for their full address of usual residence and the full address of usual residence five years ago, Statistics New Zealand only stores the area unit of the latter. This meant that area unit was the smallest geographic unit available for linking across censuses. Any records that had missing or imputed information for date of birth, sex, usual residence at the current census, and address at the previous census were not eligible for deterministic linking.

an equally exact link), and; (3) the responses for each of the blocking variables was complete and valid.[64] Approximately 69 percent of the theoretical population were deterministically linked, with all remaining records proceeding to probabilistic linking.

In probabilistic linking, records with the same year of birth and area unit of usual residence were compared on how close they were on day of birth, month of birth, and sex. Links were deemed to be true matches if the probability they were a match exceeded a certain cut-off weight. An additional three percent of records were linked in the probabilistic stage. During the linking process, records are assigned a unique reference ID number, which enables linked records to be joined across pairs and across specific subsets of the longitudinal series. Records that failed to be linked (deterministically or probabilistically) are indicated as non-linked in the NZLC dataset.

Across the five census pairs between 1981 and 2006, the average link rate (the percentage of the theoretical population at any given census linked to the previous census) was 72 percent. The link rate falls as the timespan increases. For example, moving from pairs of censuses to any three consecutive censuses, the average link rate drops to 57 percent. Looking at the records linked from 2006 back to 1981, the link rate is 32 percent; that is, only about one-third of the theoretical population at the 2006 census are linked all the way back to 1981.[65] The linking methodology used to construct the NZLC links individuals across *consecutive* censuses, so there is no 'leap-frog' linking across non-consecutive censuses. Thus, a person who is absent for an intermediate census cannot be linked from the prior census to the one afterwards.

Which groups of people are less likely to be linked across censuses? First and foremost, people who change addresses frequently are less likely to be linked, mainly because they are more likely to incorrectly recall their address at the previous census, or provide incomplete or erroneous information on key variables in the census, or not be counted at the previous census. Similarly, people who were overseas during one or more censuses (hence did not complete census forms) cannot be linked back over the period they were away (Didham & Nissen, 2015; Didham et al., 2014).

Analysis by Didham et al. (2014) found that link rates also differed by age, sex, and ethnicity. Link rates are higher for children and middle-aged adults, and lower for young adults, especially people aged in their twenties, with link rates bottoming-out for people in their late twenties. For instance, in the 2006-2001 census pair, link rates for people aged between 50 and 75 years were almost 80 percent, whereas for people aged 25 to 29 years they were about 50 percent, a pattern typical for all census pairs. Cohort effects are also present, such that link rates for the same age groups have tended to decline over time, reflecting increases over time in residential mobility and international migration, and decreases in public compliance in census form completion. Males have lower link rates than females at virtually all ages, partly because they are more likely than females to be missed by the census or to have missing or different data between one census and

[64] Note that deterministically linked records "merely indicate that those records had consistent information over time for the blocking variables" and that there is "no absolute certainty that the links refer to the same person since…neither names nor exact addresses [were] available to use for linking of historic census records" (Didham et al., 2014, p. 11). Thus, while the majority of the linked record pairs are expected to be true matches, it is also expected that a small portion of links will not be true matches (Didham et al., 2014).

[65] The link rate for the 2013-2006 pair has not been published.

the next. Māori (58 percent link rate, on average), Pacific (57 percent), and Asian groups (65 percent) have lower link rates compared to the total population (72 percent).

Link rates also differed by several other demographic and socio-economic characteristics. Link rates were generally lower for: overseas-born respondents compared to those born in New Zealand; non-partnered compared to partnered respondents, except for those aged under 30 years; people who did not own a home compared to home-owners, except for those aged under 25 years; respondents without a formal educational qualification compared to those with one; respondents not in employment compared to employed people; and respondents in 'blue-collar' occupations (agriculture, fishery, and trades workers, plant and machinery operators, labourers) compared to 'white-collar' occupations (managers, professionals, technicians, clerks) (Didham et al., 2014). Link rates also varied by region of residence in New Zealand, being lower among regions with large seasonal and semi-skilled or unskilled labour forces and above-average unemployment rates. Regional variations are, in turn, explained in large part by variation in link rates between different degrees of area-level socioeconomic deprivation. As the level of deprivation among regions increases, there is an almost uniform decline in link rates (Didham et al., 2014).

In short, there is differential attrition in the NZLC. Some groups of people are more likely to 'drop out' of the NZLC over time, meaning the characteristics of those who *are* linked across censuses differ systematically from those who are not linked, termed 'linkage bias'. Perhaps the overriding commonality among individuals not linked in the NZLC is that they are relatively residentially mobile (whether voluntarily or involuntarily), chief among whom are young people transitioning from education to work and students in short-term tenancies. Such linkage bias has the potential to introduce selection bias into statistical analyses. Section 3.6 discusses linkage bias and other limitations of the NZLC data in more detail, and what impact these may have on the results reported in this chapter.

### *3.1.2 Linking offspring with parents*

The ideal dataset for a study of intergenerational income mobility would contain complete income histories over the lifetimes of parents and their offspring. In contrast, at its maximum the NZLC spans 32 years. Consequently, a sample of offspring is selected from the earliest census available – the 1981 census – in order that these children can be observed in the later censuses (particularly 2006 and 2013) at adult ages when their current incomes are suitable for proxying their lifetime average incomes (deemed to be the mid-thirties for sons and around age 40 for daughters, discussed in detail in section 3.2.1 below). While a sample of children could be selected from, say, the 1986 or 1991 censuses, most of these children would only have reached their twenties or early thirties by the 2006 and 2013 censuses, which is likely too young for their current incomes to be reliable guides to their lifetime average incomes. Thus, the analysis focuses on New Zealand-born[66] children aged under 15 years in the 1981 census.

[66] Foreign-born children are not considered here because the focus of the analysis is on how New Zealand's social and cultural institutions affect the transmission of economic advantage and disadvantage from one generation to the next.

To conduct intergenerational analyses, children in the 1981 census need to be matched to their parents, which can be achieved by reference to variables identifying families. The 1981 census defined a family as:

> …a husband and wife with or without never married children of any age or a lone parent with one or more never married children, living in a private household. The term 'children' used here includes step children and adopted children, but not foster children. Grandchildren can also be included provided they are usually resident at the same address. Generally, if grandchildren are included, it will be when both parents are absent. A family is not necessarily all related people in a household, but only those related by blood, marriage, or adoption, who normally live together as a single family unit, as defined above, and who are present on census night. Census respondents are not asked to state family membership, this is coded from answers to other questions (Department of Statistics, 1985, p. 188).

There are several important aspects to this definition. Firstly, the definition of a family is restricted to people living in the same household, and, while not explicitly stated in the definition above, the children in the household do not have partners or children of their own living in the same household. This definition is sometimes referred to as a 'census family'. Secondly, the children can be stepchildren or adopted children, which means the 'fathers' and 'mothers' in the analysis may not be birth parents or even biological relatives, they could be step-parents or adoptive parents. Thirdly, respondents had to be present at the family dwelling on census night to be counted, or enumerated, in their family. Family members who were absent from their home on census night 1981 – identified on the census dwelling form as usually living in that dwelling but absent on the night – were not enumerated in their family, even if they filled out a census form elsewhere in New Zealand on census night (that is, absentees were not 'repatriated' back to their family or usual dwelling when the census forms were processed). As a result, there is a small proportion of families that contain children but no parents (because they were not enumerated in the family). These children cannot be paired to a parent and therefore cannot be included in the analysis.[67]

The 1981 census contains a 'Family ID number' variable (unique to each family) and a 'Role in family' variable (i.e. parent, child, visitor, etc.). Together with the 'Sex' variable, these three variables enabled the linking of sons and daughters with their fathers and mothers. Some children could be paired to both their father and mother, whereas others could only be paired to one parent, either because their other parent was missed at the 1981 census or absent from home on census night or because the child lived in a single-parent family. This means that when, for instance, sons are linked to fathers and (separately) linked to mothers, the resulting son-father and son-mother pairs *will not contain the same set of sons*; while many sons will appear in both pairs, some sons will appear in one pair but not the other. In other words, the two pairs contain *overlapping but not identical* sets of sons, and the fathers and mothers in each are not all couples (that is, they are not all spousal dyads).

The fathers and mothers identified in the 1981 census are used as the parents in this analysis. Once matched to these parents, any subsequent changes in parents over time are disregarded (that is, this definition of parents is held fixed regardless of any change in the offspring's caregivers after 1981, as in Chetty et al. (2014)). The bottom and top one percent of the

[67] A final point to note is that the definition of a family used in 1981 excludes same-sex couples with children.

distribution of fathers' and mothers' ages at the birth of their child are trimmed, which equates to dropping fathers who were younger than 17 years or older than 48 years at the birth of their child, and mothers who were younger than 16 years or older than 42 years at the birth of their child.[68]

Among the remaining offspring, sons aged zero and 1 in 1981 are simply too young to be observed in their mid-thirties at the later censuses, and similarly daughters aged zero to 6 are too young to be observed around age 40, even if these sons and daughters are linked across the full period to 2013. They are therefore not considered in the analysis, leaving sons aged 2 to 14 and daughters aged 7 to 14 in 1981 (this is elaborated in section 3.2 below). Likewise, fathers aged over 33 years and mothers aged over 38 years in 1981 are too old to have their incomes observed across multiple censuses centred around the mid-thirties for fathers and age 40 for mothers[69] (this too is elaborated in section 3.2). The data of the remaining 1981-enumerated fathers and mothers is then appended to their sons' and daughters' records in the NZLC. Since parents may have multiple children in the 1981 census, the analysis includes siblings. The end result is a target population consisting of New Zealand-born sons aged 2 to 14 and daughters aged 7 to 14 in 1981 who were home on census night (24th March) 1981 and had a father and/or a mother who (a) was also home on census night (hence was enumerated in the family), and (b) in the case of fathers was at least 19 but not older than 33 in 1981 and in the case of mothers was at least 18 but not older than 38 in 1981 (the lower age cut-off demarcating the age-at-birth-of-offspring criteria, the higher age cut-off demarcating the need to observe parents at a point in the lifecycle when their current incomes are suitable for proxying their lifetime average incomes).

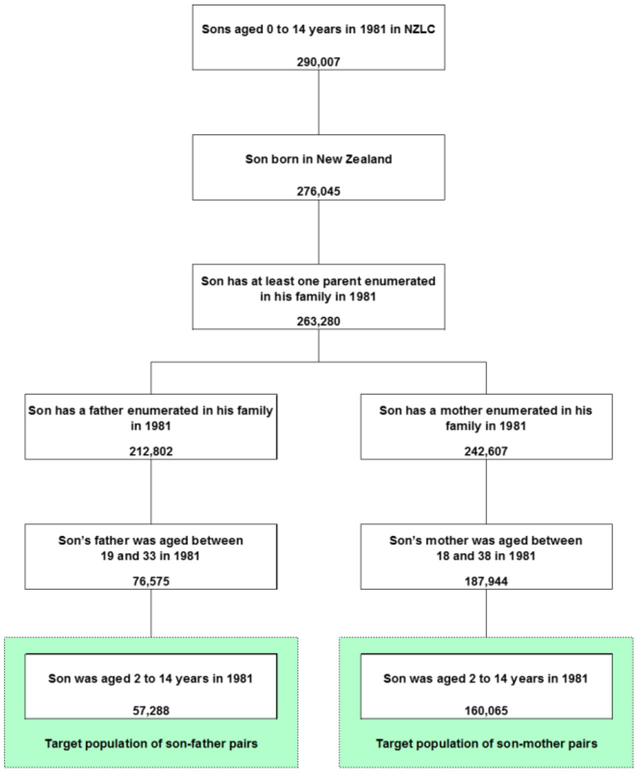
In total, there were 57,288 son-father pairs, 160,065 son-mother pairs, 86,004 daughter-mother pairs, and 18,474 daughter-father pairs that met these criteria and were therefore eligible for the analysis. It is from these target populations that the four analytical samples are drawn. Figure 3 (sons) and Figure 4 (daughters) are flowcharts showing how these populations of interest were arrived at.

As mentioned, of the population of sons aged 2 to 14 and daughters aged 7 to 14 in 1981, only those who are actually observed in the NZLC in their mid-thirties (for sons) and around age 40 (for daughters) can have their lifetime average incomes reliably estimated, which is the key sample selection criterion. The 2001 census is the earliest census at which these sons begin to reach their mid-thirties, while the 2006 census is the earliest census at which daughters reach age 40. Linkage up to at least the 2001 census for sons and the 2006 census for daughters is therefore a criterion for inclusion in the sample. Figure 5 (sons) and Figure 6 (daughters) display census linkage among the sons and daughters in the four target populations. Figure 5 shows that of the 57,288 sons in the son-father population, only 15,291 are linked through to 2001, by which point the oldest of these sons will have reached their mid-thirties (the bulk of the sons will have reached their mid-thirties even later at the 2006 census). For the son-mother population, the

[68] For example, there are some 'fathers' whose age at the birth of their child was extremely young (e.g., less than 10 years, with a minimum age of one year at the birth of their 'child'). Cases like this arise when, for example, the child's mother has repartnered with a younger male, who becomes (or identifies as) the new 'father' or father figure to the child, despite there being only a small age difference between the new father figure and the child.

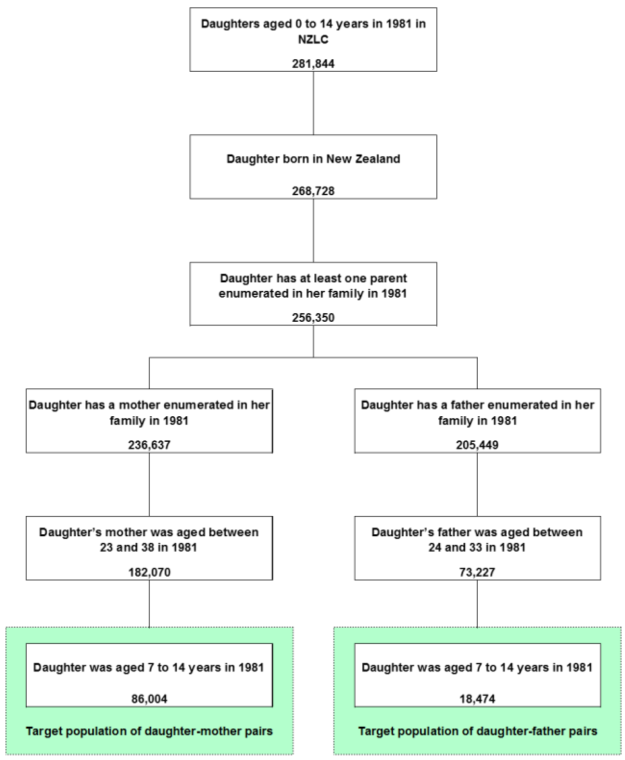
[69] Fathers aged under 19 and mothers aged under 18 in 1981 do not meet the age-at-birth-of-offspring criterion and are therefore excluded from the sample.

**Figure 3**. *Flowchart showing son-father and son-mother target populations of interest.*



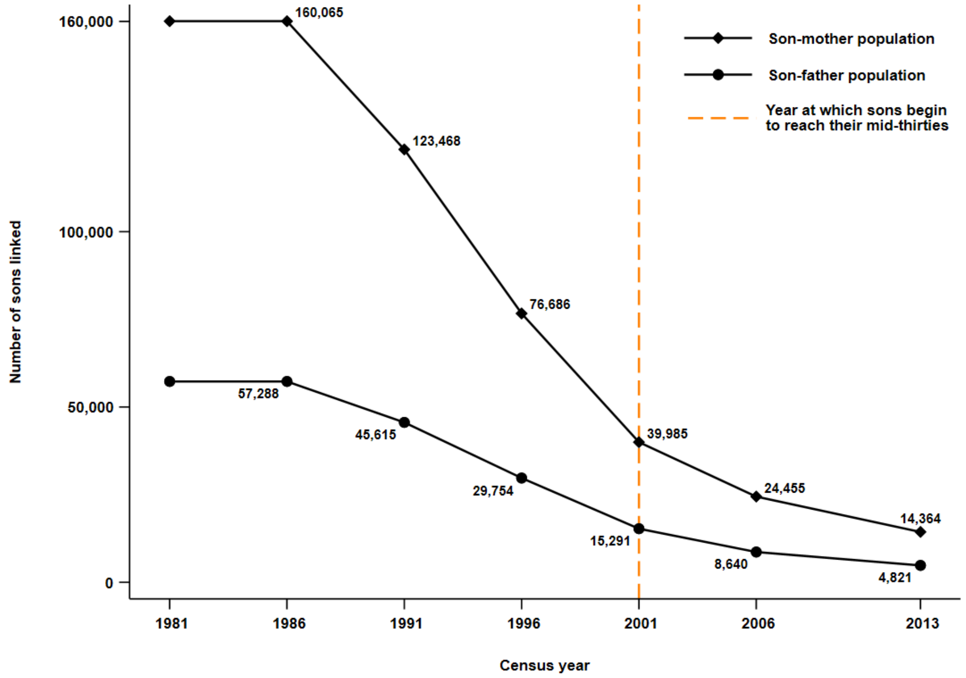

**Source:** New Zealand Longitudinal Census dataset.

**Figure 4**. *Flowchart showing daughter-mother and daughter-father target populations of interest.*



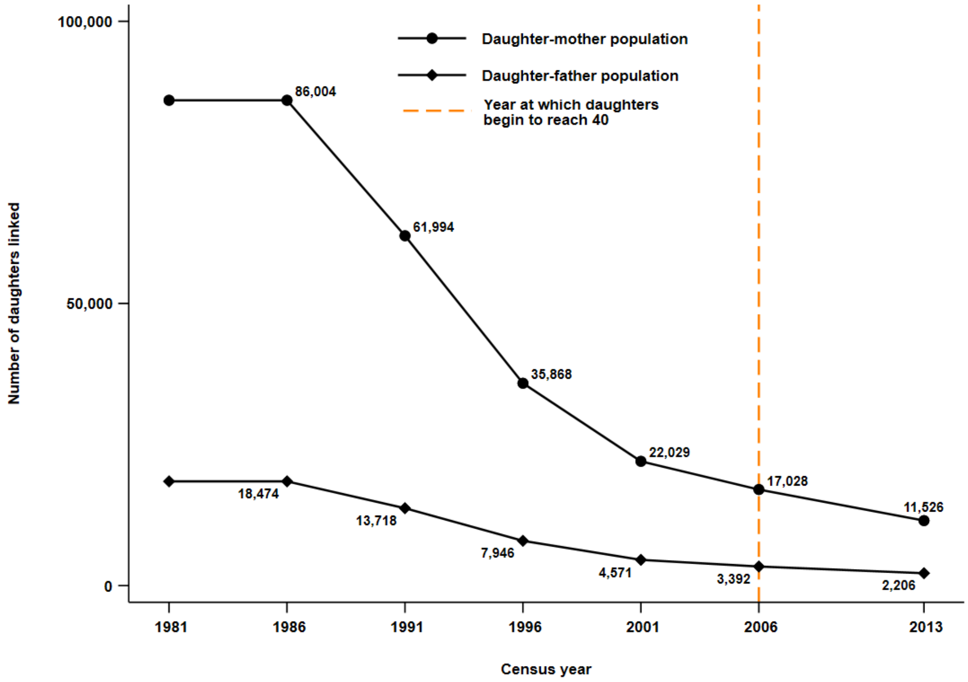

**Source:** New Zealand Longitudinal Census dataset.

**Figure 5**. *Census linkage among sons in the son-father and son-mother populations.*




**Source:** New Zealand Longitudinal Census dataset.

**Figure 6**. *Census linkage among daughters in the daughter-mother and daughter-father populations.*




**Source:** New Zealand Longitudinal Census dataset.

number of sons linked through to 2001 is 39,985. Similarly, Figure 6 shows that of the 86,004 daughters in the daughter-mother population, only 17,028 are linked through to 2006, by which point the oldest of these daughters will have reached their fortieth birthday (most of the daughters will have reached age 40 even later, at the 2013 census). For the daughter-father population the number of daughters linked through to 2006 is 3,392. Sons and daughters not linked through to 2001 and 2006, respectively, cannot be considered in the analysis. The method by which proxies for lifetime average income are constructed, and hence sample selection criteria applied to the four target populations, is returned to in section 3.2.

### *3.1.3 Income concept and measurement*

All censuses from 1981 to 2013 collected information on individuals' 'total personal income', measured as gross (before tax) annual income received from all sources for the 12 months ended 31 March of the census year. The income question is asked of all census respondents aged 15 years and over. A question about sources of income precedes the total personal income question in all these censuses, partly to prompt individuals to include these sources in their answer for personal income. See Statistics New Zealand (2009) for a precise definition of the income concept that underpins the total personal income variable[70] and Errington, Cotterell, von Randow and Milligan (2008) for a comparison of census variables over the period 1981 to 2006 including changes to the definition and question wording of the total personal income variable.[71]

Total personal income is collected in income bands rather than actual dollar amounts. That is, the data is interval-censored, with the maximum open-ended category (e.g., "$100,001 or more") censored from above and the minimum category censored from below. The bands have changed over censuses with respect to their number (ranging from 13 to 24), the amounts by which they increment, and their minimum and maximum categories. From 1981 to 1991 the lowest income category was "Nil or loss", while from 1996 onwards these were split out into separate "Loss" (negative income) and "Zero income" categories. Since 1986, Statistics New Zealand has calculated 'representative values' for each income band using data from surveys which collect more detailed income data than the census. Statistics New Zealand derives these representative values so that they can calculate income at the family and household level (by tallying the representative values of each member of the family or household). The representative value they assign to each band is the *median* value for those in that income band in the more detailed survey. Data from the Household Economic Survey were used to calculate the medians for the 1986, 1991, 1996, and 2013 censuses, tax data were used for the 2001 census, and data from the Survey of Family, Income, and Employment were used for the 2006 census. More detailed survey data for calculating band medians was not available for the 1981 census, so Statistics New Zealand has used mid-points for 1981.

[70] For example, it excludes irregular lump sum payments (such as superannuation payouts, insurance payouts, matrimonial property settlements, lottery wins, and inheritances) and money given by members of the same household to each other (intra-household transfers), exclusions that have been mentioned in the census form instructions or accompanying guide notes.

[71] For example, Errington et al. (2008) discuss changes across censuses in the treatment of fringe benefits/income-in-kind and tax credits. Overall, however, Errington et al. (2008, p. 98) conclude that the total personal income data is "broadly comparable" over the period 1981 to 2006 (and by extension, to 2013, since the question wording and guide notes were the same as in 2006).

In the current analysis, respondents at each census are assigned the representative value of their income band (the median, or mid-point for 1981) which is then adjusted for inflation by deflating to 2012 September quarter dollars using the consumer price index (CPI). The 'Nil or loss' categories used in 1981, 1986, and 1991 and the 'Zero income' categories used from 1996 onwards are allocated representative values of zero by Statistics New Zealand, and the 'Loss' (negative income) categories used from 1996 onwards are bottom-coded to representative values ranging between -$2,440 in 2001 to -$20,716 in 2013 (similarly, the maximum open-ended categories are assigned representative values that top-code the data). Appendix Table 1 presents the total personal income classifications (bands) used in each census from 1981 to 2013, the representative values assigned to each band, and the CPI-deflated representative values. Following common practice (e.g., Chetty et al., 2014, p. 1571; Gregg, Macmillan, et al., 2017, p. 88; Landersø & Heckman, 2017, p. 185), 'zero incomes' are recoded to one dollar so that when the natural logarithm of income is taken these individuals will not be eliminated from the sample (however, the robustness of the results to this treatment of zero incomes is checked). Losses are preserved at their (negative) representative values.

### 3.2 Method

#### *3.2.1 Constructing a proxy for lifetime average income*

To construct a proxy for lifetime average income (hereafter, 'proxy income') for offspring and parents in the analysis, a simple method advanced by Mazumder (2016) is used. In datasets containing multiple observations of current (e.g., annual) income, Mazumder (2016) advocates taking a *multiyear average of income observations centred on an age at which current income is known to be representative of lifetime average income* (hereafter, the 'proxy age'). Doing so, he argues, will minimise lifecycle and attenuation biases.

As discussed in section 2.3.5 of chapter 2, on the basis of the international empirical evidence estimating the age at which current income best approximates lifetime average income, proxy incomes in the current analysis are constructed as a multiyear average *centred as close as possible to age 35 for men* (sons and fathers) and *centred as close as possible to age 40 for women* (daughters and mothers).[72] Not all sons and fathers (daughters and mothers) are observed in the NZLC at exactly age 35 (age 40), hence the caveat that the proxies are centred *as close as possible* to these ages. Because not every age cohort of offspring and parents are observed at exactly their proxy age, a decision is required about how much deviation from the proxy age is acceptable. A cut-off of 'the proxy age plus or minus one year' is chosen as delimiting acceptable proxies, that is, proxy incomes for sons and fathers are permitted to be centred anywhere between ages 34 and 36 years, while those for daughters and mothers are permitted to be centred anywhere between 39 and 41 years.[73]

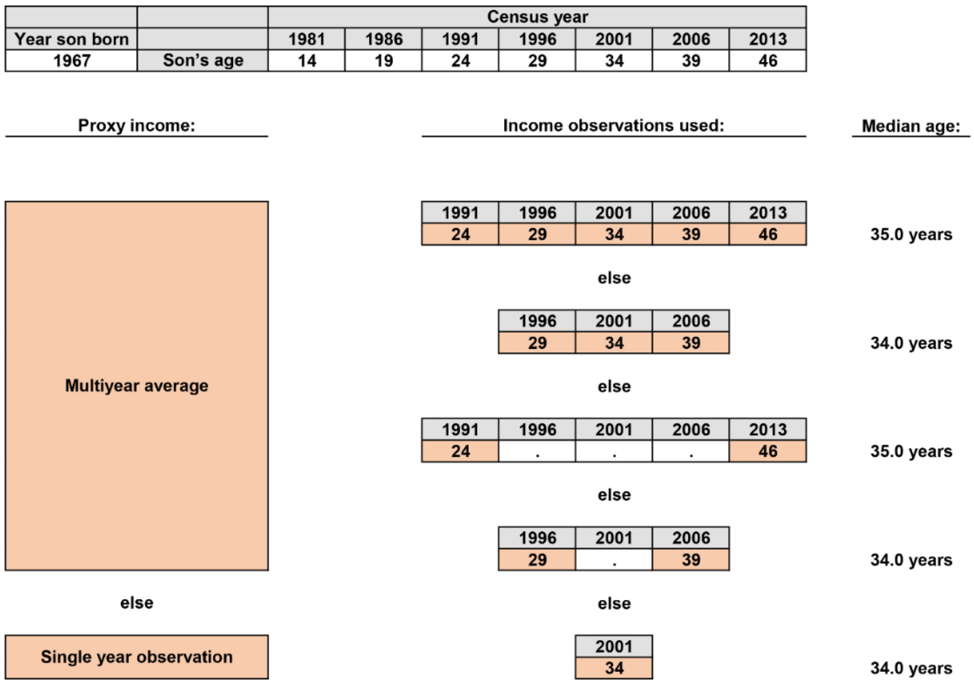
[72] While some of the studies in Table 1 suggest income observed after age 40 and up to age 50 provide the best proxies for women, using this age range is not feasible in the current analysis because it would hugely reduce the sample size of daughters, since only daughters aged 12 to 14 years in 1981 are observed in this age range (none are observed beyond 46 years of age).

[73] Proxy incomes centred within this tolerance (of one year either side of the proxy age) can be calculated for all offspring and parents in the samples, with one exception for sons and one exception for daughters: for sons aged 5 years in 1981 the proxy income is centred on age 33.5 years and for daughters aged 10 years in 1981 the proxy income is centred on

Figure 7 below uses sons aged 14 years in 1981 (born 1967) as an example of the steps followed to construct proxy incomes:

1. In the first instance, the longest consecutive multiyear average centred as close as possible to age 35 is taken, which in the example case is an average of incomes over the five censuses from 1991 to 2013 (for which the central or median age is exactly 35 years)
2. If that is not possible (because, for example, the son is only linked through to 2006), the next-longest consecutive time-average is taken, in this case over 1996 to 2006 (median age 34.0 years, which falls within the one-year tolerance).
3. Failing that, the longest possible *non-consecutive* time-average is taken, which would be an average of the 1991 and 2013 observations (where the intervening income values – from 1996 to 2006 – are missing; median age 35.0).
4. Or else the next-longest non-consecutive time-average is taken (an average of the 1996 and 2006 observations, median age 34.0).
5. If a multiyear average cannot be constructed, a single year observation is taken as the proxy income, in this case the 2001 value (observed at age 34.0).

**Figure 7**. *Construction of proxy for lifetime average income of offspring, using sons aged 14 years in 1981 as an example.*




As previously mentioned, this procedure can be applied to sons aged 2 to 14 years on census night 1981 (born between 1967 and 1979) and daughters aged 7 to 14 years on census night 1981 (born between 1967 and 1974), but not sons aged zero to 1 and daughters aged zero to 6 in 1981, who are all too young to have reached their proxy age by the latest census (even

age 38.5 years, there being no proxy income centred within the 34 to 36 and 39 to 41 age windows, respectively, for these particular cohorts of offspring. See Appendix Table 2 for details.

assuming they are linked that far) and thus cannot be included in the analysis. Figure 8 (sons) and Figure 9 (daughters) show the particular census income observations utilised in the construction of proxy incomes for each age cohort.[74] Multiyear averages are possible for most offspring (maximum of five income observations spanning 22 years), except for sons aged 2 to 4 and daughters aged 7 to 9, for whom single-year observations are the only possible proxy.

**Figure 8**. *Age cohorts of sons for which a proxy income can be constructed, and census income observations used in its calculation.*

| | | Census year | | | | | | | Eligibility for analysis |
|---|---|---|---|---|---|---|---|---|---|
| **Year son born** | | **1981** | **1986** | **1991** | **1996** | **2001** | **2006** | **2013** | |
| 1981 | Son's age | 0 | 5 | 10 | 15 | 20 | 25 | 32 | These sons *do not* have incomes observed at one or more ages which can be centred within one year of age 35. |
| 1980 | | 1 | 6 | 11 | 16 | 21 | 26 | 33 | |
| 1979 | | 2 | 7 | 12 | 17 | 22 | 27 | 34 | These cohorts of sons have incomes observed at one or more ages which can be centred within one year of age 35 (i.e. 34 to 36).[1] |
| 1978 | | 3 | 8 | 13 | 18 | 23 | 28 | 35 | |
| 1977 | | 4 | 9 | 14 | 19 | 24 | 29 | 36 | |
| 1976 | | 5 | 10 | 15 | 20 | 25 | 30 | 37 | |
| 1975 | | 6 | 11 | 16 | 21 | 26 | 31 | 38 | |
| 1974 | | 7 | 12 | 17 | 22 | 27 | 32 | 39 | |
| 1973 | | 8 | 13 | 18 | 23 | 28 | 33 | 40 | |
| 1972 | | 9 | 14 | 19 | 24 | 29 | 34 | 41 | |
| 1971 | | 10 | 15 | 20 | 25 | 30 | 35 | 42 | |
| 1970 | | 11 | 16 | 21 | 26 | 31 | 36 | 43 | |
| 1969 | | 12 | 17 | 22 | 27 | 32 | 37 | 44 | |
| 1968 | | 13 | 18 | 23 | 28 | 33 | 38 | 45 | |
| 1967 | | 14 | 19 | 24 | 29 | 34 | 39 | 46 | |

**Notes:**
[1] Calculating a proxy income within one year of the proxy age was possible for all sons except those aged 5 years in 1981, who were instead assigned a proxy within 1.5 years of their proxy age.

**Figure 9.** *Age cohorts of daughters for which a proxy income can be constructed, and census income observations used in its calculation.*

| | | Census year | | | | | | | Eligibility for analysis |
|---|---|---|---|---|---|---|---|---|---|
| **Year daughter born** | | **1981** | **1986** | **1991** | **1996** | **2001** | **2006** | **2013** | |
| 1981 | Daughter's age | 0 | 5 | 10 | 15 | 20 | 25 | 32 | These daughters *do not* have incomes observed at one or more ages which can be centred within one year of age 40. |
| 1980 | | 1 | 6 | 11 | 16 | 21 | 26 | 33 | |
| 1979 | | 2 | 7 | 12 | 17 | 22 | 27 | 34 | |
| 1978 | | 3 | 8 | 13 | 18 | 23 | 28 | 35 | |
| 1977 | | 4 | 9 | 14 | 19 | 24 | 29 | 36 | |
| 1976 | | 5 | 10 | 15 | 20 | 25 | 30 | 37 | |
| 1975 | | 6 | 11 | 16 | 21 | 26 | 31 | 38 | |
| 1974 | | 7 | 12 | 17 | 22 | 27 | 32 | 39 | These cohorts of daughters have incomes observed at one or more ages which can be centred within one year of age 40 (i.e. 39 to 41).[1] |
| 1973 | | 8 | 13 | 18 | 23 | 28 | 33 | 40 | |
| 1972 | | 9 | 14 | 19 | 24 | 29 | 34 | 41 | |
| 1971 | | 10 | 15 | 20 | 25 | 30 | 35 | 42 | |
| 1970 | | 11 | 16 | 21 | 26 | 31 | 36 | 43 | |
| 1969 | | 12 | 17 | 22 | 27 | 32 | 37 | 44 | |
| 1968 | | 13 | 18 | 23 | 28 | 33 | 38 | 45 | |
| 1967 | | 14 | 19 | 24 | 29 | 34 | 39 | 46 | |

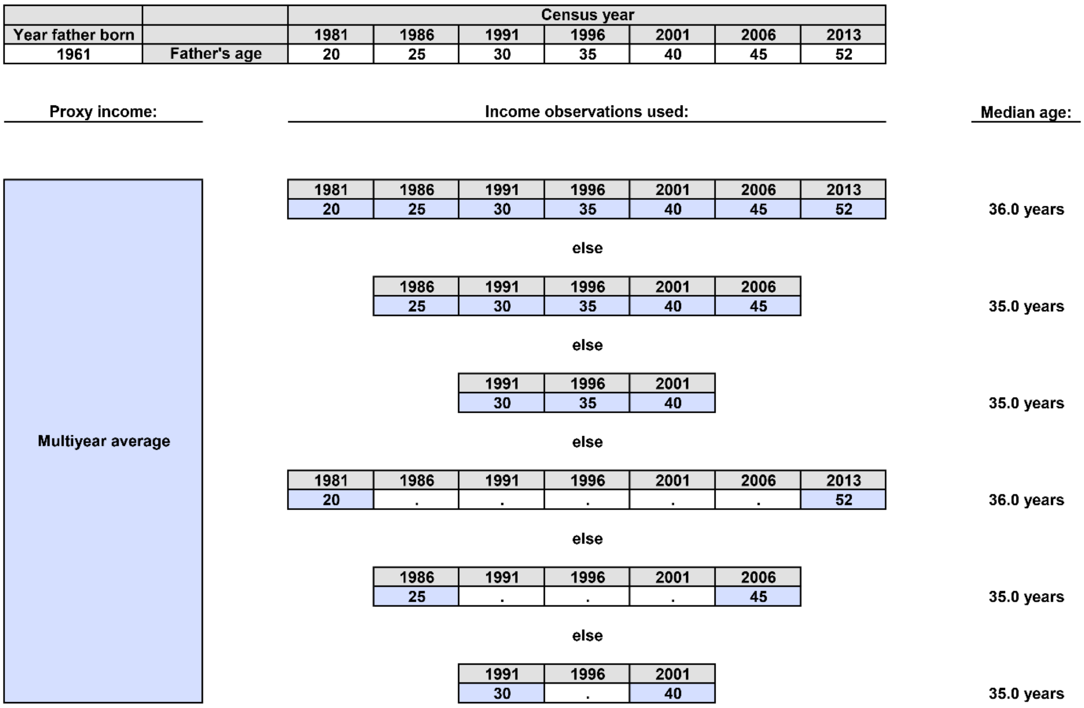
**Notes:**
[1] Calculating a proxy income within one year of the proxy age was possible for all sons except those aged 5 years in 1981, who were instead assigned a proxy within 1.5 years of their proxy age.

To calculate a proxy income for fathers and mothers, the same procedure outlined above is repeated, with one exception: the last step is omitted. That is, parents are not permitted to have single-year income observations as proxies. Permitting only multiyear averages for parents is consistent with the classic linear regression mismeasurement model in which measurement error in the right-hand side variable (in this case, parents' income) attenuates estimates of the coefficient on that variable (in this case, the IGE), whereas measurement error in the left-hand-side variable (offspring's income) results in no such attenuation (the estimate of the IGE is

[74] Note that these cohorts are defined by *age on census night 1981*, *not year of birth*. For example, of all offspring born in 1967, only those who had reached their fourteenth birthday by census night (24th March) 1981 are eligible for inclusion in the analysis.

unbiased), only less statistical precision in its estimation (e.g., Hausman, 2001). Figure 10 below uses fathers aged 20 years in 1981 (born 1961) as an example of how proxy incomes for parents are calculated.

**Figure 10**. *Construction of proxy for lifetime average income of parents, using fathers aged 20 years in 1981 as an example.*




Some parents are observed in the NZLC across an age range that is too old (beyond the proxy age) for a multiyear average income to be calculated. Figure 11 (fathers) and Figure 12 (mothers) show that construction of a proxy income is possible for fathers aged 17 to 33 in 1981 and mothers aged 16 to 38 in 1981. Fathers aged over 33 years and mothers aged over 38 years in 1981 cannot be assigned a proxy income and thus are excluded from the analysis. Fathers younger than 17 in 1981 and mothers younger than 16 in 1981 fail the age-at-birth-of-offspring criteria and thus are also excluded from the analysis. Figures 11 and 12 also show, as previously stated, that fathers and mothers are not permitted to have single-year observations as proxy incomes, hence the minimum number of observations used in their proxy incomes is two (with a maximum of seven spanning the full 32-year period). Appendix Table 2 provides details of the proxy income construction procedure, listing all multiyear averages (and single year proxies, in the case of offspring) attempted for each cohort of sons, daughters, fathers, and mothers.

*3.2.2 Samples of offspring-parent pairs*

Having established which specific age cohorts of sons, daughters, fathers, and mothers a proxy income can be constructed for, and hence are eligible for inclusion in the analysis, the samples can now be drawn from these eligible offspring and parents. The first criterion for selection into the samples is that offspring and parents must be linked across the time period represented by

**Figure 11**. *Age cohorts of fathers for which a proxy income can be constructed, and census income observations used in its calculation.*

| | | Census year | | | | | | | Eligibility for analysis |
|---|---|---|---|---|---|---|---|---|---|
| **Year father born** | | **1981** | **1986** | **1991** | **1996** | **2001** | **2006** | **2013** | |
| 1966 | Father's age | 15 | 20 | 25 | 30 | 35 | 40 | 47 | These fathers do not meet the age-at-birth-of-offspring criterion, hence are not eligible for inclusion in the analysis. |
| 1965 | | 16 | 21 | 26 | 31 | 36 | 41 | 48 | |
| 1964 | | 17 | 22 | 27 | 32 | 37 | 42 | 49 | |
| 1963 | | 18 | 23 | 28 | 33 | 38 | 43 | 50 | |
| 1962 | | 19 | 24 | 29 | 34 | 39 | 44 | 51 | These fathers do not meet the age-at-birth-of-daughter criterion, but do meet the criterion for sons, hence are eligible for inclusion in the son-father analysis but not the daughter-father analysis. |
| 1961 | | 20 | 25 | 30 | 35 | 40 | 45 | 52 | |
| 1960 | | 21 | 26 | 31 | 36 | 41 | 46 | 53 | |
| 1959 | | 22 | 27 | 32 | 37 | 42 | 47 | 54 | |
| 1958 | | 23 | 28 | 33 | 38 | 43 | 48 | 55 | |
| 1957 | | 24 | 29 | 34 | 39 | 44 | 49 | 56 | These cohorts of fathers have incomes observed at two or more ages which can be centred within one year of age 35. |
| 1956 | | 25 | 30 | 35 | 40 | 45 | 50 | 57 | |
| 1955 | | 26 | 31 | 36 | 41 | 46 | 51 | 58 | |
| 1954 | | 27 | 32 | 37 | 42 | 47 | 52 | 59 | |
| 1953 | | 28 | 33 | 38 | 43 | 48 | 53 | 60 | |
| 1952 | | 29 | 34 | 39 | 44 | 49 | 54 | 61 | |
| 1951 | | 30 | 35 | 40 | 45 | 50 | 55 | 62 | |
| 1950 | | 31 | 36 | 41 | 46 | 51 | 56 | 63 | |
| 1949 | | 32 | 37 | 42 | 47 | 52 | 57 | 64 | |
| 1948 | | 33 | 38 | 43 | 48 | 53 | 58 | 65 | |
| 1947 | | 34 | 39 | 44 | 49 | 54 | 59 | 66 | These cohorts of fathers *do not* have incomes observed at two or more ages which can be centred within one year of age 35. |
| 1946 | | 35 | 40 | 45 | 50 | 55 | 60 | 67 | |
| 1945 | | 36 | 41 | 46 | 51 | 56 | 61 | 68 | |
| 1944 | | 37 | 42 | 47 | 52 | 57 | 62 | 69 | |

**Figure 12**. *Age cohorts of mothers for which a proxy income can be constructed, and census income observations used in its calculation.*

| | | Census year | | | | | | | Eligibility for analysis |
|---|---|---|---|---|---|---|---|---|---|
| **Year mother born** | | **1981** | **1986** | **1991** | **1996** | **2001** | **2006** | **2013** | |
| 1966 | Mother's age | 15 | 20 | 25 | 30 | 35 | 40 | 47 | These mothers do not meet the age-at-birth-of-offspring criterion, hence are not eligible for inclusion in the analysis. |
| 1965 | | 16 | 21 | 26 | 31 | 36 | 41 | 48 | |
| 1964 | | 17 | 22 | 27 | 32 | 37 | 42 | 49 | |
| 1963 | | 18 | 23 | 28 | 33 | 38 | 43 | 50 | These mothers do not meet the age-at-birth-of-daughter criterion, but do meet the criterion for sons, hence are eligible for inclusion in the son-mother analysis but not the daughter-mother analysis. |
| 1962 | | 19 | 24 | 29 | 34 | 39 | 44 | 51 | |
| 1961 | | 20 | 25 | 30 | 35 | 40 | 45 | 52 | |
| 1960 | | 21 | 26 | 31 | 36 | 41 | 46 | 53 | |
| 1959 | | 22 | 27 | 32 | 37 | 42 | 47 | 54 | |
| 1958 | | 23 | 28 | 33 | 38 | 43 | 48 | 55 | These cohorts of mothers have incomes observed at two or more ages which can be centred within one year of age 40. |
| 1957 | | 24 | 29 | 34 | 39 | 44 | 49 | 56 | |
| 1956 | | 25 | 30 | 35 | 40 | 45 | 50 | 57 | |
| 1955 | | 26 | 31 | 36 | 41 | 46 | 51 | 58 | |
| 1954 | | 27 | 32 | 37 | 42 | 47 | 52 | 59 | |
| 1953 | | 28 | 33 | 38 | 43 | 48 | 53 | 60 | |
| 1952 | | 29 | 34 | 39 | 44 | 49 | 54 | 61 | |
| 1951 | | 30 | 35 | 40 | 45 | 50 | 55 | 62 | |
| 1950 | | 31 | 36 | 41 | 46 | 51 | 56 | 63 | |
| 1949 | | 32 | 37 | 42 | 47 | 52 | 57 | 64 | |
| 1948 | | 33 | 38 | 43 | 48 | 53 | 58 | 65 | |
| 1947 | | 34 | 39 | 44 | 49 | 54 | 59 | 66 | |
| 1946 | | 35 | 40 | 45 | 50 | 55 | 60 | 67 | |
| 1945 | | 36 | 41 | 46 | 51 | 56 | 61 | 68 | |
| 1944 | | 37 | 42 | 47 | 52 | 57 | 62 | 69 | |
| 1943 | | 38 | 43 | 48 | 53 | 58 | 63 | 70 | |
| 1942 | | 39 | 44 | 49 | 54 | 59 | 64 | 71 | These cohorts of mothers *do not* have incomes observed at two or more ages which can be centred within one year of age 40. |
| 1941 | | 40 | 45 | 50 | 55 | 60 | 65 | 72 | |
| 1940 | | 41 | 46 | 51 | 56 | 61 | 66 | 73 | |
| 1939 | | 42 | 47 | 52 | 57 | 62 | 67 | 74 | |

the shaded cells in Figures 8, 9, 11, and 12 (or a portion of, as illustrated in Figures 7 and 10). This is equivalent to saying that sons and fathers need to be linked through to at least age 34 and daughters and mothers linked through to at least age 39, since these are the youngest ages at which a proxy income is permitted to be centred. Continuing the example used in Figure 7, sons aged 14 years in 1981 will need to be linked through to at least the 2001 census (and have a non-missing income observation at that census) – when they will have reached their $34^{th}$ birthday – in order to be assigned a proxy income. This can be seen at the bottom of Figure 7; a son linked only to 2001 and no further will have his 2001 single-year observation assigned as his proxy income. Sons linked beyond 2001 (to either 2006 or 2013) are eligible for one of the four multiyear proxies, while sons linked only to 1996 (or earlier censuses) cannot be assigned a proxy, as they are not observed close enough to the proxy age.

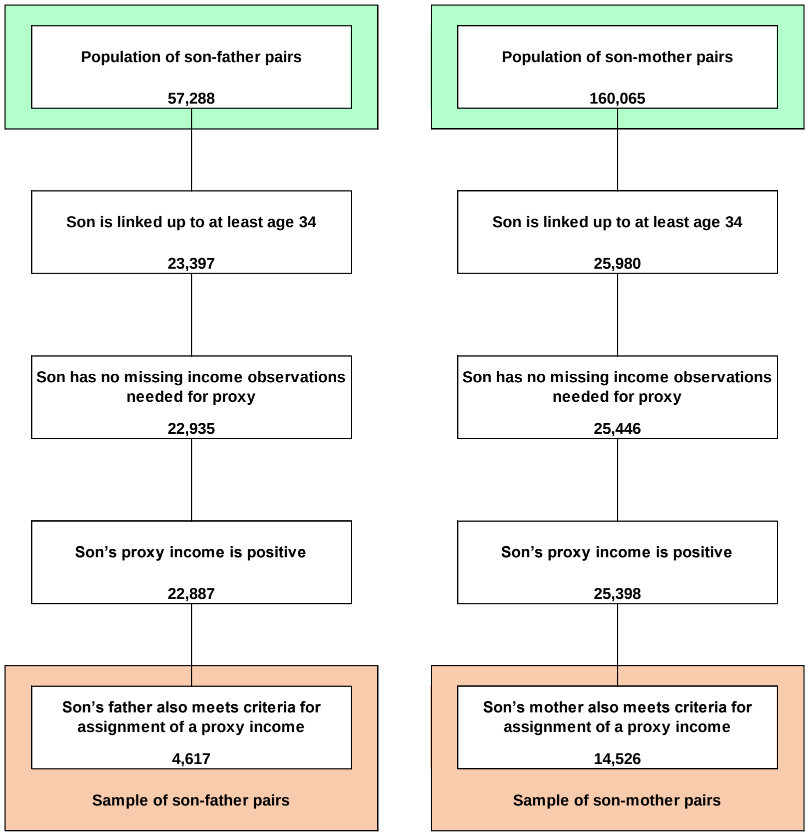
Sons who are linked through to at least age 34 and daughters linked through to at least age 39, who have no missing income observations needed for the calculation of their proxy income, are then assigned a proxy income following the procedure outlined above. Because losses (negative incomes) are valid responses in the censuses from 1996 to 2013, *it is possible for a person's proxy income to be negative*. Since the estimation of the intergenerational elasticity relies on a log-log specification, and since the natural logarithm of a negative number is undefined, any offspring whose proxy income is negative is dropped from the sample. Fathers and mothers must also be appropriately linked through the NZLC, with no missing income observations needed for the construction of proxy incomes, and with positive proxy incomes, to be included in the sample. Offspring and parents who meet all these criteria are included in the samples of offspring-parent pairs, summarised in Figure 13 (sons) and Figure 14 (daughters) below, which continue directly from Figures 3 and 4 (which showed how the target populations of interest were determined). Figure 13 shows there are 4,617 son-father pairs and 14,526 son-mother pairs who satisfy all the conditions for selection into the sample. Figure 14 shows there are 9,312 daughter-mother pairs and 1,944 daughter-father pairs who satisfy the same conditions. These are the four samples analysed here.

### 3.3 Descriptive statistics of samples

#### *3.3.1 Sociodemographic characteristics of offspring in childhood and parents in 1981*

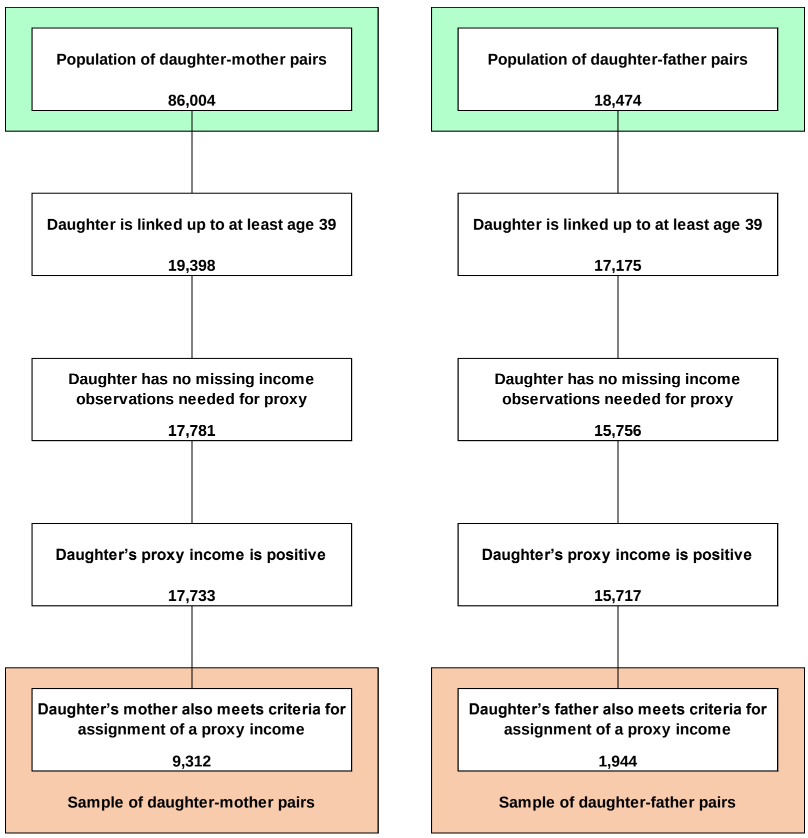
This section provides descriptive statistics of the four offspring-parent populations, their corresponding samples, and a comparison of the two. Table 3 (sons) and Table 4 (daughters) present sociodemographic characteristics of offspring, for each population and its corresponding sample, drawn from variables available in the 1981 census, as well as census enumeration and census linkage. A chi-square goodness of fit test compares the distribution of characteristics between each population and its sample; a statistically significant chi-square statistic indicates that the sample distribution differs significantly from the population distribution. Note that with sufficiently large sample sizes, such as those in the current analysis, differences between the population and sample may be *statistically* significant but small in absolute terms. Also note that the goodness of fit test cannot be performed when the number of categories containing non-zero observations differs between the population and the sample. Taking son-father pairs as an example, the *population* of sons contains a very small number of sons whose household income in 1981 was reported as 'Nil', while the corresponding *sample* contained none of these sons (or a number so small that it was suppressed on confidentiality grounds). As a result, the household income variable has 11 valid categories in the population, but only 10 valid categories in the sample, meaning the goodness of fit test cannot be performed (since it requires the same number of valid categories in each).

**Figure 13**. *Flowchart showing selection of son-father and son-mother samples.*




**Source:** New Zealand Longitudinal Census dataset.

**Figure 14**. *Flowchart showing selection of daughter-mother and daughter-father samples.*




**Source:** New Zealand Longitudinal Census dataset

Among sons, the *populations* are similar between son-father and son-mother pairs, being overwhelmingly from two-parent families, from households containing two or three children, and from families who had both parents enumerated in the family on census night 1981. Over three-quarters are of European ethnicity, about four percent Māori, two percent Pacific, and roughly 15 percent of mixed ethnic origin ('Two or more ethnic groups').[75] Majorities of the son populations lived in households in the lower-to-middle income brackets (spanning $10,000 to $24,999), with about 10 to 15 percent living in households receiving more than this, about 10 to 15 percent in households receiving less than this (virtually no households reporting no income), and a substantial minority (about 15 percent) whose household income information was missing ('Not specified').[76] However, sons in the son-father population are more heavily weighted towards the two to six year age range and have a higher proportion in two-parent families, whereas those in the son-mother population are more evenly spread across the age range and have a higher proportion living in one-parent families, reflecting the fact that mothers are more likely to be single parents (that is, there were relatively few single fathers in 1981).

The *samples* of sons closely resemble their respective populations (even if the distributions of variables were *statistically* distinguishable), with the following exceptions. Sons of European ethnicity and in the $10,000 to $15,999 household income bracket are over-represented in the samples, while sons of mixed ethnic origin and those with missing household income are under-represented. Among the son-mother sample, sons from one-parent families and from families in which only the mother was enumerated are also under-represented (these sets of families will largely overlap). All chi-square tests that were able to be performed were statistically significant, except for household composition among sons in son-father pairs where no significant difference was detected.

Among daughters, the populations are, like sons, overwhelmingly from two-parent families, from households containing two or three children, and from families who had both parents enumerated in the family on census night 1981. The ethnic make-up of the daughter populations was similar to the son populations (being 74 to 80 percent European), as were the distributions across regions and household income brackets. As with sons, daughters in the daughter-father population are more heavily drawn from the younger age groups (seven to nine years for daughters) and have a higher proportion in two-parent families, whereas those in the daughter-mother population are more evenly spread across the age range and more likely to live in one-parent families (about 10 percent, the highest of all four populations).

[75] The NZLC contains ethnic group indicator variables which are used here to code offspring's ethnicity into a 'single and combination' output (as opposed to a total response classification), which counts each respondent in *one ethnic group only*, based on the ethnic group or combination of ethnic groups they belong to. The ethnicity concept used in the 1981 census was 'ethnic origin', which the help notes defined as the 'blood mixture of races within a person', which is similar to 'race'. This concept is different to 'ethnicity', which emphasises cultural affiliation over ancestry. See Errington et al. (2008, pp. 44-47). Note that many offspring will either have had their ethnicity information filled out by their parent on their behalf (this will clearly apply to the younger age groups who could not fill out a census form on their own) or have sought their parents' guidance when responding to the ethnicity question.

[76] As household income is derived by aggregating total personal income across all members of the household aged 15 years or older, missing personal income information for any one household member will result in household income being coded as missing (unless the combined non-missing incomes fall in the highest income bracket). This means household income has a higher non-response rate than for personal income (see Statistics New Zealand, 2014).

**Table 3**. *Sociodemographic characteristics in 1981 and census linkage of sons in son-father and son-mother populations and samples.*

| Variable | Category | Son-father pairs | | | | Son-mother pairs | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Population | Sample | Chi-square goodness of fit | | Population | Sample | Chi-square goodness of fit | |
| | | % | % | $\chi^2$ | *p* | % | % | $\chi^2$ | *p* |
| Son's age in 1981 | 2 years | 14.8 | 15.5 | 168.99 | <0.001 | 8.3 | 7.4 | 1357.36 | <0.001 |
| | 3 years | 14.3 | 13.6 | | | 8.8 | 6.9 | | |
| | 4 years | 13.7 | 12.2 | | | 9.2 | 7.0 | | |
| | 5 years | 12.5 | 11.6 | | | 9.2 | 7.3 | | |
| | 6 years | 11.3 | 10.2 | | | 9.4 | 7.0 | | |
| | 7 years | 9.6 | 8.5 | | | 9.3 | 7.3 | | |
| | 8 years | 7.9 | 6.3 | | | 9.1 | 6.7 | | |
| | 9 years | 6.1 | 8.6 | | | 8.7 | 11.0 | | |
| | 10 years | 4.3 | 6.2 | | | 7.9 | 10.3 | | |
| | 11 years | 2.8 | 3.5 | | | 6.8 | 9.1 | | |
| | 12 years | 1.6 | 1.9 | | | 5.6 | 7.8 | | |
| | 13 years | 0.8 | 1.1 | | | 4.3 | 5.9 | | |
| | 14 years | 0.4 | 0.8 | | | 3.4 | 6.3 | | |
| Son's household composition in 1981 | Two-parent family[1] | 98.9 | 99.2 | 2.81 | 0.2454 | 91.6 | 94.1 | 118.17 | <0.001 |
| | One-parent family[1] | 1.0 | 0.8 | | | 8.3 | 5.9 | | |
| | Three or more families | 0.1 | S | | | 0.1 | S | | |
| Parent(s) enumerated in son's family in census 1981 | Father and mother | 91.0 | 93.2 | 27.25 | <0.001 | 79.3 | 84.2 | 209.49 | <0.001 |
| | Father only | 9.0 | 6.8 | | | n/a | n/a | | |
| | Mother only | n/a | n/a | | | 20.7 | 15.8 | | |
| Number of children aged 0 to 14 in son's household in 1981 | 1 | 7.4 | 7.5 | 39.56 | <0.001 | 7.6 | 7.8 | 77.07 | <0.001 |
| | 2 | 47.2 | 50.4 | | | 41.4 | 43.6 | | |
| | 3 | 31.7 | 31.0 | | | 33.1 | 33.0 | | |
| | 4 | 10.0 | 8.6 | | | 12.7 | 11.7 | | |
| | 5 or more | 3.7 | 2.5 | | | 5.2 | 3.9 | | |
| Number of siblings enumerated in son's family in census 1981 | 0 | 13.9 | 13.2 | 16.11 | 0.0065 | 13.9 | 12.9 | 22.94 | 0.0003 |
| | 1 | 49.4 | 51.4 | | | 44.9 | 45.9 | | |
| | 2 | 27.9 | 28.0 | | | 29.5 | 30.1 | | |
| | 3 | 7.2 | 6.0 | | | 9.2 | 8.9 | | |
| | 4 | 1.4 | 1.2 | | | 2.0 | 1.7 | | |
| | 5 or more | 0.2 | 0.2 | | | 0.5 | 0.4 | | |
| Son's ethnicity in 1981 | European only | 76.4 | 84.6 | Test not performed[4] | | 79.5 | 87.2 | 650.07 | <0.001 |
| | Māori only | 3.8 | 1.9 | | | 3.5 | 1.4 | | |
| | Pacific peoples only | 2.0 | 1.1 | | | 1.7 | 1.0 | | |
| | Asian only | 0.5 | 0.8 | | | 0.6 | 1.0 | | |
| | MELAA[2] only | S | S | | | S | S | | |
| | Other ethnicity only | 0.0 | S | | | 0.0 | S | | |
| | Two or more ethnic groups | 16.7 | 11.0 | | | 14.1 | 9.1 | | |
| | Ethnicity not stated | 0.6 | 0.6 | | | 0.6 | 0.4 | | |
| Son's region of usual residence in 1981[3] | Auckland | 21.0 | 20.9 | Test not performed[4] | | 23.2 | 23.4 | 138.65 | <0.001 |
| | Waikato | 11.3 | 10.1 | | | 10.6 | 9.4 | | |
| | Wellington | 11.3 | 12.1 | | | 11.7 | 12.7 | | |
| | Rest of North Island | 27.8 | 25.2 | | | 26.5 | 23.5 | | |
| | Canterbury | 13.7 | 15.8 | | | 13.8 | 15.9 | | |
| | Otago | 6.2 | 6.6 | | | 5.9 | 6.4 | | |
| | Rest of South Island | 8.7 | 9.2 | | | 8.3 | 8.7 | | |
| | Area outside of region | 0.0 | S | | | 0.0 | S | | |
| Son's total household income in 1981 | Nil | 0.0 | S | Test not performed[4] | | 0.2 | 0.1 | 532.72 | <0.001 |
| | $1 - $4,999 | 1.5 | 1.2 | | | 4.4 | 3.5 | | |
| | $5,000 - $9,999 | 8.9 | 7.5 | | | 9.9 | 8.7 | | |
| | $10,000 - $15,999 | 33.1 | 37.4 | | | 25.8 | 28.7 | | |
| | $16,000 - $19,999 | 17.7 | 18.8 | | | 16.3 | 18.3 | | |
| | $20,000 - $24,999 | 12.2 | 13.6 | | | 13.4 | 15.6 | | |
| | $25,000 - $29,999 | 4.7 | 4.7 | | | 6.2 | 6.9 | | |
| | $30,000 - $39,999 | 3.6 | 3.4 | | | 5.0 | 5.4 | | |
| | $40,000 - $49,999 | 1.2 | 1.2 | | | 1.7 | 1.7 | | |
| | $50,000 or over | 1.8 | 1.6 | | | 2.2 | 2.4 | | |
| | Not specified | 15.2 | 10.6 | | | 14.7 | 8.9 | | |
| Son's census linkage | 1981-1986 | 20.4 | n/a | Test not performed[4] | | 22.9 | n/a | Test not performed[4] | |
| | 1981-1991 | 27.7 | n/a | | | 29.2 | n/a | | |
| | 1981-1996 | 25.2 | n/a | | | 22.9 | n/a | | |
| | 1981-2001 | 11.6 | 0.2 | | | 9.7 | 1.4 | | |
| | 1981-2006 | 6.7 | 7.3 | | | 6.3 | 16.6 | | |
| | 1981-2013 | 8.4 | 92.5 | | | 9.0 | 82.0 | | |
| Total number of observations | | 57,288 | 4,617 | | | 160,065 | 14,526 | | |

**Symbols:**
S suppressed
n/a not applicable

**Notes:**
[1] These are two-parent and one-parent families with or without others in the household.
[2] Middle Eastern/Latin American/African.
[3] Region of usual residence based on Regional Council 2001 boundaries.
[4] The chi-square test cannot be performed when the number of categories containing non-zero observations differs between the population and the sample. In the case of census linkage, sons must be linked to at least 2001 to be selected into the sample, therefore the 'n/a' categories are equivalent to zero.

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

**Table 4**. *Sociodemographic characteristics in 1981 and census linkage of daughters in daughter-mother and daughter-father populations and samples.*

| Variable | Category | Daughter-mother pairs | | | | Daughter-father pairs | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Population | Sample | Chi-square goodness of fit | | Population | Sample | Chi-square goodness of fit | |
| | | % | % | $\chi^2$ | *p* | % | % | $\chi^2$ | *p* |
| Daughter's age in 1981 | 2 years | n/a | n/a | 68.2 | <0.001 | n/a | n/a | 7.02 | 0.4273 |
| | 3 years | n/a | n/a | | | n/a | n/a | | |
| | 4 years | n/a | n/a | | | n/a | n/a | | |
| | 5 years | n/a | n/a | | | n/a | n/a | | |
| | 6 years | n/a | n/a | | | n/a | n/a | | |
| | 7 years | 16.9 | 16.1 | | | 28.5 | 28.3 | | |
| | 8 years | 16.6 | 14.5 | | | 23.6 | 23.5 | | |
| | 9 years | 15.8 | 15.1 | | | 18.4 | 18.2 | | |
| | 10 years | 14.3 | 14.3 | | | 12.9 | 13.4 | | |
| | 11 years | 12.3 | 13.1 | | | 8.2 | 7.4 | | |
| | 12 years | 10.1 | 11.2 | | | 4.8 | 5.9 | | |
| | 13 years | 7.7 | 8.4 | | | 2.4 | 2.2 | | |
| | 14 years | 6.3 | 7.3 | | | 1.2 | 1.1 | | |
| Daughter's household composition in 1981 | Two-parent family[1] | 90.4 | 92.7 | 57.23 | <0.001 | 98.6 | 99.1 | Test not performed[4] | |
| | One-parent family[1] | 9.6 | 7.3 | | | 1.4 | 0.9 | | |
| | Three or more families | 0.0 | S | | | 0.0 | S | | |
| Parent(s) enumerated in daughter's family in census 1981 | Father and mother | 77.6 | 81.8 | 91.61 | <0.001 | 90.0 | 93.2 | 20.39 | <0.001 |
| | Father only | n/a | n/a | | | 10.0 | 6.8 | | |
| | Mother only | 22.4 | 18.2 | | | n/a | | | |
| Number of children aged 0 to 14 in daughter's household in 1981 | 1 | 6.7 | 7.1 | 115.58 | <0.001 | 4.2 | 4.3 | 25.66 | <0.001 |
| | 2 | 37.9 | 41.0 | | | 39.3 | 41.9 | | |
| | 3 | 34.4 | 35.2 | | | 35.7 | 35.0 | | |
| | 4 | 14.5 | 12.4 | | | 14.7 | 15.3 | | |
| | 5 or more | 6.5 | 4.4 | | | 6.1 | 3.5 | | |
| Number of siblings enumerated in daughter's family in census 1981 | 0 | 12.8 | 11.8 | 43.58 | <0.001 | 10.0 | 8.6 | 14.27 | 0.014 |
| | 1 | 42.5 | 44.5 | | | 44.1 | 45.5 | | |
| | 2 | 31.0 | 31.7 | | | 32.5 | 33.0 | | |
| | 3 | 10.6 | 9.6 | | | 10.7 | 11.3 | | |
| | 4 | 2.4 | 1.9 | | | 2.2 | 1.5 | | |
| | 5 or more | 0.7 | 0.5 | | | 0.5 | S | | |
| Daughter's ethnicity in 1981 | European only | 79.9 | 86.7 | Test not performed[4] | | 74.0 | 83.3 | Test not performed[4] | |
| | Māori only | 3.8 | 1.5 | | | 5.1 | 2.6 | | |
| | Pacific peoples only | 1.3 | 1.0 | | | 1.4 | 1.1 | | |
| | Asian only | 0.5 | 0.6 | | | 0.2 | S | | |
| | MELAA[2] only | S | S | | | S | S | | |
| | Other ethnicity only | S | S | | | S | S | | |
| | Two or more ethnic groups | 13.8 | 9.7 | | | 18.7 | 12.3 | | |
| | Ethnicity not stated | 0.6 | 0.5 | | | 0.7 | 0.6 | | |
| Daughter's region of usual residence in 1981[3] | Auckland | 22.6 | 21.7 | Test not performed[4] | | 21.3 | 18.0 | Test not performed[4] | |
| | Waikato | 10.7 | 8.8 | | | 11.3 | 9.1 | | |
| | Wellington | 11.5 | 12.6 | | | 10.9 | 12.5 | | |
| | Rest of North Island | 26.7 | 23.9 | | | 28.3 | 26.8 | | |
| | Canterbury | 14.1 | 16.3 | | | 13.9 | 16.9 | | |
| | Otago | 5.9 | 7.6 | | | 6.1 | 7.4 | | |
| | Rest of South Island | 8.3 | 9.2 | | | 8.2 | 9.4 | | |
| | Area outside of region | 0.0 | S | | | 0.0 | S | | |
| Daughter's total household income in 1981 | Nil | 0.2 | 0.1 | 382.25 | <0.001 | 0.1 | S | 75.98 | <0.001 |
| | $1 - $4,999 | 4.2 | 3.7 | | | 1.2 | 0.6 | | |
| | $5,000 - $9,999 | 9.4 | 8.3 | | | 7.4 | 6.8 | | |
| | $10,000 - $15,999 | 22.8 | 26.2 | | | 29.7 | 33.6 | | |
| | $16,000 - $19,999 | 16.3 | 19.5 | | | 18.9 | 22.7 | | |
| | $20,000 - $24,999 | 15.1 | 16.7 | | | 15.3 | 16.2 | | |
| | $25,000 - $29,999 | 7.4 | 7.4 | | | 6.0 | 5.1 | | |
| | $30,000 - $39,999 | 5.9 | 5.9 | | | 4.0 | 3.4 | | |
| | $40,000 - $49,999 | 1.9 | 1.7 | | | 1.3 | 1.1 | | |
| | $50,000 or over | 2.4 | 1.9 | | | 1.8 | 1.2 | | |
| | Not specified | 14.6 | 8.5 | | | 14.5 | 9.4 | | |
| Daughter's census linkage | 1981-1986 | 27.9 | n/a | Test not performed[4] | | 25.7 | n/a | Test not performed[4] | |
| | 1981-1991 | 30.4 | n/a | | | 31.2 | n/a | | |
| | 1981-1996 | 16.1 | n/a | | | 18.3 | n/a | | |
| | 1981-2001 | 5.8 | n/a | | | 6.4 | n/a | | |
| | 1981-2006 | 6.4 | S | | | 6.4 | S | | |
| | 1981-2013 | 13.4 | 100.0 | | | 11.9 | 100.0 | | |
| Total number of observations | | 86,004 | 9,312 | | | 18,474 | 1,944 | | |

**Symbols:**
S suppressed
n/a not applicable

**Notes:**
[1] These are two-parent and one-parent families with or without others in the household.
[2] Middle Eastern/Latin American/African.
[3] Region of usual residence based on Regional Council 2001 boundaries.
[4] The chi-square test cannot be performed when the number of categories containing non-zero observations differs between the population and the sample. In the case of census linkage, daughters must be linked to at least 2006 to be selected into the sample, therefore the 'n/a' categories are equivalent to zero.

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

Among the samples of daughters, we see the same groups under-represented as for sons: daughters of European ethnic origin and in the $10,000 to $15,999 household income bracket are over-represented in the samples, while daughters of mixed ethnic origin, from one-parent families, from families in which only the mother was enumerated, and from households with missing household income, are all under-represented. All chi-square tests that could be performed were statistically significant, except for daughters' age among daughter-father pairs where no significant difference was detected.

Overall, across all four offspring-parent pairs, the four samples consistently over-represent offspring who are European, from families in which both parents were enumerated, and with household income in the $10,000 to $15,999 bracket, and consistently under-represent offspring of mixed ethnicity, from families in which only one parent was enumerated, and with missing household income.

Turning to parent characteristics, Table 5 (fathers) and Table 6 (mothers) present sociodemographic characteristics of parents, for each population and its corresponding sample, drawn from variables available in the 1981 census, as well as census enumeration and census linkage. Majorities of the father populations were aged 30 to 33 in 1981 (born 1948 to 1951) and aged 23 to 28 at the birth of their son and 21 to 26 at the birth of their daughter. Approximately half of fathers had no formal educational qualifications in 1981, about one-quarter had a trade qualification, and about half were employed in production and labouring occupations (the next most common occupation being agricultural work). About three percent of fathers received a (non-universal) social security benefit in 1981.

The samples of fathers closely resemble their corresponding populations, but older fathers (aged 32 to 33 in 1981) and those linked over the full 1981-to-2013 period are over-represented, while younger fathers (aged 19 to 29 in 1981) and those linked over two or three censuses (i.e., up to 1991 only) are under-represented.

Majorities of mothers were aged 31 to 36 in 1981 among the son-mother population and 33 to 38 among the daughter-mother population (thus, born between 1943 and 1950) and gave birth to their sons and daughters between ages 21 and 28. As with fathers, about half of mothers had no formal qualifications, 17 to 19 percent had the lowest ($5^{th}$ form/Year 11) school qualification, and about 15 percent had a trade qualification. Just over half of mothers in the son-mother population, and just under half in the daughter-mother population, had no occupation, and among those who did have an occupation, the most common was clerical work. About eight percent of mothers were receiving a social security benefit in 1981.

The samples of mothers closely resemble their corresponding populations, but older mothers (33 to 38 in 1981) and those linked over the full 1981-to-2013 period are over-represented, while younger mothers (aged 18 to 30) and those linked over two or three censuses are under-represented.

**Table 5**. Sociodemographic characteristics in 1981 and census linkage of fathers in son-father and daughter-father populations and samples.

| Variable | Category | Son-father pairs | | | | Daughter-father pairs | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Population | Sample | Chi-square goodness of fit | | Population | Sample | Chi-square goodness of fit | |
| | | % | % | $\chi^2$ | *p* | % | % | $\chi^2$ | *p* |
| Father's age group in 1981 | 19-25 years | 8.5 | 4.3 | 258.68 | <0.001 | 1.3 | 0.8 | 60.03 | <0.001 |
| | 26-27 years | 11.1 | 8.6 | | | 5.7 | 3.8 | | |
| | 28-29 years | 18.9 | 16.6 | | | 15.2 | 11.8 | | |
| | 30-31 years | 27.0 | 26.1 | | | 29.3 | 27.2 | | |
| | 32-33 years | 34.5 | 44.4 | | | 48.4 | 56.4 | | |
| Father's age group at offspring's birth | 17-18 years | 2.8 | 1.8 | 61.73 | <0.001 | 5.9 | 4.3 | 31.69 | <0.001 |
| | 19-20 years | 8.8 | 6.7 | | | 17.5 | 13.7 | | |
| | 21-22 years | 15.9 | 16.3 | | | 28.5 | 30.0 | | |
| | 23-24 years | 21.6 | 22.7 | | | 30.1 | 32.4 | | |
| | 25-26 years | 22.4 | 21.6 | | | 18.0 | 19.7 | | |
| | 27-28 years | 17.5 | 17.5 | | | n/a | n/a | | |
| | 29-31 years | 11.2 | 13.4 | | | n/a | n/a | | |
| Father's highest educational qualification in 1981 | No school qualification | 44.0 | 41.8 | 38.51 | <0.001 | 50.7 | 49.9 | 15.68 | 0.0471 |
| | Qualification in 5th form/Year 11 | 12.9 | 13.0 | | | 11.8 | 12.2 | | |
| | Qualification in the 6th Form/Year 12 | 6.1 | 5.9 | | | 4.8 | 4.1 | | |
| | Qualification in the 7th Form/Year 13 | 2.3 | 2.0 | | | 1.9 | 1.7 | | |
| | Other school qualification | 0.1 | S | | | 0.1 | S | | |
| | Trade certificate/vocational qualification | 26.0 | 29.3 | | | 24.6 | 27.2 | | |
| | Undergraduate degree | 5.3 | 5.3 | | | 3.4 | 2.9 | | |
| | Postgraduate degree/higher degree | 1.2 | 1.3 | | | 0.8 | 0.8 | | |
| | Unidentifiable or not specified | 2.0 | 1.3 | | | 1.9 | 1.1 | | |
| Father's occupation in 1981[1] | Professional, technical and related workers | 10.5 | 11.2 | 105.74 | <0.001 | 7.9 | 8.4 | 40.25 | <0.001 |
| | Administrative and managerial workers | 4.5 | 5.2 | | | 4.7 | 5.2 | | |
| | Clerical and related workers | 5.9 | 7.1 | | | 5.5 | 6.4 | | |
| | Sales workers | 8.4 | 8.8 | | | 8.4 | 9.4 | | |
| | Service workers | 5.4 | 4.7 | | | 5.8 | 5.6 | | |
| | Agriculture and other workers[2] | 13.9 | 10.4 | | | 12.9 | 8.7 | | |
| | Production and other workers[3] | 48.8 | 51.4 | | | 52.3 | 54.7 | | |
| | Not elsewhere included[4] | 2.6 | 1.2 | | | 2.5 | 1.6 | | |
| Father received a non-universal social security benefit in 1981[5] | Did not receive a benefit | 96.5 | 97.1 | 8.27 | 0.016 | 96.4 | 97.5 | 6.46 | 0.0396 |
| | Received a benefit | 3.2 | 2.4 | | | 3.2 | 2.2 | | |
| | Benefits undefined or not specified | 0.3 | 0.4 | | | 0.4 | 0.3 | | |
| Father's census linkage | 1981-1986 | 18.2 | 3.0 | 1171.08 | <0.001 | 18.6 | 5.6 | 276.11 | <0.001 |
| | 1981-1991 | 16.3 | 10.3 | | | 16.9 | 14.0 | | |
| | 1981-1996 | 12.1 | 10.9 | | | 12.6 | 13.3 | | |
| | 1981-2001 | 9.3 | 10.1 | | | 9.6 | 10.8 | | |
| | 1981-2006 | 11.6 | 15.0 | | | 11.0 | 14.0 | | |
| | 1981-2013 | 32.5 | 50.8 | | | 31.4 | 42.4 | | |
| Total number of observations | | 43,608 | 4,416 | | | 15,435 | 1,887 | | |

**Symbols**:
S suppressed
n/a not applicable

**Notes**:
[1] Occupations are coded to level one of the *New Zealand Standard Classification of Occupations 1968*. Note that NZSCO68 is structurally different to, and thus cannot be easily compared with, subsequent versions of NZSCO.
[2] The full category label is 'Agriculture, animal husbandry and forest workers, fishermen and hunters'.
[3] The full category label is 'Production and related workers, transport equipment operators and labourers'.
[4] This category includes the non-valid responses of unidentifiable, not specified, or missing (respondent had no occupation).
[5] This variable denotes whether the father received at least one social security benefit excluding 'Family Benefit' and 'National Superannuation' during the year ended 31 March 1981.

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

Table 7 reports, for each population and its corresponding sample, the mean and standard deviation of age in 1981 of offspring and parents, and age of parents at the birth of their offspring. The average age of sons in 1981 is 5.8 in the son-father sample and 8.1 in the son-mother sample, while the average age of daughters is 10.0 in the daughter-mother sample and 8.8 in the daughter-father sample. Given sons aged zero to 1 year and daughters aged zero to 6 years are excluded from all samples, it is not surprising that in all four samples the mean age of offspring is higher than in their corresponding populations (except for daughter-father pairs where the two are similar). On average, fathers were aged about 31 in 1981 and about 25 at the birth of their sons and 22 at the birth of their daughters, while mothers were aged about 33 in 1981 and about 25 at the birth of their sons and 24 at the birth of their daughters. Fathers and mothers in the samples are slightly older than their corresponding populations.

**Table 6**. *Sociodemographic characteristics in 1981 and census linkage of mothers in daughter-mother and son-mother populations and samples.*

| Variable | Category | Son-mother pairs | | | | Daughter-mother pairs | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Population | Sample | Chi-square goodness of fit | | Population | Sample | Chi-square goodness of fit | |
| | | % | % | $\chi^2$ | $p$ | % | % | $\chi^2$ | $p$ |
| Mother's age group in 1981 | 18-24 years | 6.6 | 2.8 | 1049.92 | <0.001 | 0.6 | 0.3 | 288.78 | <0.001 |
| | 25-26 years | 7.5 | 4.3 | | | 2.9 | 1.3 | | |
| | 27-28 years | 10.7 | 8.1 | | | 6.7 | 4.3 | | |
| | 29-30 years | 14.0 | 11.8 | | | 11.6 | 9.2 | | |
| | 31-32 years | 16.4 | 16.9 | | | 17.4 | 17.1 | | |
| | 33-34 years | 17.7 | 20.9 | | | 22.1 | 24.0 | | |
| | 35-36 years | 14.9 | 18.4 | | | 20.7 | 22.3 | | |
| | 37-38 years | 12.2 | 16.8 | | | 18.1 | 21.5 | | |
| Mother's age group at offspring's birth | 16-18 years | 5.3 | 3.0 | 286.71 | <0.001 | 6.3 | 3.7 | 178.04 | <0.001 |
| | 19-20 years | 10.3 | 8.4 | | | 12.5 | 10.1 | | |
| | 21-22 years | 16.1 | 15.9 | | | 19.4 | 19.2 | | |
| | 23-24 years | 20.8 | 24.1 | | | 24.1 | 26.3 | | |
| | 25-26 years | 19.6 | 21.2 | | | 20.5 | 22.8 | | |
| | 27-28 years | 14.2 | 14.8 | | | 11.9 | 12.4 | | |
| | 29-30 years | 8.3 | 7.6 | | | 4.8 | 4.9 | | |
| | 31-32 years | 3.8 | 3.4 | | | 0.6 | 0.6 | | |
| | 33-34 years | 1.4 | 1.3 | | | n/a | n/a | | |
| | 35-36 years | 0.3 | 0.4 | | | n/a | n/a | | |
| Mother's highest educational qualification in 1981 | No school qualification | 50.5 | 51.5 | 25.51 | 0.0013 | 56.4 | 56.9 | 37.64 | <0.001 |
| | Qualification in 5th form/Year 11 | 19.0 | 19.7 | | | 17.1 | 16.0 | | |
| | Qualification in the 6th Form/Year 12 | 6.4 | 6.1 | | | 5.3 | 4.8 | | |
| | Qualification in the 7th Form/Year 13 | 1.3 | 1.0 | | | 1.2 | 0.8 | | |
| | Other school qualification | 0.6 | 0.6 | | | 0.7 | 0.8 | | |
| | Trade certificate/vocational qualification | 16.7 | 14.5 | | | 14.7 | 12.0 | | |
| | Undergraduate degree | 2.2 | 2.2 | | | 1.6 | 1.5 | | |
| | Postgraduate degree/higher degree | 0.4 | 0.4 | | | 0.3 | 0.2 | | |
| | Unidentifiable or not specified | 2.8 | 2.3 | | | 2.9 | 1.8 | | |
| Mother's occupation in 1981[1] | Professional, technical and related workers | 7.2 | 7.3 | 109.21 | <0.001 | 8.1 | 7.9 | 91.05 | <0.001 |
| | Administrative and managerial workers | 0.4 | 0.4 | | | 0.5 | 0.4 | | |
| | Clerical and related workers | 12.2 | 14.8 | | | 15.0 | 17.7 | | |
| | Sales workers | 5.8 | 6.2 | | | 7.5 | 7.9 | | |
| | Service workers | 9.4 | 9.4 | | | 11.1 | 12.0 | | |
| | Agriculture and other workers[2] | 5.5 | 5.3 | | | 6.0 | 4.8 | | |
| | Production and other workers[3] | 6.4 | 6.8 | | | 8.2 | 8.7 | | |
| | Not elsewhere included[4] | 53.1 | 49.8 | | | 43.5 | 40.7 | | |
| Mother received a non-universal social security benefit in 1981[5] | Did not receive a benefit | 92.1 | 95.1 | 161.51 | <0.001 | 91.9 | 93.9 | 46.04 | <0.001 |
| | Received a benefit | 7.6 | 4.8 | | | 7.8 | 5.9 | | |
| | Benefits undefined or not specified | 0.3 | 0.2 | | | 0.3 | 0.2 | | |
| Mother's census linkage | 1981-1986 | 14.3 | 1.1 | 3758.14 | <0.001 | 14.1 | 1.5 | 1896.73 | <0.001 |
| | 1981-1991 | 14.1 | 6.4 | | | 14.7 | 8.4 | | |
| | 1981-1996 | 11.5 | 9.0 | | | 11.4 | 10.7 | | |
| | 1981-2001 | 9.1 | 9.7 | | | 9.1 | 10.0 | | |
| | 1981-2006 | 12.5 | 16.4 | | | 12.3 | 15.0 | | |
| | 1981-2013 | 38.5 | 57.5 | | | 38.5 | 54.5 | | |
| Total number of observations | | 116,622 | 13,641 | | | 67,506 | 8,892 | | |

**Symbols**:
S suppressed
n/a not applicable

**Notes**:
[1] Occupations are coded to level one of the *New Zealand Standard Classification of Occupations 1968*. Note that NZSCO68 is structurally different to, and thus cannot be easily compared with, subsequent versions of NZSCO.
[2] The full category label is 'Agriculture, animal husbandry and forest workers, fishermen and hunters'.
[3] The full category label is 'Production and related workers, transport equipment operators and labourers'.
[4] This category includes the non-valid responses of unidentifiable, not specified, or missing (respondent had no occupation).
[5] This variable denotes whether the father received at least one social security benefit excluding 'Family Benefit' and 'National Superannuation' during the year ended 31 March 1981.

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

**Table 7**. *Mean and standard deviation of age in 1981 (in years) of offspring and parents, and average age of parents at birth of offspring, in offspring-parent populations and samples.*

| | Son-father pairs | | | | Son-mother pairs | | | |
|---|---|---|---|---|---|---|---|---|
| | Population (n=57,288) | | Sample (n=4,617) | | Population (n=160,062) | | Sample (n=14,526) | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Age of offspring in 1981 | 5.5 | 2.8 | 5.8 | 3.0 | 7.2 | 3.4 | 8.1 | 3.6 |
| Age of parent in 1981 | 30.0 | 2.7 | 30.5 | 2.4 | 31.7 | 4.1 | 32.7 | 3.7 |
| Age of parent at offspring's birth | 24.5 | 3.1 | 24.8 | 3.1 | 24.4 | 3.7 | 24.6 | 3.4 |
| | **Daughter-mother pairs** | | | | **Daughter-father pairs** | | | |
| | **Population (n=86,004)** | | **Sample (n=9,312)** | | **Population (n=18,477)** | | **Sample (n=1,944)** | |
| | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** |
| Age of offspring in 1981 | 9.8 | 2.1 | 10.0 | 2.2 | 8.8 | 1.7 | 8.8 | 1.7 |
| Age of parent in 1981 | 33.3 | 3.2 | 33.8 | 3.0 | 31.0 | 1.9 | 31.3 | 1.8 |
| Age of parent at offspring's birth | 23.5 | 3.1 | 23.8 | 2.9 | 22.2 | 2.2 | 22.5 | 2.1 |

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

*3.3.2 Sociodemographic characteristics of offspring in adulthood*

A more contemporary picture of the offspring in the samples is now provided, presenting a selection of their sociodemographic characteristics as adults. Some of these adult attributes are used in the estimation of intergenerational persistence by groups (section 3.4.3) and to analyse transmission pathways that mediate the association between parent's income and offspring's income (chapter 5). Table 8 (sons) and Table 9 (daughters) present offspring's region of residence, ethnic group, highest educational qualification, and occupation, as reported in *the latest available census (where the response is non-missing) for each offspring*. The tables also present their parent's highest educational qualification, as reported in the latest available census for each parent (used to estimate group-specific IGEs in section 3.4.3). As in Tables 3 and 4, ethnic group is coded to a 'single and combination' output, counting each respondent in one ethnic group only based on the ethnic group or combination of ethnic groups they belong to. Highest educational qualification is coded to Errington and colleagues' (2008) nine-category classification harmonised across the 1981 to 2006 censuses, ranging from 'no school qualification' up to 'postgraduate or higher degree'. Occupational groups are coded to level 1 of the *New Zealand Standard Classification of Occupations 1999* (Statistics New Zealand, 2001) for the 2001, 2006, and 2013 censuses and the *New Zealand Standard Classification of Occupations 1995* (Statistics New Zealand, 1995) for the 1996 census, the level 1 output categories being the same for each.

Tables 8 and 9 shows that, by adulthood, the majority of offspring are living in the North Island, mainly Auckland and other parts of the North Island outside of Waikato and Wellington. In aggregate, the proportions of offspring in each region by adulthood are very similar to those for 1981. Whereas in 1981 the ethnicity of many offspring would have been reported by their parent, by adulthood these same offspring were self-identifying their own ethnicity in the census. The vast majority (about 80 percent) of offspring still identify as European (only), but this proportion has decreased a little since 1981, when about 85 percent of offspring were identified as European. In contrast, the proportion of offspring identifying as Māori has more than doubled since 1981 (e.g., in 1981 only two percent of sons in son-father pairs were identified as Māori, but by adulthood about five percent of these sons now identified as Māori only).

By adulthood, approximately one-third of sons have a school qualification (attained in years 11 to 13) as their highest qualification and another third have a trade certificate or other vocational qualification as their highest qualification. About 20 percent have a degree, and the remainder – about 13 percent – have no qualifications. Sons are spread widely across occupational categories. The majority are in 'white-collar' jobs, particularly as managers (about 20 percent), professionals (about 14 percent), and technicians (12 percent), although trades workers are the second most common occupation (17 percent). Relatively small proportions of sons are working in clerical occupations (about five percent) or elementary occupations (about six percent).

For daughters, roughly 45 percent have attained only a school qualification by adulthood, 20 percent have a trade qualification, about one-quarter have a degree, and the remaining 10 percent have no qualifications. Daughters are heavily concentrated in white-collar occupations (approximately 90 percent), particularly as professionals (about 21 percent) or clerks (21 percent).

**Table 8**. *Sons' adult characteristics and their parents' highest educational qualification.*

| Variable | Category | Sample | |
|---|---|---|---|
| | | Son-father | Son-mother |
| | | % | % |
| Son's region of usual residence (at latest census)[1,2] | Auckland | 23.4 | 25.6 |
| | Waikato | 10.3 | 9.5 |
| | Wellington | 12.9 | 13.8 |
| | Rest of North Island | 22.1 | 20.9 |
| | Canterbury | 17.0 | 16.7 |
| | Otago | 6.0 | 5.9 |
| | Rest of South Island | 8.3 | 7.6 |
| | Area outside region | S | S |
| Son's ethnic group (at latest census)[1] | European only | 78.9 | 79.8 |
| | Māori only | 4.8 | 4.0 |
| | Pacific peoples only | 1.4 | 1.3 |
| | Asian only | 1.6 | 1.6 |
| | MELAA[3] only | 0.1 | 0.1 |
| | Other ethnicity only | 3.6 | 5.2 |
| | Two or more ethnic groups | 9.3 | 7.7 |
| | Ethnicity not stated | 0.3 | 0.4 |
| Son's highest educational qualification (at latest census)[1] | No school qualification | 13.2 | 13.4 |
| | Qualification in 5th form/Year 11 | 14.0 | 12.9 |
| | Qualification in the 6th Form/Year 12 | 12.7 | 12.6 |
| | Qualification in the 7th Form/Year 13 | 8.3 | 7.1 |
| | Overseas school qualification | 0.6 | 0.5 |
| | Other school qualification | 0.0 | 0.0 |
| | Trade certificate/vocational qualification | 32.0 | 33.7 |
| | Undergraduate degree | 14.0 | 14.0 |
| | Postgraduate degree/higher degree | 5.1 | 5.7 |
| Son's occupation (at latest census)[1] | Legislators, administrators and managers | 18.8 | 19.4 |
| | Professionals | 13.7 | 14.9 |
| | Technicians and associate professionals | 12.3 | 12.0 |
| | Clerks | 4.8 | 5.2 |
| | Service and sales workers | 7.4 | 7.7 |
| | Agriculture and fishery workers | 7.5 | 7.1 |
| | Trades workers | 17.9 | 17.1 |
| | Plant and machine operators and assemblers | 11.8 | 10.7 |
| | Elementary occupations | 5.8 | 5.9 |
| Parent's highest educational qualification (at latest census)[1] | No school qualification | 37.3 | 44.8 |
| | Qualification in 5th form/Year 11 | 12.9 | 20.3 |
| | Qualification in the 6th Form/Year 12 | 5.1 | 5.5 |
| | Qualification in the 7th Form/Year 13 | 2.5 | 1.7 |
| | Overseas school qualification | 1.0 | 2.3 |
| | Other school qualification | S | 0.2 |
| | Trade certificate/vocational qualification | 33.7 | 18.8 |
| | Undergraduate degree | 4.9 | 4.3 |
| | Postgraduate degree/higher degree | 2.7 | 2.1 |
| Total number of observations | | 4,617 | 14,526 |

**Symbols**:
S suppressed

**Notes**:
[1] Responses are those from each son's/parent's latest available census where the response is non-missing.
[2] Region of usual residence based on Regional Council 2006 boundaries.
[3] Middle Eastern/Latin American/African.

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

Very few are working in trades (one percent) or other manual occupations. The distribution of educational qualifications among offspring contrasts starkly with that of parents, reflecting secular increases in educational attainment. Between 40 and 45 percent of fathers have no qualifications, about 20 percent have a school qualification, about one-third have a trade qualification, and the remainder (between four to eight percent) have a degree. Roughly 45 to 50 percent of mothers have no qualifications, about 30 percent have a school qualification, 18 percent have a trade qualification, and the remaining five to six percent have a degree.

**Table 9**. *Daughters' adult characteristics and their parents' highest educational qualification.*

| Variable | Category | Sample | |
|---|---|---|---|
| | | Daughter-mother | Daughter-father |
| | | % | % |
| Daughter's region of usual residence (at latest census)[1,2] | Auckland | 22.7 | 19.1 |
| | Waikato | 9.1 | 10.0 |
| | Wellington | 13.3 | 12.0 |
| | Rest of North Island | 22.8 | 25.3 |
| | Canterbury | 16.7 | 17.7 |
| | Otago | 7.0 | 7.4 |
| | Rest of South Island | 8.4 | 8.3 |
| | Area outside region | S | S |
| Daughter's ethnic group (at latest census)[1] | European only | 83.8 | 80.7 |
| | Māori only | 3.6 | 5.2 |
| | Pacific peoples only | 1.4 | 1.5 |
| | Asian only | 1.1 | 0.8 |
| | MELAA[3] only | 0.1 | S |
| | Other ethnicity only | 1.6 | 1.7 |
| | Two or more ethnic groups | 8.2 | 10.0 |
| | Ethnicity not stated | 0.1 | S |
| Daughter's highest educational qualification (at latest census)[1] | No school qualification | 9.9 | 10.2 |
| | Qualification in 5th form/Year 11 | 17.5 | 17.0 |
| | Qualification in the 6th Form/Year 12 | 19.7 | 19.6 |
| | Qualification in the 7th Form/Year 13 | 5.7 | 8.0 |
| | Overseas school qualification | 0.3 | 0.5 |
| | Other school qualification | 0.0 | 0.0 |
| | Trade certificate/vocational qualification | 22.2 | 22.1 |
| | Undergraduate degree | 18.1 | 17.6 |
| | Postgraduate degree/higher degree | 6.5 | 5.1 |
| Daughter's occupation (at latest census)[1] | Legislators, administrators and managers | 17.4 | 17.1 |
| | Professionals | 22.3 | 20.2 |
| | Technicians and associate professionals | 16.5 | 16.3 |
| | Clerks | 21.2 | 21.2 |
| | Service and sales workers | 13.5 | 13.9 |
| | Agriculture and fishery workers | 3.3 | 3.7 |
| | Trades workers | 0.9 | 1.1 |
| | Plant and machine operators and assemblers | 1.5 | 2.4 |
| | Elementary occupations | 3.3 | 4.0 |
| Parent's highest educational qualification (at latest census)[1] | No school qualification | 50.7 | 45.0 |
| | Qualification in 5th form/Year 11 | 18.9 | 12.0 |
| | Qualification in the 6th Form/Year 12 | 4.8 | 3.9 |
| | Qualification in the 7th Form/Year 13 | 1.5 | 2.5 |
| | Overseas school qualification | 2.0 | 0.9 |
| | Other school qualification | 0.2 | S |
| | Trade certificate/vocational qualification | 17.2 | 31.3 |
| | Undergraduate degree | 3.2 | 2.6 |
| | Postgraduate degree/higher degree | 1.4 | 1.8 |
| Total number of observations | | 9,312 | 1,944 |

**Symbols**:
S suppressed

**Notes**:
[1] Responses are those from each daughter's/parent's latest available census where the response is non-missing.
[2] Region of usual residence based on Regional Council 2006 boundaries.
[3] Middle Eastern/Latin American/African.

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

### *3.3.3 Descriptive statistics of proxy for lifetime average income*

Table 10 presents descriptive statistics of the proxy incomes of offspring and parents in each offspring-parent pair. Sons have the highest mean proxy incomes, followed by fathers, daughters, and then mothers, although fathers have higher median incomes than sons. Uniquely, there are no fathers with proxy incomes of zero. The dispersion of proxy incomes is higher among offspring than parents, due to greater income inequality in the later census years (2006 and 2013) compared to the earlier ones (1981 and 1986, as noted in section 1.3 of chapter 1). The proportions of offspring and parents with proxy incomes of $1 (reflecting a single-year proxy of zero or a multiyear average of zero) are small, the highest being just under five percent among daughters in the daughter-father sample, and the lowest being zero for fathers in both samples. On average, sons' proxy incomes use about two observations and are centred slightly under age 35, while father's' proxy incomes use about 2.5 observations are centred slightly over age 35. Daughters' proxy incomes use 1.3 observations among daughter-father pairs and 1.7 among daughter-mother pairs, and are centred just under age 40, while mothers' incomes use just over three observations and are centred on age 40, on average. Appendix Table 3 and Appendix Table 4 provide further descriptive statistics on sons' and daughters' proxy incomes (respectively), including a breakdown by age, census linkage, and number of income observations used in the proxy, as well as descriptive statistics of the natural logarithm of proxy income.

**Table 10.** *Descriptive statistics of proxy for lifetime average income for offspring and parents in samples.*

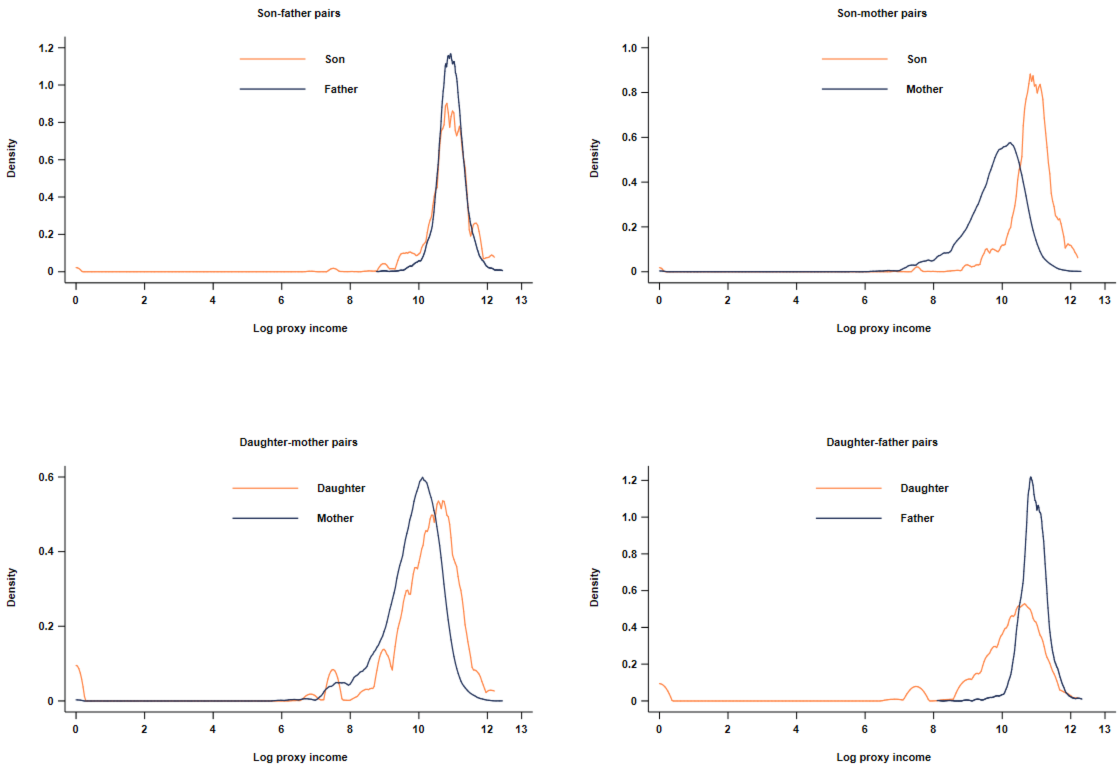
| Offspring-parent pairs | | n | Proxy for lifetime average income (in 2012 Q3 dollars) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Minimum | Median | Maximum | Mean | SD | No. (%) with proxy incomes of $1[1] | Mean no. observations used in proxy | Mean age (years) at which proxy is centred |
| Son-father | Sons | 4,617 | 1 | 54,429 | 200,000 | 60,779 | 34,634 | 24 (0.52%) | 1.8 | 34.7 |
| | Fathers | | 6,419 | 56,351 | 253,969 | 60,293 | 24,686 | 0 (0.00%) | 2.7 | 35.1 |
| Son-mother | Sons | 14,526 | 1 | 54,429 | 200,000 | 61,195 | 34,385 | 51 (0.35%) | 2.3 | 34.8 |
| | Mothers | | 1 | 20,492 | 219,279 | 23,940 | 17,501 | 15 (0.10%) | 3.4 | 40.0 |
| Daughter-mother | Daughters | 9,312 | 1 | 32,349 | 200,000 | 37,690 | 30,787 | 312 (3.35%) | 1.7 | 39.7 |
| | Mothers | | 1 | 20,656 | 252,140 | 23,871 | 16,921 | 9 (0.10%) | 3.2 | 40.1 |
| Daughter-father | Daughters | 1,944 | 1 | 32,349 | 200,000 | 37,224 | 29,922 | 90 (4.63%) | 1.3 | 39.7 |
| | Fathers | | 3,309 | 56,188 | 225,233 | 59,483 | 23,663 | 0 (0.00%) | 2.5 | 35.2 |

**Notes**:
[1] These are individuals with proxy incomes of zero (either a single-year proxy of zero or a multiyear average of zero), which have then been recoded to $1 to retain them in the log-log model specification used to estimate the IGE.

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

Figure 15 displays kernel density plots showing the probability density function of offspring's and parents' log proxy incomes. Unsurprisingly, the offspring and parent curves for same-sex pairs (son-father, daughter-mother) resemble one another more closely than do the curves for opposite-sex pairs (son-mother, daughter-father). Examining the curves for offspring in all pairs, there is some density at a log income of zero (real proxy income of $1), then the curves flatline, representing zero density at very low incomes. Mothers also display this long tail on the left (but with negligible density at zero log income), whereas fathers have no density at zero and no tail on the left. The two curves for sons abruptly stop at about 12.2 (real income of $200,000 in 2012 Q3 dollars) which reflects top-coding (that is, sons whose proxy income was the highest income band in 2013). In the middle of the distribution of log income, sons and fathers are similarly distributed with tall, spiked curves, whereas daughters and mothers have more dispersed (shorter, wider) curves. The curves for sons and daughters are 'noisier' than for parents who have smoother curves, due in part to allowing offspring to have single-year proxies which results in groups of offspring bunched at common proxies, especially at a log income of about 11.7 for sons ($120,000 in 2012 Q3 dollars, which is the representative value of the $100,001 to $150,000 band in the 2013 census) and at about 7.5 for daughters ($1,787 in 2012 Q3 dollars, which is the representative value of the $1 to $5,000 band in the 2013 census). Appendix Figure 1 presents, for each offspring-parent pair, scatterplots of offspring's versus parents' log proxy income with prediction lines fitted with linear regression.

**Figure 15**. *Kernel density plots of log proxy income for offspring and parents in each offspring-parent pair.*




**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

## 3.4 Results

### *3.4.1 Estimates of the intergenerational income elasticity*

To estimate the IGE, ordinary least squares is applied to equation (6), a variation of equation (1):

$$Y_{ij}^{\text{offspring}} = \alpha + \beta Y_i^{\text{parent}} + Age^{\text{offspring}} + Age^{2\ \text{offspring}} + Age^{\text{parent}} + Age^{2\ \text{parent}} + \varepsilon_i, \quad (6)$$

where $Y_{ij}^{\text{offspring}}$ is log lifetime average income of offspring $i$ in family $j$, $Y_i^{\text{parent}}$ is log lifetime average income of a parent in family $j$, $Age$ is the age of the offspring and parent in 1981, $\varepsilon_i$ is an error term capturing factors orthogonal to parental income, and $\beta$ is the intergenerational income elasticity. Standard errors are adjusted for clustering within families (to account for the presence of siblings in each sample). The regression results are presented in Table 11.

**Table 11**. *Intergenerational income elasticity estimates for offspring-parent pairs in the NZLC.*

| | Sample | | | |
|---|---|---|---|---|
| | **Son-father** | **Son-mother** | **Daughter-mother** | **Daughter-father** |
| Parent's log proxy income (IGE) | 0.239 *** | 0.054 *** | 0.145 *** | 0.135 |
| | (0.038) | (0.007) | (0.026) | (0.131) |
| Offspring's age | 0.053 ** | 0.073 *** | 0.340 *** | 0.365 |
| | (0.019) | (0.012) | (0.097) | (0.261) |
| Offspring's age-squared | -0.003 * | -0.004 *** | -0.012 ** | -0.013 |
| | (0.001) | (0.001) | (0.005) | (0.013) |
| Parent's age | 0.221 | 0.051 | 0.048 | -1.172 |
| | (0.116) | (0.033) | (0.143) | (0.832) |
| Parent's age-squared | -0.004 | -0.001 | -0.001 | 0.020 |
| | (0.002) | (0.001) | (0.002) | (0.014) |
| Constant | 4.826 ** | 9.114 *** | 5.558 * | 23.159 |
| | (1.727) | (0.537) | (2.321) | (12.673) |
| R-squared | 0.012 | 0.010 | 0.015 | 0.012 |
| Observations | 4,617 | 14,526 | 9,312 | 1,944 |

**Symbols**:
***p<0.001 **p<0.01 *p<0.05

**Notes**:
Robust standard errors in parentheses, adjusted for 4,416 clusters in father's unique ID number for son-father pairs, 13,638 clusters in mother's unique ID number for son-mother pairs, 8,892 clusters in mother's unique ID number for daughter-mother pairs, and 1,887 clusters in father's unique ID number in daughter-father pairs.
**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

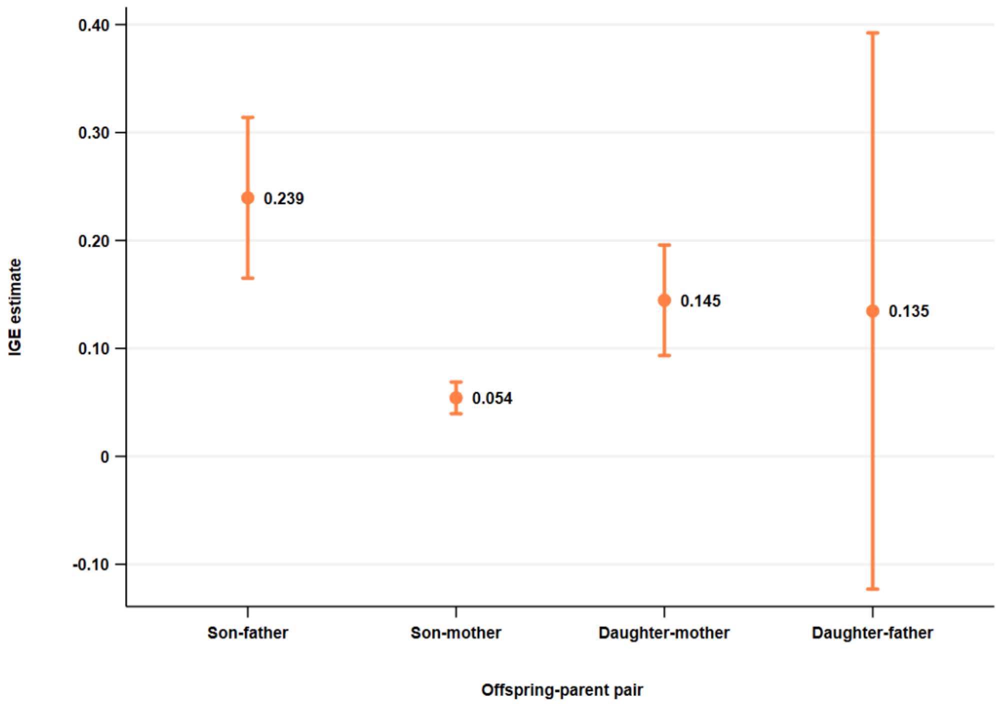
The son-father IGE is 0.239, the son-mother IGE is 0.054, the daughter-mother IGE is 0.145, and the daughter-father IGE is 0.135. This means that a 10 percent increase in fathers' long-run income is associated with, on average, a 2.39 percent increase in sons' long-run income and a 1.35 percent increase in daughters' long-run income, while a 10 percent increase in mother's long-run income is associated with a 0.54 percent increase in son's long-run income and a 1.45 percent increase in daughters' long-run income.[77] Put another way, if one father has twice the income of another father who is the same age, then the son of the higher-income father will, as an adult, make roughly 24 percent more income than the (same-aged) son of the lower-income father. Similarly, the daughter of the higher-income father will make roughly 14 percent more than the daughter of the lower-income father. Likewise, if one mother has twice the income of another mother, then the son (daughter) of the higher-income mother will, as an adult, make roughly five percent (15 percent) more income than the son (daughter) of the lower-income mother.[78]

All IGE estimates are statistically significant except for the daughter-father elasticity ($p$=0.306), meaning the null hypothesis that the association between daughters' and fathers' incomes is zero cannot be rejected. Gibbons (2010) also finds the association between daughters' and fathers' incomes to be not statistically significant (for daughters living in New Zealand at age 32). Offspring's age and its square term are also statistically significant in all pairs except daughter-father pairs, indicating that the age of sons and daughters continues to exert an effect on their log proxy incomes despite them being centred on the proxy age (plus or minus one year). On the

[77] This is the conventional interpretation. As discussed in chapter 2, Mitnik and Grusky (2020) state that the correct interpretation of these results is as follows: a 10 percent increase in fathers' long-run income is associated with a 2.39 percent increase in *the geometric mean* of sons' long-run income.
[78] These interpretations assume the relationship between the log incomes of offspring and parents is linear.

other hand, parents' age and its square term had no statistically significant effect on offspring's proxy income. This suggests that the method of taking a time-average of income centred on a particular age was largely successful in stripping out age effects for parents, but was not successful for offspring, likely due to allowing offspring to have single-year proxies. Figure 16 presents the IGE estimates pictorially, along with their 95 percent confidence intervals. The son-mother confidence interval is relatively narrow, owing to the larger sample size, whereas the interval for daughter-father pairs is very wide, spanning -0.12 to 0.39.

**Figure 16**. *Intergenerational income elasticity point estimates and 95% confidence intervals for offspring-parent pairs.*




**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

The daughter-father IGE estimate is just over half the size of the son-father IGE estimate, and the son-mother estimate just over one-third the size of the daughter-mother estimate. In other words, increases in fathers' income have a much larger effect on sons' than daughters' incomes (compare the left-most and right-most estimates), while increases in mothers' incomes have a much larger effect on daughters' than sons' incomes (compare the two middle estimates). The same pattern among the same four offspring-parent pairs is found in Deutscher and Mazumder's (2020) analysis of Australian tax data, albeit with smaller differences between sons and daughters. At the same time, fathers' and mothers' incomes have similar effects on daughters' incomes (compare the two estimates for daughters on the right), but dissimilar effects on sons' incomes (a much larger effect from fathers, compare the two estimates for sons on the left). Again, Deutscher and Mazumder (2020) find a similar pattern; their elasticities for daughter-mother (0.105) and daughter-father (0.133) pairs differ by 27 percent, whereas their son-father IGE estimate (0.157) is more than double the size of their son-mother estimate (0.065).

The pattern is also consistent with results from Raaum et al. (2007, p. 24) who estimate the intergenerational correlation in completed years of schooling between son-father, son-mother, daughter-mother, and daughter-father pairs in Denmark, Finland, Norway, Great Britain, and the US, finding that "men's years of schooling seem to be more strongly correlated with the attainment of their father than their mother in all countries, while for women the correlation is more or less the same with respect to both parents".

The explanation for the pattern in Figure 16 is not clear. Using US data from some of the National Longitudinal Surveys, Altonji and Dunn (2000, p. 245) find that the intergenerational correlation in annual hours worked is much higher for same-sex pairs (0.23 between fathers and sons, 0.24 between mothers and daughters) than for opposite-sex pairs (0.07 between mothers and sons, 0.00 between fathers and daughters), and that these are primarily due to within-family correlations in preferences, concluding that "there are strong family linkages in hours preferences between family members of the same sex". Intergenerational correlations in work hours that run along gender lines (which may reflect an unequal division of labour within households or occupational sex segregation) may explain the dissimilar son-father and son-mother IGE estimates found in the current analysis but cannot explain the closeness of the daughter-mother and daughter-father estimates. On the other hand, assortative mating (marital sorting) might explain why mothers and fathers exert similar influences on daughters' incomes but cannot explain the disparate father and mother effects found among sons.

*3.4.2 Robustness checks of intergenerational income elasticity estimates*

Figures 17, 18, 19, and 20 display, for each offspring-parent pair, the point estimates and 95 percent confidence intervals of the models ("Model 1", the left-most estimates) followed by six alternative specifications:

Model 2. This model drops all age variables (including the square terms) from the model, since some analysts consider age to be endogenous to income (see footnote 32 in section 2.3.1).

Model 3. This model recodes all zero incomes to \$0.01 instead of \$1, and then re-derives the proxy incomes, given that the way zero incomes are treated can change the IGE dramatically (see section 2.3.2).

Model 4. This model drops any offspring and parents with proxy incomes of \$1. In practice, these arise either from offspring having a single-year proxy of zero income (then recoded to \$1) or offspring or parents having a multi-year average of zero income (each year of zero income recoded to \$1, then time-averaged to \$1). Thus, it is equivalent to retaining offspring and parents with positive proxy incomes only (before recoding). It is difficult to gauge a person's average income over their lifetime if the particular year(s) they were observed happened to be periods when they received no income. Assigning these individuals a lifetime average income of \$1, while having the practical benefit of retaining them in the analysis, is not very realistic since it implies these individuals have no earnings capacity whatsoever over their lifetime.

Model 5. This model imposes the requirement that offspring's proxy incomes must – like their parents' – be a multi-year average. This has the effect of reducing the sample size, because all offspring with single-year proxy incomes are dropped from the models.

Model 6. This model uses Mitnik's alternative estimator – the IGE of *expected* income - using Poisson pseudo-maximum likelihood (PPML) estimation.[79]

Model 7. This model relaxes the restrictions on parental age imposed by the requirement to observe parents around particular ages (35 for fathers, 40 for mothers) designed to minimise lifecycle bias. Figure 11 shows that this minimisation-of-lifecycle-bias requirement has the effect of restricting the age range of fathers to those who were aged 19 to 33 in 1981 for those paired with sons (meaning fathers who had their sons when aged 17 to 31) and to those aged 24 to 33 in 1981 for those paired with daughters (meaning fathers who had their daughters when aged 22 to 31). Similarly, Figure 12 shows that this requirement restricts the age of mothers to those who were 18 to 38 in 1981 for those paired with sons (meaning mothers who gave birth to their sons between 16 and 36) and to those who were 23 to 38 in 1981 for those paired with daughters (meaning mothers who gave birth to their daughters between 21 and 36). Thus, fathers and (to a lesser extent) mothers who had children at relatively older ages – beyond age 31 for fathers and beyond 36 for mothers – are excluded from the samples (because their incomes cannot be observed around ages 35 and 40, respectively). This exclusion may induce a downward bias in the estimate of the IGE arising from a relatively homogenous or unrepresentative sample (see section 2.3.4). Therefore, in Model 7, instead of holding fixed the age around which parental income is measured, the *years* of parental income are held fixed (an average of parental income is taken over the 1981 and 1986 censuses) and no restrictions on parental age are imposed (except those trimming the top and bottom one percent of the age-at-birth-of-offspring distribution). Sample sizes consequently increase considerably. No changes are made to how offspring income is measured.

[79] Arguably, Mitnik's PPML specification is not an 'alternative specification' of the preferred model, but a *completely different model* (a different estimator with a different interpretation than the estimator from the conventional log-log specification). Note also that the PPML estimates are identical whether offspring with zero incomes have their incomes imputed at $1 or kept at zero.

For son-father pairs only, the following additional robustness check is performed:

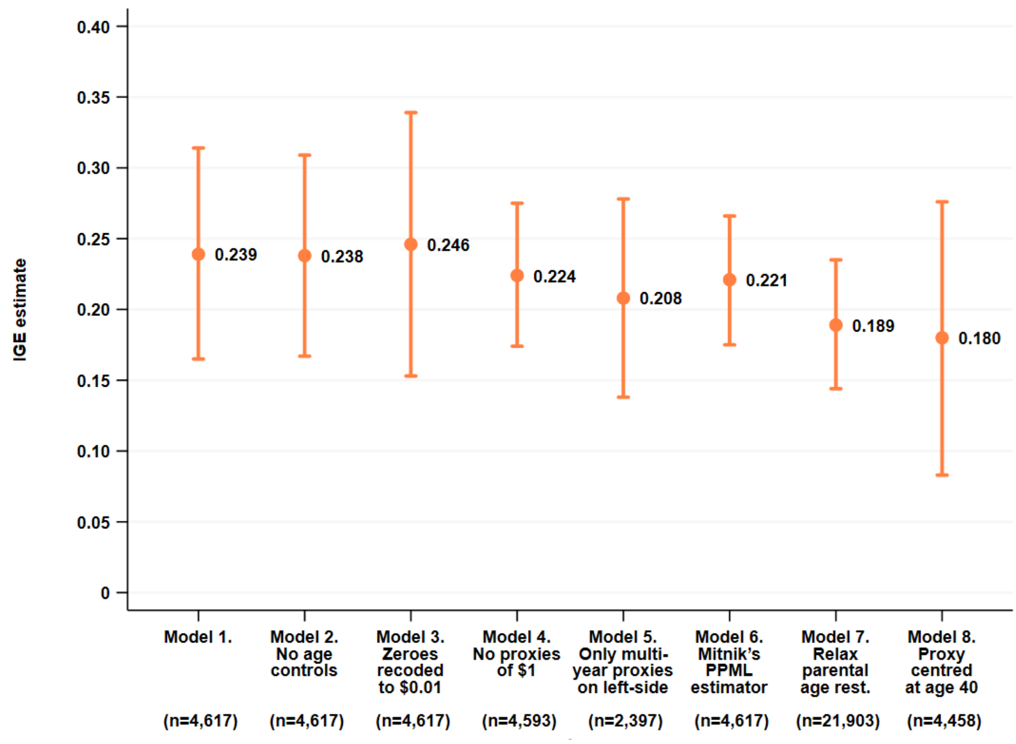
Model 8. All sons' and fathers' proxy incomes are centred around age 40 rather than age 35, since this is a common age used in empirical studies (e.g., Aaronson & Mazumder, 2008; Mazumder, 2016) and falls within empirically-estimated age windows for minimising life-cycle bias (see Table 1 in section 2.3.5). There are constraints to attempting a similar robustness check for women – say, centring all daughters' and mothers' proxies at age 45 rather than age 40 – due to there being only a relatively small number of daughters in 1981 who are observed in the NZLC at age 45 (Appendix Table 2 shows that this would only be possible for 12 to 14-year-old daughters in 1981) and none are observed in their fifties.

Across all offspring-parent pairs, removing the age controls has little effect on the IGE (the point estimates decrease by 0.001 for all pairs except for daughter-mother pairs where it increases by 0.002, and the confidence intervals remain about the same for all pairs). This may be due to centring the proxy incomes at particular ages for both offspring and parents, which minimises the effects of age on the relationship between incomes across generations. Recoding zero incomes to $0.01 (instead of $1) also has marginal impacts on the IGE point estimates (changing them in the range of -0.007 to +0.007), but the confidence intervals tend to widen.

Dropping offspring and parents with proxy incomes of $1 has small-to-moderate effects on the IGE point estimates but tightens their confidence intervals. The son-father IGE decreases by six percent to 0.224, the daughter-mother IGE decreases by 21 percent to 0.115, the son-mother IGE increases by 13 percent to 0.061, and the daughter-father IGE increases by 10 percent to 0.149. Estimates of the IGE are known to be sensitive to the treatment of zero incomes (Chetty et al., 2014; Mitnik & Grusky, 2020). As indicated in Table 10, these effects are driven mainly by offspring with zero income at their proxy age, rather than parents. Of the daughters falling into this category (312 in daughter-mother pairs and 90 in daughter-father pairs), the vast majority were not in the labour force at their proxy age, were not studying, and reported that they had 'looked after a child who is a member of my household' as an unpaid activity in the past four weeks, implying that these daughters are stay-at-home mothers. Of the sons falling into this category (24 in son-father pairs and 51 in son-mother pairs), the majority were not in the labour force and not studying at their proxy age, and were not receiving a social welfare benefit (reporting 'No source of income' in the previous 12 months). A small number appeared to be stay-at-home fathers, but for most it was not clear why they reported zero income.

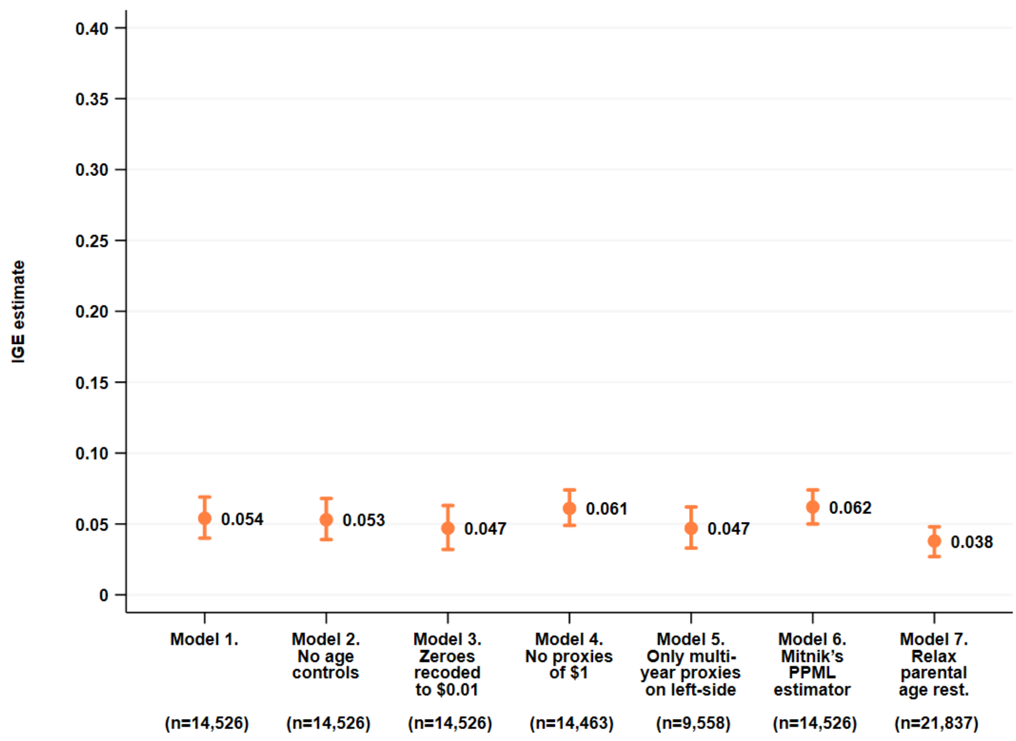
With the exception of daughter-father pairs, restricting the sample to offspring with multi-year proxies lowers the IGE point estimate by between 10 and 13 percent, but the confidence intervals overlap substantially with those of the first models. For daughter-father pairs, the point estimate changes dramatically, increasing by 158 percent to 0.348. This is because the daughters dropped from the sample all had single-year proxy incomes of $1 (zero incomes), the vast majority of whom were not in the labour force. Dropping these daughters tightened the link between the incomes of the remaining daughters and their fathers. Using Mitnik's PPML estimator had moderate to large effects on the IGE point estimates, decreasing them by eight percent for son-

**Figure 17**. *Robustness checks of IGE estimate for son-father pairs.*



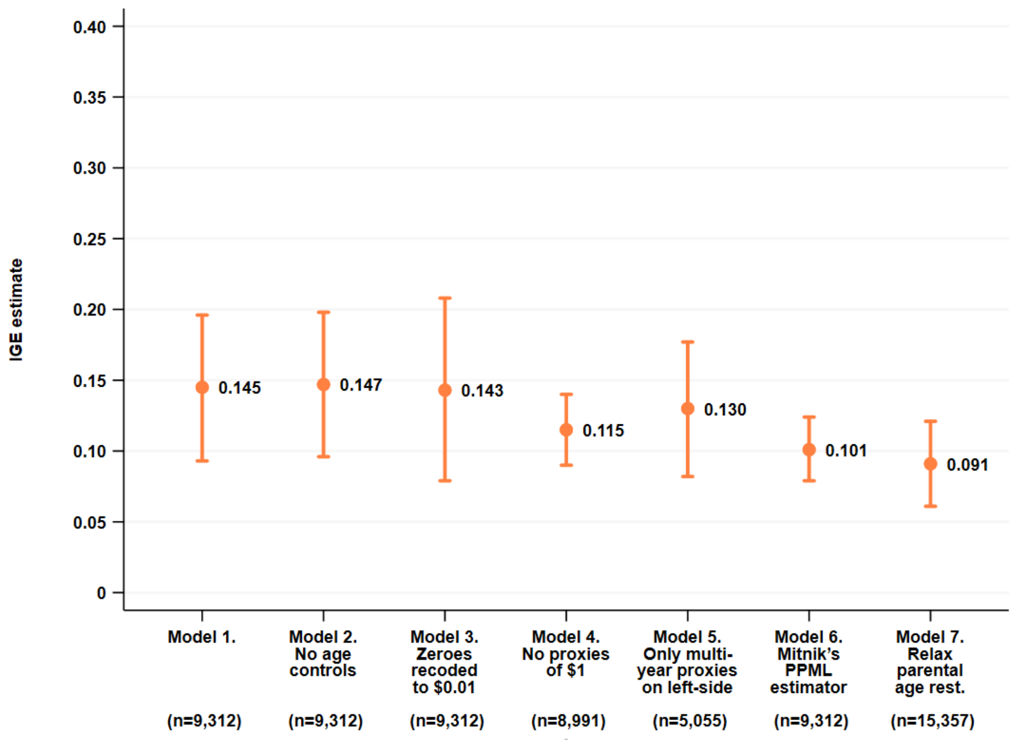

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

**Figure 18**. *Robustness checks of IGE estimate for son-mother pairs.*



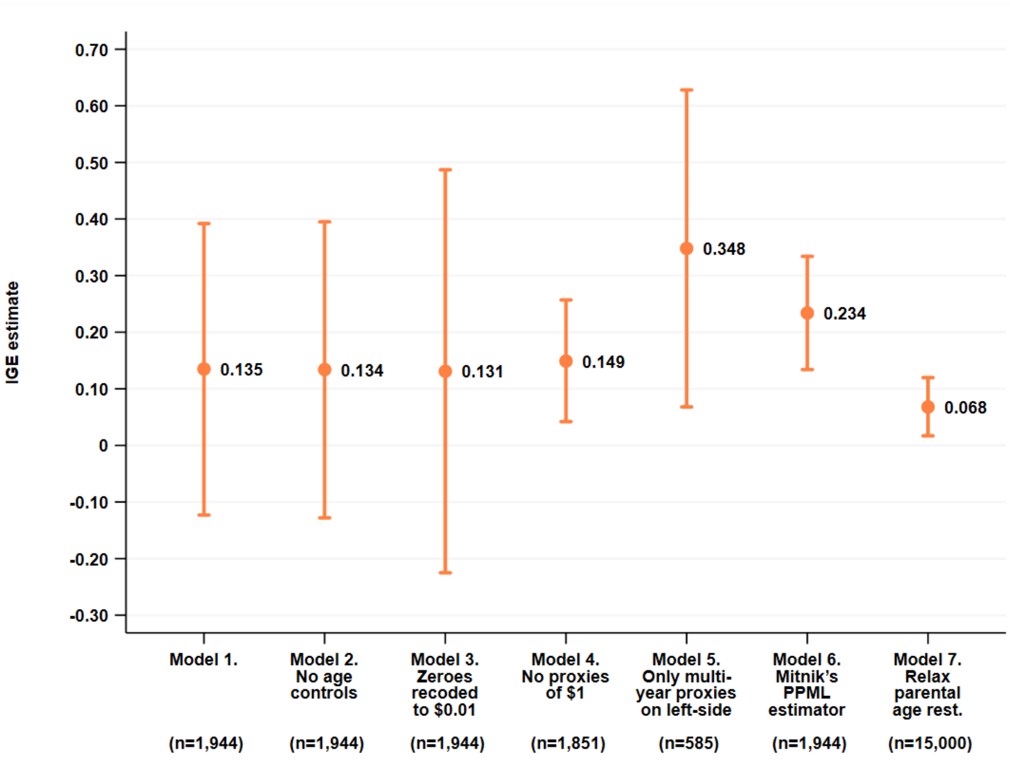

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

**Figure 19**. *Robustness checks of IGE estimate for daughter-mother pairs.*




**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

**Figure 20**. *Robustness checks of IGE estimate for daughter-father pairs.*




**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

father pairs and 30 percent for daughter-mother pairs, and increasing them by 15 percent for son-mother pairs and 73 percent for daughter-father pairs, with narrower confidence intervals for all pairs (but overlapping with the intervals of the preferred models). Relaxing the restrictions on parental age results in the mean age of fathers (in 1981) increasing from 30.5 to 39.0 among son-father pairs and from 31.3 to 40.4 among daughter-father pairs and the mean age of mothers increasing from 33.8 to 37.3 among daughter-mother pairs and from 32.7 to 36.1 among son-mother pairs. Relaxing the restrictions also has large effects on the IGE estimates, decreasing them by between 21 percent (among son-father pairs) and 50 percent (among daughter-father pairs). Thus, any attenuation bias avoided by having a sample of parents from a wider age range (hence presumably a more representative sample) is offset by larger biases in the opposite direction from averaging incomes over only two observations and from measuring incomes at older stages of the lifecycle, on average. Finally, centring the proxy incomes at age 40 for sons and fathers changed the size and composition of the son-father sample and lowered the IGE point estimate by 25 percent, but with a wide confidence interval overlapping with the interval of model 1.

Overall, the model 1 IGE estimates for son-father and son-mother pairs appear reasonably robust, with no dramatic differences across alternative specifications. For daughter-mother pairs, estimation of the IGE appears sensitive to the inclusion or exclusion of individuals who report zero incomes, as indicated by the point estimates from both the ‘no proxy incomes of $1’ model and Mitnik’s PPML model (which is more robust to the inclusion of individuals with zero income). For daughter-father pairs, two of the alternative point estimates differ substantially from the other four, the confidence intervals are very wide, and some straddle an IGE of zero while others do not. The model 1 IGE estimate for daughter-father pairs thus appears to be very sensitive to model specification.

### *3.4.3 Group-specific estimates of the intergenerational income elasticity*

The IGEs estimated for offspring-parent pairs may conceal heterogeneity between groups, so this section estimates intergenerational persistence for different groups. Hertz (2005) notes that stratifying and calculating within-group IGEs does not produce a meaningful answer to the question of which groups are most and least mobile. In a subsequent paper describing his method for calculating group-specific measures of intergenerational persistence, Hertz (2008, p. 415) explains why simply comparing within-group IGEs is erroneous:

> …any long-run difference in group [income] means...is ignored under stratification, and this can lead to the false impression that a persistently disadvantaged group is actually highly mobile. Children of the disadvantaged group may be quite mobile with respect to their parents when measured by their within-group intergenerational elasticity. But if both parents and children in that group are generally confined to the lower end of the income distribution, then the children will typically remain relatively disadvantaged, which is to say, they will display little mobility.

When mean income differs between groups, estimating the IGE separately for each group and then comparing the results will produce a misleading comparison of intergenerational persistence between the groups. To deal with this, Hertz (2008, p. 415) develops a group-specific IGE statistic which measures “both the degree to which parents and children have similar incomes within a group, as well as the degree to which people in a given group tend to fall above or below the sample mean [the average income pooled across all groups]".

Let groups be indexed by $g = 1 \dots G$ and let $\hat{\pi}_g$ represent group $g$'s share of the total number of offspring-parent pairs in the sample. Let $x$ and $y$ be the lifetime average incomes of the parent and offspring generations, respectively, with sample means $\overline{x}$ and $\overline{y}$ and estimated variances $\hat{\sigma}_x^2$ and $\hat{\sigma}_y^2$. Let $\overline{x}_g$ and $\overline{y}_g$ represent the sample means for group $g$'s parents and offspring, and $\hat{\sigma}_{x(g)}^2$ and $\hat{\sigma}_{y(g)}^2$ their respective variances. Let $\hat{\beta}_g$ represent the IGE from a regression of $y$ on $x$ in group $g$. Let $\hat{\delta}_G$ represent the group-size-weighted between-group regression coefficient, defined as the coefficient from the regression of the group means $\overline{y}$ on the group means $\overline{x}$:[80]

$$\hat{\delta}_G = \sum_g \hat{\pi}_g \left(\overline{y}_g - \overline{y}\right)\left(\overline{x}_g - \overline{x}\right) / \sum_g \hat{\pi}_g \left(\bar{x}_g - \bar{x}\right)^2 \tag{7}$$

Group-specific IGEs are calculated by decomposing the pooled IGE into a *within-group component* (the degree to which parents and children have similar incomes within a group) and a *between-group component* (the degree to which people in a given group tend to fall above or below the average income across all groups) and then allocating each group a share of each component, as shown in equation (8):

$$\hat{\beta} = \sum_g \hat{\pi}_g \left( \hat{\beta}_g \frac{\hat{\sigma}_{x(g)}^2}{\hat{\sigma}_x^2} + \frac{\left(\overline{x}_g - \overline{x}\right)\left(\overline{y}_g - \overline{y}\right)}{\hat{\sigma}_x^2} \right) \tag{8}$$

For each group, the within-group component is the group's within-group IGE scaled by the group's contribution to the variance in the parent generation (the term on the left within the brackets in equation (8)), while the between-group component is the group's contribution to the between-group covariance (the term on the right within the brackets in equation (8)). The result is a metric "by which income persistence may meaningfully be compared across groups, even when those groups have very different average outcomes" (Hertz, 2008, p. 415).[81]

Table 12 (sons) and Table 13 (daughters) present IGEs for different socio-demographic groups of offspring, along with the pooled IGE for comparison (all IGEs shaded in orange). The analysis is performed for:

1. *Offspring's ethnic group*: Due to very small numbers for some level one ethnic groups, this variable is collapsed into three groupings – European only, Māori only, and Other ethnic groups. The latter includes all other single ethnic groups, such as Pacific only and

[80] In their textbook on multilevel modelling, Snijders and Bosker (2012, pp. 26-28) also discuss within- and between-group relations within a regression context including an explanation of the within-group and between-group coefficients.

[81] However, as noted in section 2.3.5 on country- and sex-specific proxy ages, Nybom and Stuhler (2016, p. 266) issue a note of caution: "The comparison of elasticities across subgroups of a population can be compromised when the age pattern in income profiles differs, which may, for example, be the case when groups are classified by education, sex, or immigration status".

Asian only, as well as combinations of ethnic groups, such as European and Māori. The ethnic group variable is based on responses from the latest available census (where the response is non-missing) for each offspring. For brevity, offspring who are European only, Māori only, or from Other ethnic groups are referred to as ‘European’, ‘Māori’ and ‘Other ethnic’, respectively.

This ‘single and combination’ ethnic classification results in proportionately small numbers falling into the ‘Māori only’ category (since many Māori belong to other ethnic groups, thus placing them in the ‘Two or more ethnic groups’ category). Therefore, the following robustness check is performed: offspring are dichotomised into ‘Any Māori’ ethnic identification (all offspring who mark the ‘Māori’ ethnic group box at their latest available census, regardless of whether they mark any other ethnic groups) versus ‘No Māori’ (all other offspring, which equates to offspring who indicate no Māori ethnic identification). The numbers of offspring falling into the ‘Any Māori’ category is roughly, and on average, 2.5 times the numbers falling into the ‘Māori only’ category. The results of this robustness check are presented in Appendix Table 5.

2. *Rurality in 1981*: A rurality indicator variable denoting whether the offspring lived in an urban or rural area in 1981, derived from Atkinson, Shaw, Blakely, Stanley and Sloane’s (2010, pp. 146-147) aggregation of the ‘Urban Area’ variable in the 1981 census.

3. *Parent’s receipt of a social security benefit in 1981*: This variable denotes whether the parent received at least one social security benefit during the year ended 31 March 1981, *excluding* the universal payments of ‘Family Benefit’ and ‘National Superannuation’.

4. *Parent’s highest educational qualification*: This variable is collapsed into four categories: no school qualification, school qualification (a qualification gained in years 11, 12, or 13, or an overseas equivalent or other school qualification), trade qualification (trade certificate or vocational qualification), or degree (undergraduate or postgraduate or higher degree). Parent’s education is based on responses from each parent’s latest available census where the response is non-missing.

Following Hertz (2008), each group’s sample share, mean and variance of (log proxy) income, and within-group IGE are used to derive within-group and between-group components, which sum to a group-specific IGE. To derive standard errors for these group IGEs, the estimates are bootstrapped, resampling observations from the data (with replacement) with 100 replications, which are then used to derive confidence intervals.

Beginning with sons, Table 12 shows that European sons had higher log proxy incomes, on average, than Māori and Other ethnic sons, and their fathers had higher average incomes than the fathers of non-European sons (but the same pattern was not observed among their mothers). However, the group-specific IGEs suggest income persistence across generations does not differ drastically between these ethnic groups. The IGEs are a little lower for Māori sons (IGE=0.196

among son-father pairs, 0.020 among son-mother pairs) and Other ethnic sons (IGE=0.210 among son-father pairs, 0.032 among son-mother pairs) compared to European sons (IGE=0.248 among son-father pairs, 0.058 among son-mother pairs). The confidence intervals are very wide, especially for the Māori and Other ethnic groups who have relatively small samples sizes.

The robustness check comparing 'Any Māori' identification versus 'No Māori' identification (Appendix Table 5) reverses the 'Māori only' result for sons in son-father pairs: Māori sons now have *higher* income persistence ('Any Māori' IGE=0.314) than sons who have no Māori identification ('No Māori' IGE=0.226), although the confidence intervals overlap. However, persistence among Māori sons in son-mother pairs does not materially change (IGE=0.023).

Sons from urban and rural backgrounds (in 1981) in both son-parent pairs had broadly similar average incomes, but income persistence from fathers to sons was about 80 percent higher among rural sons (IGE=0.396) compared to urban sons (IGE=0.219), while income persistence from mothers to sons was about 60 percent higher among rural sons (IGE=0.082) compared to urban sons (IGE=0.050). While the wide confidence intervals indicate these estimates are not significantly different from each other, the point estimates do suggest that, broadly speaking, intergenerational persistence is higher among rural families than urban families.

Only a small minority of sons had a father (two percent) or mother (five percent) in receipt of a non-universal social security benefit in the year prior to the 1981 census. These sons had lower average incomes than sons whose parents did not receive any such benefits. Income persistence from fathers to sons was three times higher among fathers receiving a benefit (IGE=0.715) compared to fathers not receiving a benefit (IGE=0.227). The opposite pattern is observed among son-mother pairs: income persistence was weaker (indeed negative) for mothers who received a benefit (IGE=-0.010, meaning higher incomes among mothers are associated with lower incomes among sons) compared to mothers who did not receive a benefit (IGE=0.057).

On average, fathers and mothers with no qualifications had the lowest incomes and fathers and mothers with degrees had the highest incomes. Sons' average incomes increased monotonically with father's education level (a similar pattern was observed in relation to mother's education level, except that sons with mothers from the two middle education categories had similar average incomes). Sons whose fathers had no qualifications (IGE=0.246) or only school qualifications (IGE=0.251) showed similar levels of income persistence, which in turn were similar to the overall (pooled) level of persistence. Persistence among sons whose fathers had trade qualifications (IGE=0.117) was about half of the overall level of persistence, indicating higher mobility among these sons. Persistence among sons whose fathers had degrees (IGE=0.668) was nearly three times as high as overall persistence, indicating intergenerational transmission of advantage to these sons (given that these sons and fathers had the highest average incomes). A similar pattern was observed in relation to mother's education: persistence among sons whose mothers had no qualifications or only school qualifications was similar to overall persistence, persistence among

**Table 12**. *Group-specific IGE estimates for son-father and son-mother pairs.*

| | Son-father pairs | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Son's ethnic group**[1] | | | **Rurality in 1981** | | **Parent's benefit receipt 1981** | | **Father's education**[1] | | | |
| | **Pooled** | **European only** | **Māori only** | **Other ethnic**[2] | **Urban** | **Rural** | **No benefit** | **Benefit** | **No school qual.** | **School qual.** | **Trade qual.** | **Degree** |
| Sample size | 4,617 | 3,639 | 222 | 756 | 4,059 | 561 | 4,506 | 111 | 1,722 | 990 | 1,554 | 351 |
| Sample shares (%) | 100.0 | 78.8 | 4.8 | 16.4 | 87.9 | 12.1 | 97.6 | 2.4 | 37.3 | 21.4 | 33.6 | 7.6 |
| Mean log proxy income: sons | 10.80 | 10.84 | 10.42 | 10.70 | 10.80 | 10.77 | 10.80 | 10.64 | 10.72 | 10.81 | 10.84 | 10.98 |
| Mean log proxy income: fathers | 10.93 | 10.95 | 10.86 | 10.88 | 10.95 | 10.82 | 10.94 | 10.55 | 10.82 | 11.00 | 10.95 | 11.24 |
| Within-group component of pooled IGE | **0.239** | 0.244 | 0.001 | 0.175 | 0.219 | 0.376 | 0.227 | 0.305 | 0.183 | 0.248 | 0.113 | 0.300 |
| Between-group component of pooled IGE | | 0.005 | 0.195 | 0.035 | 0.000 | 0.020 | 0.000 | 0.410 | 0.063 | 0.003 | 0.004 | 0.368 |
| **Group-specific IGE** | | **0.248** | **0.196** | **0.210** | **0.219** | **0.396** | **0.227** | **0.715** | **0.246** | **0.251** | **0.117** | **0.668** |
| Bootstrapped 95% confidence interval | | (0.176, 0.320) | (-0.135, 0.527) | (-0.040, 0.461) | (0.145, 0.294) | (0.140, 0.653) | (0.161, 0.293) | (0.291, 1.139) | (0.117, 0.375) | (0.132, 0.370) | (0.017, 0.217) | (0.395, 0.940) |
| | Son-mother pairs | | | | | | | | | | | |
| | | **Son's ethnic group**[1] | | | **Rurality in 1981** | | **Parent's benefit receipt 1981** | | **Mother's education**[1] | | | |
| | **Pooled** | **European only** | **Māori only** | **Other ethnic**[2] | **Urban** | **Rural** | **No benefit** | **Benefit** | **No school qual.** | **School qual.** | **Trade qual.** | **Degree** |
| Sample size | 14,526 | 11,592 | 579 | 2,355 | 12,786 | 1,743 | 13,830 | 681 | 6,504 | 4,362 | 2,736 | 927 |
| Sample shares (%) | 100.0 | 79.8 | 4.0 | 16.2 | 88.0 | 12.0 | 95.3 | 4.7 | 44.8 | 30.0 | 18.8 | 6.4 |
| Mean log proxy income: sons | 10.83 | 10.86 | 10.50 | 10.72 | 10.83 | 10.80 | 10.84 | 10.65 | 10.76 | 10.87 | 10.86 | 10.94 |
| Mean log proxy income: mothers | 9.80 | 9.79 | 9.85 | 9.83 | 9.78 | 9.90 | 9.79 | 10.00 | 9.66 | 9.82 | 9.94 | 10.20 |
| Within-group component of pooled IGE | **0.054** | 0.059 | 0.040 | 0.037 | 0.050 | 0.086 | 0.057 | 0.035 | 0.053 | 0.048 | 0.011 | 0.023 |
| Between-group component of pooled IGE | | 0.000 | -0.020 | -0.004 | 0.000 | -0.004 | 0.000 | -0.045 | 0.011 | 0.002 | 0.007 | 0.058 |
| **Group-specific IGE** | | **0.058** | **0.020** | **0.032** | **0.050** | **0.082** | **0.057** | **-0.010** | **0.064** | **0.050** | **0.018** | **0.081** |
| Bootstrapped 95% confidence interval | | (0.042, 0.075) | (-0.076, 0.116) | (-0.007, 0.072) | (0.034, 0.067) | (0.042, 0.122) | (0.043, 0.071) | (-0.077, 0.057) | (0.041, 0.088) | (0.022, 0.078) | (-0.015, 0.051) | (0.010, 0.152) |

**Notes**:
[1] For these variables, responses are those from the latest available census where the response is non-missing for each son and parent.
[2] 'Other ethnic' includes all other single ethnic groups (at classification level one, e.g. Pacific only, Asian only, etc.) plus combinations of ethnic groups (e.g. European and Māori).

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

**Table 13**. *Group-specific IGE estimates for daughter-mother and daughter-father pairs.*

| | Daughter-mother pairs | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Daughter's ethnic group[1] | | | Rurality in 1981 | | Parent's benefit receipt 1981 | | Mother's education[1] | | | |
| | **Pooled** | **European only** | **Māori only** | **Other ethnic[2]** | **Urban** | **Rural** | **No benefit** | **Benefit** | **No school qual.** | **School qual.** | **Trade qual.** | **Degree** |
| Sample size | 9,312 | 7,806 | 339 | 1,167 | 8,367 | 942 | 8,757 | 546 | 4,722 | 2,556 | 1,608 | 423 |
| Sample shares (%) | 100.0 | 83.8 | 3.6 | 12.5 | 89.9 | 10.1 | 94.1 | 5.9 | 50.7 | 27.5 | 17.3 | 4.6 |
| Mean log proxy income: daughters | 9.90 | 9.88 | 10.13 | 9.97 | 9.89 | 10.03 | 9.89 | 10.01 | 9.86 | 9.93 | 9.91 | 10.07 |
| Mean log proxy income: mothers | 9.80 | 9.79 | 9.78 | 9.88 | 9.80 | 9.81 | 9.79 | 10.02 | 9.66 | 9.87 | 9.99 | 10.20 |
| Within-group component of pooled IGE | **0.145** | 0.144 | 0.123 | 0.130 | 0.139 | 0.189 | 0.149 | 0.033 | 0.143 | 0.126 | 0.145 | -0.028 |
| Between-group component of pooled IGE | | 0.000 | -0.006 | 0.007 | 0.000 | 0.002 | 0.000 | 0.033 | 0.007 | 0.003 | 0.003 | 0.090 |
| **Group-specific IGE** | | **0.144** | **0.116** | **0.137** | **0.139** | **0.191** | **0.150** | **0.066** | **0.150** | **0.129** | **0.148** | **0.062** |
| Bootstrapped 95% confidence interval | | (0.084, 0.204) | (-0.011, 0.244) | (0.008, 0.267) | (0.085, 0.193) | (0.024, 0.359) | (0.094, 0.205) | (-0.042, 0.174) | (0.081, 0.219) | (0.041, 0.218) | (0.052, 0.243) | (-0.169, 0.293) |
| | Daughter-father pairs | | | | | | | | | | | |
| | | Daughter's ethnic group[1] | | | Rurality in 1981 | | Parent's benefit receipt 1981 | | Father's education[1] | | | |
| | **Pooled** | **European only** | **Māori only** | **Other ethnic[2]** | **Urban** | **Rural** | **No benefit** | **Benefit** | **No school qual.** | **School qual.** | **Trade qual.** | **Degree** |
| Sample size | 1,944 | 1,569 | 102 | 273 | 1,722 | 222 | 1,899 | 39 | 876 | 372 | 612 | 87 |
| Sample shares (%) | 100.0 | 80.8 | 5.2 | 14.0 | 88.6 | 11.4 | 98.0 | 2.0 | 45.0 | 19.2 | 31.4 | 4.4 |
| Mean log proxy income: daughters | 9.77 | 9.76 | 9.61 | 9.91 | 9.77 | 9.79 | 9.77 | 9.62 | 9.74 | 9.83 | 9.76 | 9.91 |
| Mean log proxy income: fathers | 10.92 | 10.94 | 10.81 | 10.84 | 10.93 | 10.84 | 10.93 | 10.48 | 10.84 | 10.99 | 10.95 | 11.27 |
| Within-group component of pooled IGE | **0.135** | 0.246 | 0.159 | -0.339 | 0.056 | 0.763 | 0.152 | -0.910 | 0.061 | 0.519 | -0.070 | -0.250 |
| Between-group component of pooled IGE | | -0.002 | 0.116 | -0.079 | 0.000 | -0.010 | 0.000 | 0.435 | 0.018 | 0.026 | -0.002 | 0.339 |
| **Group-specific IGE** | | **0.244** | **0.275** | **-0.419** | **0.056** | **0.753** | **0.153** | **-0.475** | **0.079** | **0.545** | **-0.072** | **0.090** |
| Bootstrapped 95% confidence interval | | (-0.065, 0.553) | (-0.782, 1.332) | (-1.001, 0.164) | (-0.183, 0.294) | (-0.152, 1.658) | (-0.101, 0.407) | (-4.033, 3.082) | (-0.300, 0.457) | (-0.135, 1.225) | (-0.425, 0.281) | (-2.427, 2.606) |

**Notes**:
[1] For these variables, responses are those from the latest available census where the response is non-missing for each daughter and parent.
[2] 'Other ethnic' includes all other single ethnic groups (at classification level one, e.g. Pacific only, Asian only, etc.) plus combinations of ethnic groups (e.g. European and Māori).

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

sons with trade-qualified mothers was lower than overall persistence, and persistence among sons with degree-qualified mothers was highest (however, the substantial advantage of having a degree-qualified parent was weaker for mothers than for fathers). Note that the large IGE for sons with degree-qualified fathers is the only estimate with a confidence interval that does not overlap with those of the other groups, indicating a statistically significant result (the only such result across all offspring-parent pairs).

Turning to daughters, Table 13 shows that Māori and Other ethnic daughters had higher average incomes than European daughters (which may reflect ethnic differences in labour force participation, or alternatively, differential attrition in the NZLC by ethnicity). For daughter-mother pairs, the group-specific IGEs differed little by ethnic group, being somewhat lower among Māori and Other ethnic daughters. For daughter-father pairs, the pooled IGE concealed a negative IGE (-0.419) for Other ethnic daughters and higher IGEs for European (0.244) and Māori (0.275) daughters than for the total sample.

The robustness check comparing 'Any Māori' identification versus 'No Māori' identification (Appendix Table 5) resulted in slightly higher persistence among Māori daughters in daughter-mother pairs ('Any Māori' IGE=0.145), and a dramatic reduction of persistence (into negative persistence, which equates to intergenerational mobility) among Māori daughters in daughter-father pairs ('Any Māori' IGE= -0.089, compared to an IGE of 0.275 among 'Māori only' daughters).

Daughters from rural backgrounds had slightly higher average incomes than their urban counterparts. Persistence from fathers to daughters was very strong for daughters from rural backgrounds (IGE=0.753) and weak for daughters from urban backgrounds (IGE=0.056). Persistence from mothers to daughters was about the same for daughters from rural (IGE=0.139) and urban (IGE=0.191) backgrounds.

Small minorities of daughters had mothers (six percent) and fathers (two percent) who received a social security benefit. Unexpectedly, mothers who received a benefit had higher proxy incomes, on average, than mothers who did not receive a benefit, and the daughters of the former also had higher incomes than the daughters of the latter. This unexpected pattern did not apply to daughter-father pairs. For daughters in both pairs, persistence was weaker if the parent had received a benefit: persistence from benefit-receiving mothers to their daughters was about half (IGE=0.066) the level of persistence found among non-recipient mothers to their daughters (IGE=0.150), and persistence from benefit-receiving fathers to their daughters was strongly negative (IGE=-0.475) compared to the positive level of persistence found among non-recipient fathers to their daughters (IGE=0.153), although the former group of fathers was very small (n=39) so the result must be treated with caution.

Daughters' incomes and parents' incomes generally increased with parents' education level (for both mothers and fathers). Daughters whose mothers had no qualifications (IGE=0.150) or school qualifications (IGE=0.129) or trade qualifications (IGE=0.148) generally displayed similar levels of persistence to each other and to overall persistence. Daughters with degree-qualified mothers were the exception, with less than half the persistence in the total sample. This contrasts with the equivalent results for sons, who displayed higher persistence if their parents were degree-qualified. Daughters whose fathers had no school qualifications (IGE=0.079) or had degrees (IGE=0.090) displayed low levels of persistence. The standout results were for daughters with trade-qualified fathers, between whom the intergenerational association of incomes was negative, and for daughters whose fathers had school qualifications only, between whom there was substantial intergenerational persistence (IGE=0.545). For daughters, all group-specific estimates had confidence intervals that overlapped with those of the other groups.

### 3.5 Cross-national comparison

In this section, the NZLC IGE estimates for the four offspring-parent pairs are compared to estimates from the two other existing New Zealand studies and from studies of other countries. Estimates from comparator studies have been chosen to be as analogous as possible by selecting estimates that are (a) for the same offspring-parent pair (e.g., NZLC estimates for son-father pairs only compared to son-father estimates from other studies), and; (b) for a similar income concept (i.e., NZLC estimates for are only compared to estimates from other studies that also use individual total income, or failing that, individual labour earnings, for both generations), and; (c) where more than one study uses a comparable offspring-parent pair and comparable income concept, choosing the study that uses offspring birth cohorts and/or ages at which offspring income is observed that are closest to the NZLC birth cohorts (1967 to 1979 for sons, 1967 to 1974 for daughters) and ages of income observation (35 for sons and 40 for daughters).[82] Of course, even similar income concepts may be measured differently between studies (and in fact *are* measured differently – individual total income is defined slightly differently in each country/study considered here) and similar offspring-parent pairs may be defined differently (e.g., parents may be defined as biological parents only or defined more broadly). More generally, there are nearly always a range of differences in income definitions, sample selection criteria, time periods covered, and methodology – big or small – between any two or more studies. Consequently, precise rankings of countries are fraught with difficulty and should be interpreted with caution.

In Figures 21 to 24 below, the IGE estimates for each offspring-parent pair are plotted alongside comparable estimates from studies of New Zealand, Australia, Denmark, Finland, Norway, Sweden, Germany, Italy, Great Britain, Canada, and the US. Included beneath these figures are tables providing further details of the studies from which the estimates are taken, including the type of data source (whether the data was from an administrative or survey source), the income concept used for offspring (broadly defined), the particular birth cohorts studied (years that offspring were born), and the ages at which offspring's income is observed. For some countries, estimates plotted in one figure may come from a different source to estimates in another figure for the same country. For example, for Sweden, Björklund, Roine and Waldenström (2012) estimate the IGE of individual total income for son-father pairs, but not for any other offspring-

[82] Comparisons are also restricted to estimates from OLS estimation, as estimates from other approaches such as two-sample instrumental variables (TSIV) are upward-biased both in theory (since commonly-used instruments that are correlated with parental income can be expected to exert independent effects on offspring's income) and in practice (see the discussion in Blanden (2013) for evidence that OLS estimates tend to be smaller than IV estimates based on studies that use both approaches with the same sample), thus making estimates from the two approaches not directly comparable.

parent pairs, so daughter-mother and son-mother estimates for Sweden are taken from Österberg (2000) and a daughter-father estimate for Sweden is taken from Jäntti et al. (2006).

Deutscher and Mazumder (2020) use administrative tax data on sons and daughters born in Australia between 1978 and 1982 whose individual total incomes are observed over five-year windows across ages 29 and 37 (average age 33) and estimate the IGE of offspring's log individual total income with respect to their fathers' and mothers' log individual total incomes to be 0.157 among 414,100 son-father pairs, 0.105 among 385,800 daughter-mother pairs, 0.065 among 408,800 son-mother pairs, and 0.133 among 384,800 daughter-father pairs.[83] Munk, Bonke and Hussain (2016) use Danish administrative tax data on 261,248 sons born between 1966 and 1973 whose individual total incomes are averaged over five-year windows spanning ages 31 to 42 and estimate the IGE of sons' log individual total income with respect to their fathers' log individual total income to be 0.241.[84] Andrews and Leigh's (2008) study – discussed in section 2.4.2 of chapter 2 – uses data from the 1999 International Social Survey Programme on sons in New Zealand born between 1945 and 1974 whose hourly wage earnings are observed in 1999 (ages 25 to 54) and estimate the IGE of sons' log hourly earnings with respect to their father's imputed log hourly earnings (predicted from fathers' occupation and at age 40) to be 0.245. Gibbons' (2010) study, also discussed in section 2.4.2, uses data on 289 sons and 291 daughters drawn from the longitudinal Dunedin Study born in 1972-1973 whose individual total incomes are observed at age 32 and estimates the IGE of offspring's log individual total income with respect to their fathers' log individual total income to be 0.253 among son-father pairs and 0.167 among daughter-father pairs. Björklund et al. (2012) use Swedish administrative data on 108,277 sons born between 1960 and 1967 whose individual total incomes are observed over 10-year windows spanning ages 29 to 45 (average age 41) and estimate the IGE of sons' log individual total income with respect to their fathers' log individual total income to be 0.260.[85] Pekkarinen, Uusitalo and Kerr (2009) use Finnish administrative tax data on 20,824 sons born between 1960 and 1966 whose individual earnings are observed in the year 2000 (ages 34 to 40, average 37) and estimate the IGE of sons' log individual labour earnings with respect to their fathers' log individual labour earnings to be 0.277. Schnitzlein (2016) uses survey data on 408 sons drawn from the German Socio-economic Panel (GSOEP)[86] born between 1955 and 1976 whose individual earnings are observed over ages 35 to 42 (average of five observations, average age of 37) and estimates the IGE of sons' log individual labour earnings with respect to their fathers' log individual labour earnings to be 0.318 in Germany. Nilsen et al. (2012) use administrative tax data on 57,510 sons born in Norway between 1959 and 1962 whose individual earnings are averaged over three- to

[83] These estimates are for Australia as a whole, but Deutscher and Mazumder's (2020) focus is on estimating IGEs for regions within Australia. They are the first to use administrative data to estimate intergenerational income mobility in Australia, and the first to produce regional estimates for Australia, leading the authors to state that the paper "provides the most precise and comprehensive set of estimates of intergenerational income mobility for Australia to date" (Deutscher & Mazumder, 2020).

[84] This estimate of 0.241 is very close to Helsø's (2021) son-father estimate of 0.246 for Denmark based on administrative data on sons born between 1973 and 1975 whose individual labour earnings are averaged over windows of between one and seven years across ages 33 to 39.

[85] This estimate is virtually identical to Nybom and Stuhler's (2016) son-father estimate of 0.261 for Sweden based on administrative data on 3,460 sons born between 1955 and 1957 whose individual total incomes are averaged over ages 22 to 50 (average of 28 observations centred on age 35).

[86] The GSOEP is a nationally representative panel of approximately 11,000 households surveyed annually since 1984. The original sample consisted of 5,921 households containing 12,245 individuals who were resident in the Federal Republic of Germany (West Germany), later supplemented with samples from the German Democratic Republic (East Germany) in 1990, immigrants in 1994-1995, and other 'refreshment' and specialised subsamples thereafter.

five-year windows across ages 36 to 40 and estimate the IGE of sons' log individual labour earnings with respect to their fathers' log individual labour earnings to be 0.338. Chen et al. (2017) use administrative tax data on 261,871 sons and 216,214 daughters born between 1963 and 1966 whose individual total incomes are observed over three- to five-year windows across ages 38 to 42 (average age 40) and estimate the IGE of offspring's log individual total income with respect to their fathers' log individual total income to be 0.359 among sons and 0.254 among daughters. Bloise and Raitano (2021) use data on 718 sons drawn from the Italian component of the European Union Statistics on Income and Living Conditions household panel survey (matched to administrative social security records) who were born in Italy between 1972 and 1993 and whose individual earnings are observed in a single year between 2008 and 2014 (ages 21 to 36, average age 26) and estimate the IGE of sons' log individual labour earnings with respect to their fathers' log individual labour earnings to be 0.413. Bratsberg et al. (2007) use British survey data on 2,384 sons drawn from the National Child Development Study (NCDS)[87] whose individual earnings are observed at age 33 and/or age 41 and estimate the IGE of sons' log individual labour earnings with respect to their fathers' log individual labour earnings to be 0.450. Mazumder (2016) uses data on 327 sons drawn from the Panel Study of Income Dynamics (PSID)[88] born between 1951 and 1975 whose individual earnings are averaged over windows of between one and 10 years across ages 35 to 45 (where averages are centred on age 40) and estimates the IGE of sons' log individual labour earnings with respect to their fathers' log individual labour earnings to be 0.451 in the US.[89] Österbacka (2001) uses data on 16,070 sons and 14,763 daughters drawn from Finnish quinquennial census panel data (linked to tax records) born between 1950 and 1960 whose individual earnings are averaged over the 1985, 1990, and 1995 censuses (corresponding to ages 25, 30, and 35 for the youngest cohort and 35, 40, and 45 for the oldest cohort, with an average age of about 35 in 1990) and estimates the IGE of offspring's log individual labour earnings with respect to their mothers' log individual labour earnings to be 0.023 among daughters and 0.037 among sons. Österberg (2000) uses Swedish administrative tax data on 8,753 sons and 8,387 daughters born between 1941 and 1965 (and residing in Sweden in 1978) whose individual earnings are averaged over 1990, 1991, and 1992 (corresponding to ages 25 to 27 for the youngest cohort and 49 to 51 for the oldest cohort, with an average age of 37 in 1992),[90] and estimate the IGE of offspring's log individual labour earnings with respect to their mothers' log individual labour earnings to be 0.025 among daughters and 0.018 among sons. Bratberg, Nilsen and Vaage (2005) use data on 20,072 daughters drawn from Norwegian census panel data linked to tax records born in 1955 whose individual earnings are observed at age 40 and estimate the

[87] The NCDS is an ongoing longitudinal study tracking the health and development of a birth cohort of 17,415 children born in England, Wales, and Scotland in a single week in 1958.

[88] The PSID is an ongoing longitudinal study begun in 1968 of a nationally representative sample of about 4,800 American families interviewed annually until 1997 and biennially thereafter.

[89] This estimate is similar to US estimates by Gouskova, Chiteji and Stafford (2010), who estimated the son-father IGE at 0.410 based on PSID survey data on 535 sons born between 1948 and 1969 whose individual labour earnings are observed in a single year randomly chosen between ages 35 and 44 (on average, observed in the year 2000 at age 38), and by Schnitzlein (2016), who estimated the son-father IGE at 0.494 based also on PSID data on 462 sons born between 1955 and 1974 whose individual labour earnings are observed biennially between 1997 and 2009 encompassing ages 35 to 42 (average of 2.7 observations observed at an average age of 38).

[90] Österberg first adjusts offspring's and parents' incomes for age differences by regressing observed log earnings on age and age-squared, separately for each generation, sex, and year. The residuals from these regressions, representing age-adjusted earnings (or "income corrected for age"), are then used to compute expected earnings evaluated at the mean age for that generation plus the individual-specific residual (Österberg, 2000, p. 426). An average of age-adjusted log earnings over the three years (for all non-missing observations) is then calculated.

IGE of daughters' log individual labour earnings with respect to their fathers' log individual labour earnings to be 0.100.[91] Jäntti et al. (2006) estimate the IGE for daughter-father pairs in Denmark, Finland, Sweden, and Great Britain (in a cross-national comparison that includes estimates for other offspring-parent pairs and other countries). For Denmark, they use administrative data on 55,178 daughters drawn from the population register born between 1958 and 1960 whose individual earnings are observed in 1998 (ages 38 to 40) and 2000 (ages 40 to 42). For Finland, they use data on 5,144 daughters drawn from quinquennial census panel data born between 1958 to 1960 whose individual earnings are observed in 1993 (ages 33 to 35) and 2000 (ages 40 to 42). For Sweden, they use administrative data on 30,410 daughters drawn from the population register whose individual earnings are observed at ages 34 and 37. For Great Britain, they use data on 2,348 daughters drawn from the NCDS whose individual earnings are observed at ages 33 and 41. To adjust for age differences within and between countries, Jäntti et al. (2006) use the observed earnings data to predict what daughters and fathers in each country would have earned at the age of 40.[92] They estimate the IGE of daughters' log individual labour earnings with respect to their fathers' log individual labour earnings to be 0.034 in Denmark, 0.099 in Finland,[93] 0.204 in Sweden, and 0.331 in Great Britain. Figures 21, 22, 23, and 24 below display these IGE estimates for son-father, daughter-mother, son-mother, and daughter-father pairs, respectively, alongside the relevant NZLC estimate.

On these estimates, intergenerational persistence among sons with respect to their fathers is comparatively low in New Zealand. All three son-father IGE point estimates for New Zealand are very similar, around 0.24, positioning New Zealand as having similar levels of persistence to (three of the four) Nordic countries (Denmark, Sweden, Finland), with only Australia standing out as having lower persistence. Germany, Norway, and Canada then form a cluster of countries with IGE estimates in the 0.32 to 0.36 range, indicating higher persistence than New Zealand and the three other Nordic countries. At the least-mobile end of the spectrum, Italy, Great Britain and the US all have IGE estimates in the 0.41 to 0.45 range, substantially higher than the New Zealand estimates, although the confidence intervals of all three New Zealand point estimates overlap (with one exception) those for Italy, Great Britain, and the US. In New Zealand, sons' incomes bear a comparatively weak association to their fathers' incomes.

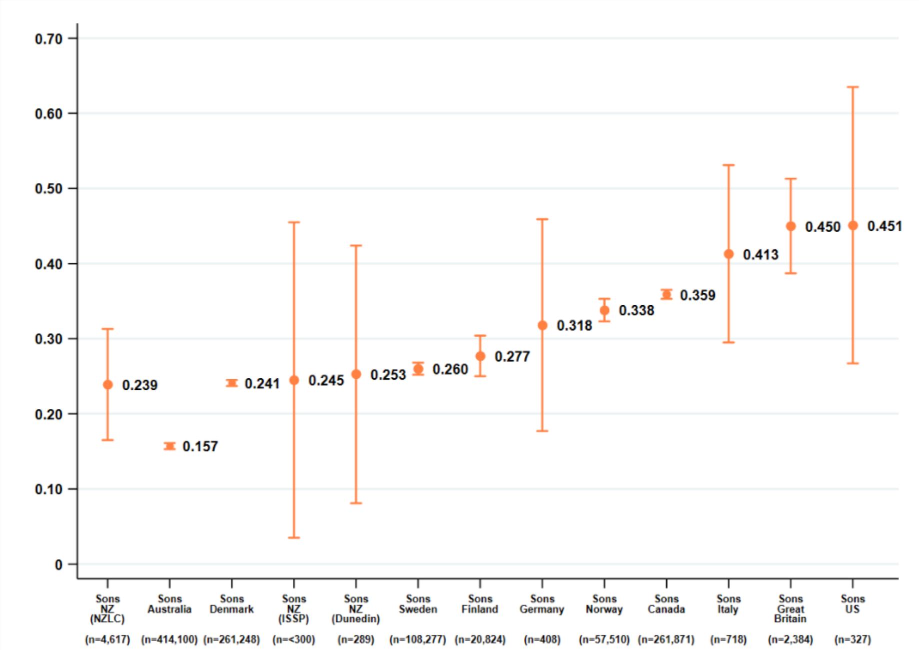
On the other hand, income persistence among daughters with respect to their mothers is comparatively high in New Zealand (albeit with only three other countries to compare against). The NZLC point estimate is 0.145, substantially higher than (and with a confidence interval that

[91] This estimate for Norway is similar to Jäntti and colleagues' (2006) daughter-father estimate of 0.121 for Norway, based on administrative data on 25,046 daughters born in 1958 whose individual earnings are observed at ages 34 and 41.

[92] The method by which this is done is described by Jäntti et al. (2006) as follows for fathers: "We inflate parental income to year 2000 values, then regress the natural log of earnings in the single outcome year [the year at which fathers' earnings are observed] on a quartic polynomial in age and record the residual from that regression. We then predict what their earnings would have been had they been 40 years old, add to this their estimated residual and take the anti-log". And for offspring: "For offspring, we also inflate the earnings to the year 2000 values, then regress the log of annual earnings on a year indicator and save the average of the OLS residual across the [two] years for each individual. We add to this the estimated time effect in the later year….and take the anti-log" (Jäntti et al., 2006, p. 6).

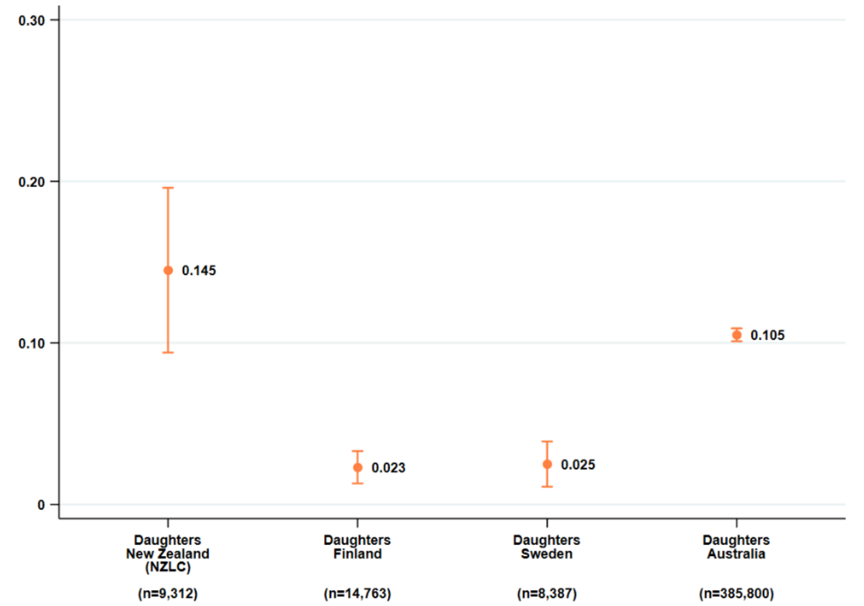
[93] This estimate for Finland is virtually identical to Österbacka's (2001) daughter-father estimate of 0.100 for Finland based on quinquennial census panel data on 20,181 daughters born between 1950 and 1960 whose individual earnings are averaged over the 1985, 1990, and 1995 censuses (corresponding to ages 25, 30, and 35 for the youngest cohort and 35, 40, and 45 for the oldest cohort, with an average age of about 35 in 1990).

**Figure 21**. *Comparison with other studies of intergenerational income elasticity estimate for sample of son-father pairs.*



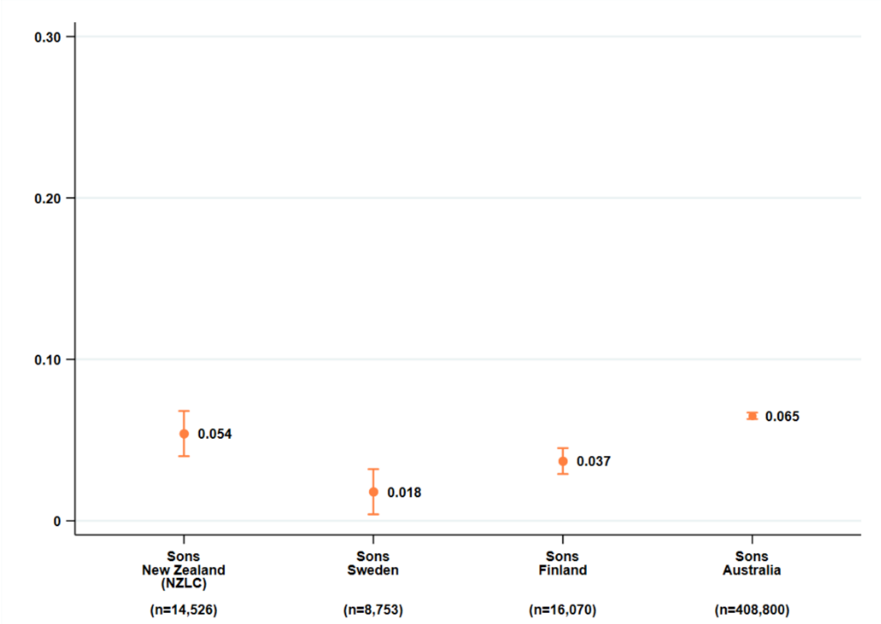

| *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|
| NZ (NZLC) | Survey | | Individual total income | 1967-1979 | 35 (centred) |
| Australia | Administrative | Deutscher & Mazumder (2020) | Individual total income | 1978-1982 | 29-37 |
| Denmark | Administrative | Munk et al. (2016, p. 32) | Individual total income | 1966-1973 | 31-42 |
| NZ (ISSP) | Survey | Andrews & Leigh (2008) | Individual labour earnings | 1945-1974 | 25-54 |
| NZ (Dunedin) | Survey | Gibbons (2010, p. 20) | Individual total income | 1972-1973 | 32 |
| Sweden | Administrative | Björklund et al. (2012, p. 478) | Individual total income | 1960-1967 | 29-45 |
| Finland | Administrative | Pekkarinen et al. (2009, p. 970) | Individual labour earnings | 1960-1966 | 34-40 |
| Germany | Survey | Schnitzlein (2016, p. 661) | Individual labour earnings | 1955-1976 | 35-42 |
| Norway | Administrative | Nilsen et al. (2012, p. 14) | Individual labour earnings | 1959-1962 | 36-40 |
| Canada | Administrative | Chen et al. (2017, p. 5) | Individual total income | 1963-1966 | 38-42 |
| Italy | Survey | Bloise et al. (2021, p. 101) | Individual labour earnings | 1972-1993 | 15-32 |
| Great Britain | Survey | Bratsberg et al. (2007, p. C83) | Individual labour earnings | 1958 | 33 and 41 |
| US | Survey | Mazumder (2016, p. 108) | Individual labour earnings | 1951-1975 | 40 (centred) |

**Figure 22**. *Comparison with other studies of intergenerational income elasticity estimate for sample of daughter-mother pairs.*



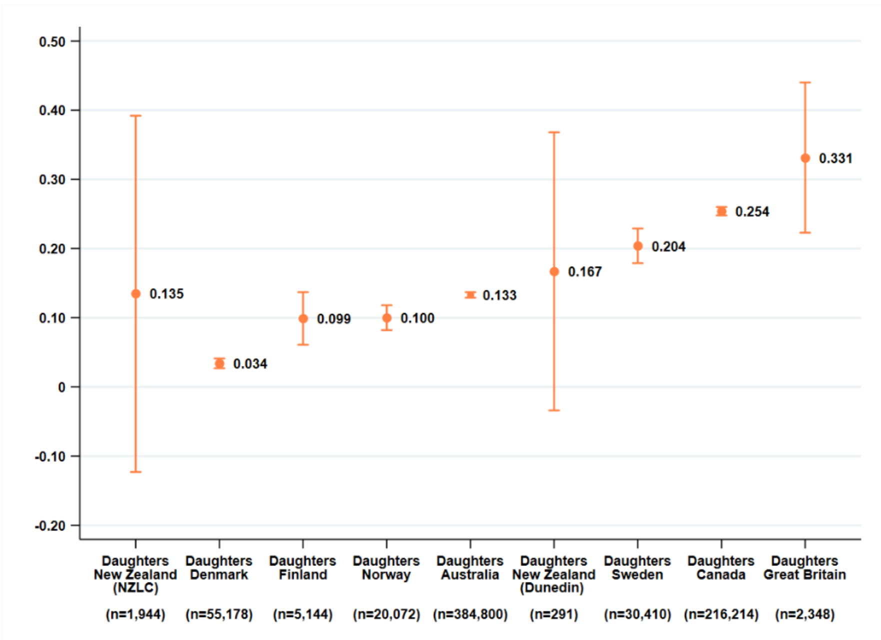

| *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|
| NZ (NZLC) | Survey | | Individual total income | 1967-1974 | 40 (centred) |
| Finland | Administrative | Österbacka (2001, p. 473) | Individual labour earnings | 1950-1960 | 25-45 |
| Sweden | Administrative | Österberg (2000, p. 426) | Individual labour earnings | 1941-1965 | 25-51 |
| Australia | Administrative | Deutscher & Mazumder (2020) | Individual total income | 1978-1982 | 29-37 |

**Figure 23**. *Comparison with other studies of intergenerational income elasticity estimate for sample of son-mother pairs.*




| *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|
| NZ (NZLC) | Survey | | Individual total income | 1967-1979 | 35 (centred) |
| Sweden | Administrative | Österberg (2000, p. 426) | Individual labour earnings | 1941-1965 | 25-51 |
| Finland | Administrative | Österbacka (2001, p. 473) | Individual labour earnings | 1950-1960 | 25-45 |
| Australia | Administrative | Deutscher & Mazumder (2020) | Individual total income | 1978-1982 | 29-37 |

**Figure 24**. *Comparison with other studies of intergenerational income elasticity estimate for sample of daughter-father pairs.*




| *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|
| NZ (NZLC) | Survey | | Individual total income | 1967-1974 | 40 (centred) |
| Denmark | Administrative | Jäntti et al. (2006, p. 13) | Individual labour earnings | 1958-1960 | 38-40 and 40-42 |
| Finland | Survey | Jäntti et al. (2006, p. 21) | Individual labour earnings | 1958-1960 | 33-35 and 40-42 |
| Norway | Administrative | Bratberg et al. (2005, p. 430) | Individual labour earnings | 1955 | 40 |
| Australia | Administrative | Deutscher & Mazumder (2020) | Individual total income | 1978-1982 | 29-37 |
| NZ (Dunedin) | Survey | Gibbons (2010, p. 20) | Individual total income | 1972-1973 | 32 |
| Sweden | Administrative | Jäntti et al. (2006, p. 21) | Individual labour earnings | 1962 | 34 and 37 |
| Canada | Administrative | Chen et al. (2017, p. 7) | Individual total income | 1963-1966 | 40 (centred) |
| Great Britain | Survey | Jäntti et al. (2006, p. 13) | Individual labour earnings | 1958 | 33 and 41 |

is non-overlapping with) the estimates of 0.02 for Finland and Sweden, and somewhat higher than (but not significantly different from) the 0.105 estimate for Australia. Thus, in New Zealand, daughters' incomes bear a comparatively strong association to their mothers' incomes (with emphasis on 'comparatively', as the absolute levels of persistence are low when considered against other offspring-parent pairs). Why might this be? Maloney et al. (2003) use data from the Christchurch Health and Development Study to estimate the intergenerational correlation in social welfare benefit receipt in New Zealand and find the correlation to be at least one-third, which is high compared to estimates in the Nordic countries and Canada (Moisio et al., 2015), and in a subsequent paper was found to be more than twice as high among daughters compared to sons (Pacheco & Maloney, 2003). An intergenerational association in welfare receipt in New Zealand that is both high by international standards and stronger among daughters may contribute to the comparatively high intergenerational income persistence among daughters with respect to their mothers found here.

With an IGE point estimate 0.054, income persistence in New Zealand among sons with respect to their mothers is similar to that in Finland and Australia (IGE estimates of 0.037 and 0.065, respectively), and somewhat higher than in Sweden (0.018). Thus, in New Zealand, the association between sons' incomes and their mothers' incomes is, comparatively speaking, neither high nor low, but middle of the range.

Finally, income persistence among daughters with respect to their fathers in New Zealand falls in the lower half of the cross-national range. The IGE point estimate of 0.135 is very similar to that for Australia[94] (0.133), and somewhat similar to the estimates for Finland and Norway (0.100) and Gibbons' (2010) estimate of 0.167 for New Zealand. Denmark stands out as having particularly low persistence (an IGE of 0.034), while persistence in Great Britain is estimated to be comparatively high (0.331), with Sweden and Canada positioned between the Nordic countries, Australia, and New Zealand, on the one hand, and Great Britain, on the other. Therefore, in New Zealand, the association between daughters' incomes and their fathers' incomes is in the low to middle range, comparatively speaking.

Taking the results for all four offspring-parent pairs together, New Zealand generally ranks as having a relatively low to middling degree of intergenerational income persistence, the exception being persistence between daughters and mothers, which is comparatively high. The finding of relatively high mobility of sons with respect to their fathers is consistent with the two other son-father estimates for New Zealand from Andrews and Leigh (2008) and Gibbons (2010). Gibbons (2010) also estimates the degree of intergenerational persistence in socioeconomic status in New Zealand (using data from the 1996 New Zealand Election Study,[95] thus a different data source from the one he uses to estimate income persistence) and compares his results to analogous results for Great Britain and Germany, finding socioeconomic mobility among both sons and

[94] Deutscher and Mazumder's (2020) Australian estimates for the same four offspring-parent pairs and with the same income concept make for an interesting comparison. Their point estimates are remarkably close to the estimates for New Zealand reported here (with the exception of son-father pairs) and the two countries occupy similar cross-national positions across all four pairs.

[95] The New Zealand Election Study is a nation-wide survey of New Zealand electors conducted triennially after each general election since 1990.

daughters (with respect to their fathers) is slightly higher in New Zealand than in Great Britain (although not statistically significantly so) and among sons only is significantly higher than in Germany, again providing some corroboration to the finding reported here. It is also consistent with Davis's (1979, pp. 55, 54-55) finding, based on an analysis of 2,530 sons drawn from a nationally-representative health survey who reported their occupation and their father's occupation when they were aged 15 which were coded to six social classes, that intergenerational social mobility among sons with respect to their fathers "is a widespread phenomenon" in New Zealand, a result which is "not consistent with any image of New Zealand as a 'closed' or rigidly stratified society"[96] (albeit a finding relating to older cohorts of sons from an earlier period in New Zealand).[97] When Davis (1979, p. 53) compares his results with those from two similar studies from Australia and the US, he finds that New Zealand shows "the weakest evidence for occupational inheritance" and is "the most 'open' of the three countries".[98] Furthermore, the low son-father IGE estimate is consistent with results from a series of studies that find intergenerational educational mobility in New Zealand to be very high by international comparison. Studies reporting cross-national estimates of intergenerational educational persistence from a regression of offspring's years of schooling on parents' years of schooling – each using different data sources for New Zealand – have found New Zealand ranks as second-most mobile of 30 countries analysed in Lee and Lee (2021), third-most mobile of 111 countries analysed in Narayan et al. (2018), and sixth-most mobile of 42 countries analysed in Hertz et al. (2007).[99]

[96] An analysis of intergenerational socioeconomic mobility among 1,096 son-father pairs by Nash (1993) – part of the 1989 Access and Opportunity in Education Project involving interviews with parents from three North Island areas (Upper Hutt, Wanganui, Hawke's Bay) who reported their own occupation and their father's occupation when they were 14 years old which were coded to six socioeconomic strata – found that 31 percent of sons occupied the same socioeconomic status as their father and that the professional and managerial strata were particularly 'open' to upward mobility from the working class (that is, had relatively low persistence). In conjunction with his results from an analysis of class and educational exogamy, Nash (1993, pp. 61, 50) surmises that "[g]iven the rate of social mobility into the middle class and the importance of educational qualifications to marriage, it hardly makes sense to speak of class closure [intergenerational persistence] in contemporary New Zealand" and concludes that his results can be seen as consistent with a "strong belief in the notion of New Zealand as an 'open' society in which those with talent and a willingness to work can achieve success regardless of their social origins".

[97] An analysis of intergenerational social class mobility among 3,795 son-father pairs in southern Dunedin over an even earlier period (1881 to 1940) - a research programme known as the Caversham Project named after the Dunedin borough which was the initial case study - concluded that intergenerational mobility was high, with sons showing "little inclination to follow in their father's footsteps" who "easily struck out on their own in different occupations" (Olssen et al., 2011, pp. 168, 214). Occupations were coded to nine social classes and cross-tabulated in a transition matrix, which revealed that 37 percent of sons were in the same social class as their father, 31 percent of sons were upwardly mobile, and 32 percent were downwardly mobile. However, rates of intergenerational persistence differed by social class, being considerably higher among employers (at the top of the class hierarchy) and manual workers (at the bottom of the class hierarchy). For instance, among small employers which included self-employed men, "[s]ome 66 per cent of these sons, on their wedding day, had fathers in the same stratum, a clear case of intergenerational transfer of cultural and financial capital" (Olssen et al., 2011, pp. 148-149). Comparing their results to methodologically-similar studies of contemporaneous England and Boston, they find intergenerational social mobility in southern Dunedin was considerably higher than in England but lower than in Boston.

[98] However, a subsequent comparative analysis of social mobility in New Zealand and Australia by Jones and Davis (1988) using the same health survey data from Davis (1979) alongside nationally-representative survey data from Australia, with occupations coded to eight social classes, found that, when compared in terms of exchange mobility (that is, abstracting from changes in the occupational structure in each country), "[m]obility patterns in Australia and New Zealand are well-nigh indistinguishable" but "to the extent that differences across countries do exist, New Zealand exhibits somewhat more openness in its mobility regime" (Jones & Davis, 1988, pp. 280, 282).

[99] See also De Broucker and Underwood (1998) and Pfeffer (2008) who use the same data source for New Zealand as Hertz et al. (2007) – the 1996 International Adult Literacy Survey – but estimate educational mobility using different methods and find that New Zealand ranks as fourth-most mobile of 11 countries analysed (De Broucker & Underwood, 1998) and third-most mobile of 20 countries analysed (Pfeffer, 2008). In addition, OECD (2018) find that among offspring whose parents have only a lower secondary education, the likelihood of these offspring attaining a tertiary education is 13 percent on average across the OECD, but 35 percent in New Zealand, the most-mobile of the 30 countries they analyse based on this measure (on other measures of educational mobility estimated by OECD (2018), New Zealand occupies a middling cross-national ranking).

### 3.6 Bias and limitations

The NZLC data has some important limitations including a number of potential sources of bias. These include linkage bias (where the characteristics of individuals who are linked across censuses differ systematically from those who are not linked), the likely presence of sets of linked records that pertain to the same individual but are not linked together (call them 'disconnected records'), and the quality of the income data including its banded and self-reported nature, each discussed in turn below.

*3.6.1 Linkage bias*

Linkage bias in the NZLC has been investigated by Clarkson (2014) and Singhal (2015) (both of whom examine linkage bias over 1981 to 2006, as the 2013 census had not been added at the time of their analyses). Clarkson (2014) identifies the factors most strongly associated with 'linkage status' (whether an individual is linked or not linked) among six different 'cohorts' or panels of different lengths in the NZLC, five of which were cohorts linked across two censuses, and one linked across three censuses. She does this by producing a correlation matrix of census variables, showing the association between each pair of variables controlling for all other variables, including a variable denoting 'linkage status' (but excluding some highly collinear variables). Based on the strength of the partial correlations with linkage status across all cohorts, she finds that probability of linkage is lower among those who had changed address since the last census, non-homeowners, males, Māori, non-partnered people, and those living in more socioeconomically deprived neighbourhoods.[100]

Note, however, that the presence of differential attrition over time in survey research does not necessarily induce selection bias, which only occurs if the exposure-outcome or predictor-outcome association differs between individuals retained in the sample compared to those who drop out (attriters). Carter, Imlach-Gunasekara, McKenzie, and Blakely (2012) demonstrate this empirically using data from waves one to three of the longitudinal Survey of Family, Income and Employment (SoFIE). Briefly, they compare an exposure-outcome association (odds ratios of poor self-rated health predicted by employment status and education level, controlling for socio-demographic confounders) between those retained in the sample and attriters. They show that while people who continue to participate in SoFIE have different characteristics to attriters, the exposure-outcome associations for each group (as measured by the odds ratios) were not statistically significantly different from each other and converged even further after adjusting for covariates in the multivariable model. Hence, they find "little evidence of selection bias" affecting their statistical analyses, despite the observed differential attrition (Carter et al., 2012, p. 220).

Singhal (2015) explores the potential for linkage bias to induce selection bias by producing a correlation matrix of census variables for the theoretical population at time t, and a correlation matrix of the same variables among census respondents linked from t to t-1, and then compares these two correlation matrices. If the individuals in the linked sample differ systematically from the theoretical population, the two sets of correlations will differ markedly. Singhal considered a

[100] Clarkson (2014) notes that because correlation coefficients only measure linear relationships, the age variable – which had a non-linear relationship to linkage status – was not found to be a significant predictor using partial correlations.

correlation to be biased if the difference in correlations was greater than 0.01. The proportion of biased correlations in a particular linked sample constitutes its linkage-induced selection bias. Singhal carries out this procedure separately for children and adults, and separately for each cohort (that is, he begins by estimating linkage bias for cohorts linked across two censuses, then across spans of three censuses, four censuses, five censuses, and finally for the single cohort of individuals linked across the six censuses from 1981 to 2006). He finds that, across the various cohorts, between 40 and 76 percent of correlations were unbiased. Of relevance to the current analysis is that *cohorts linked back to 1981 tended to be the least biased* relative to other cohorts of the same length.[101]

What, then, is the likely impact of linkage bias on the current results? The answer depends on whether intergenerational persistence differs between people linked and not linked in the NZLC, and if so, to what extent. Changing address since the previous census is the strongest predictor of linkage status, so the question arises of whether intergenerational persistence among individuals who are more residentially mobile is likely to differ from persistence among those less residentially mobile. New Zealand research has found that residential movement is more frequent among children of Māori and Pacific ethnicity, born to younger and less educated mothers, living in private rental accommodation, and living in more deprived neighbourhoods (Morton et al., 2014; Morton et al., 2020; Robertson et al., 2021).[102] Does intergenerational persistence differ by these characteristics? Results from section 3.4.3 found intergenerational persistence to be *generally slightly lower* among Māori than among European offspring (although not significantly so), and *generally similar or slightly higher* among offspring with unqualified mothers compared to qualified mothers. Given offspring with low-qualified mothers are fairly represented in the samples, but Māori offspring are under-represented (see section 3.3), the IGE estimates are likely to be slightly biased in an upward direction (overstating persistence).

*3.6.2 Emigrants*

As discussed in section 3.1, the linking methodology used to construct the NZLC means that people who were overseas during one or more censuses (hence did not complete census forms), but who subsequently returned to New Zealand, cannot be linked back over the period they were away (and plainly people who permanently emigrate from New Zealand cannot be linked over the period since their departure). This can be considered a form of linkage bias, in that people linked in the NZLC are, by construction, more likely to be people who choose to stay in New Zealand, whereas those not linked are more likely to be people who choose to leave New Zealand.[103] What

[101] Singhal (2015) also develops a set of statistical weights to compensate for linkage bias, using logistic regression models that predict linkage status by significant census variables.
[102] Morton et al. (2014) also find that children with European mothers, above-average household incomes, and who experienced a change in household income or change in parental partnership status are also more residentially mobile.
[103] Or at least those not linked are more likely to be people who emigrate *for the medium to long-term*, since short stints overseas during an intercensal period will not affect linkage provided no censuses are missed. Also note that it is possible for an individual to have two or more sets of links in the NZLC that are disconnected (not linked together), as explained by Didham and Nissen (2015, p. 1): "[S]omeone may be linked in, say, two pairs covering 1981-1991 and again in the 2001-2006 pair, as may happen for many people who were either out of New Zealand or not enumerated for any of the intervening censuses" but "[t]he possibility of linking between these two sets of links has not yet been investigated" so "people may enter and drop-out [of the NZLC], but if they re-enter in the current datasets this re-entry is currently not identified". Importantly, the NZLC samples analysed here will not contain 'duplicate individuals', that is, offspring or parents cannot appear twice (appearing as if they were unique individuals) in the samples, because the offspring populations from which the samples are drawn, and their parents, are all defined as at 1981 and subsequently followed through time until their linked record terminates (so any further sets of links that may exist for these individuals are effectively ignored.

impact might the exclusion of emigrants have on the estimates of intergenerational income persistence? The question is difficult to answer conclusively given the paucity of evidence in New Zealand. Many people (no doubt including many children aged under 15 years in 1981) move overseas to earn higher incomes than they could in New Zealand (among other 'pull' factors). But for a study of intergenerational persistence, the crucial question is *what did their parents earn?* There is little data available on the incomes of New Zealand-born emigrants and their parents to answer this question, but Gibbons' (2010) study contains some indicative evidence.

Recall that Gibbons (2010) compared the incomes of 780 offspring born in Dunedin in 1972-1973 and observed at age 32 (one-quarter of whom were living overseas at that time, mostly in Australia and Britain) with the incomes of their fathers (measured when the offspring were aged 13 and 15, and their fathers aged about 41 on average). The incomes of overseas-resident offspring observed at age 32 were converted into New Zealand dollars using purchasing power parity exchange rates, and all incomes were deflated to 2008 values. Gibbons thus had a subsample of New Zealand-born emigrants (n=188). He estimated the intergenerational income elasticity in the sample *both excluding and including these emigrants* (that is, a model of New Zealand-resident offspring only versus a model including all offspring regardless of country of residence at age 32). Any difference in the IGEs from these two models can be taken to imply greater or lesser intergenerational persistence among emigrants compared to those who stay in New Zealand. Gibbons' estimates of the son-father IGE are 0.253 for New Zealand-resident sons and 0.290 for all sons regardless of where they lived. Similarly, his estimates of the daughter-father IGE are 0.167 for New Zealand-resident daughters versus 0.215 for all daughters. Thus, as Gibbons (2010, p. 49) notes, intergenerational persistence is higher when emigrants are included in the sample, implying that persistence is somewhat stronger among those who emigrate from New Zealand compared to those who stay. Since those living overseas had higher incomes on average than those who remained in New Zealand, stronger persistence among emigrants *reflects stronger intergenerational transmission of advantage among these families*. Therefore, the current estimates of the IGE using NZLC data which excludes emigrants by construction are probably biased downwards (under-estimating persistence).

*3.6.3 Quality of income data*

Some important limitations to the analysis arise from the way income information is collected in the census. As discussed in section 3.1.3, the total personal income variable is collected in bands (interval-censored) with the upper category being open-ended and the bottom category being censored from below. What effect might this have on the IGE estimates? This question is addressed in Gregg, Macmillan, and Vittori's (2017) study estimating intergenerational income persistence among sons in Britain using data from, *inter alia*, the British Cohort Study (BCS, born 1970). In the BCS, parental income is observed twice, once in 1980 when the cohort child was aged 10 and again in 1986 at age 16, and the data is banded at both time-points. Gregg, Macmillan, et al. derive a continuous income variable by fitting a Singh-Maddala distribution to the bands, using maximum likelihood estimation to estimate an expected value for each band. They then take an average of parents' income over the two time-points, wherever possible. To assess the measurement error created from using banded data, they use administrative data (the

UK's Annual Survey of Hours and Earnings) to obtain more granular earnings information on a group of males of a similar age to the fathers in their BCS sample who are observed over a 17-year period encompassing the two years at which the BCS fathers were observed. They then group the continuously-measured earnings of these males into bands with the same proportions observed in the BCS distribution and replicate the BCS method of deriving an expected value for each band. This enables them to compare the variance in the error derived from this method with that for the original continuous data (in each of the two years separately, and overall). Gregg, Macmillan, et al. (2017, p. 4 of appendix) find that

> …the errors from our banded data at [ages] 10 and 16 are small compared to using non-banded earnings measures and non-banded average earnings at two points in time. Assuming classical measurement error, this exercise suggests that the attenuation factor from having banded data is approximately 92% and 95% that which we would have with continuous data at those given ages. Averaging across the banded data at two points in time gives approximately 95% of that which we would have from averaged continuous data.

This result implies that the effect of banded parental income data on estimates of the IGE is small.[104] Gregg, Macmillan, et al. (2017) also face the issue of an open-ended top category, where the long tail of high incomes is condensed to a single value. The total personal income variable from the NZLC is subject to the same limitation. Gregg, Macmillan, et al. (2017, p. 4 of appendix) consider this issue in their robustness checks, and judge that "averaging over two periods reduces this problem [of top-coding] and when combined with the larger issue of transitory income fluctuations through childhood, the variance of the errors based on continuous earnings averaged at age 10 and 16 and the same based on banded data are very similar". They therefore conclude that they "have explored the implications of measurement error from individuals' choosing a band rather than giving an actual value in the [BCS] by using administrative data and find this to be minimal" (Gregg, Macmillan, et al., 2017, p. 86).

In their comparison of intergenerational income persistence in Great Britain and Sweden, for which the British data are from the BCS survey and the Swedish data are from administrative tax records, Björklund et al. (2017) maximise cross-national comparability by converting their continuously-measured Swedish parental income data to discrete intervals to mimic the banded parental income data collected in the BCS. As a by-product of this data harmonisation, they are able to compare their Swedish IGE estimates when using banded versus continuous parental incomes. They find that the son-parent IGE drops by seven percent and the daughter-parent IGE drops by five percent with banding, concluding "we have also conducted a sensitivity analysis with the Swedish data by comparing the main results based on intervals with results from using the actual data" and find that "[t]he differences are small and do not affect any of our conclusions" (Björklund et al., 2017, pp. F77-F78).

Similarly, in their cross-country comparison of intergenerational earnings persistence in Denmark, Finland, Norway, Great Britain, and the US, for which the British parental income data are banded but those for the other countries are continuous, Raaum et al. (2007, p. 15) conduct "[s]ensitivity

[104] In an earlier paper using the same BCS dataset, Blanden, Gregg, and Macmillan (2011, p. 44 of appendix B) conduct the same robustness check on the use of banded data, this time using continuous income data from the British Household Panel Survey (BHPS), which they group into bands containing the same proportions as in the BCS sample, and find "negligible differences" in their equivalent variance analysis between the continuous and banded BHPS data.

checks based on data from the other countries in which we mimic the income bracketing of the UK data" and find these checks "yield results that are very similar to those reported [the main results] based on non-bracketed parental earnings data".

Ueda's (2009, p. 14) study of intergenerational income persistence in Japan uses panel survey data that contains both banded and continuous data on fathers' income for a subset of participants, enabling her to quantify the effect of banding on estimates of the IGE, from which she concludes that "the use of grouped parental income slightly reduces the estimates [of the IGE]".

Thus, based on these previous findings, the current IGE estimates may be slightly biased downwards, understating the true level of intergenerational persistence (although recall that a multi-year average is used for parents' proxy income, which produces a semi-continuous parental income variable, thus reducing the impact that banding has on estimation).

A final limitation is that the total personal income variable from the census is, of course, self-reported, and hence subject to respondent errors: individuals could misunderstand the income question, intentionally or unintentionally misreport their income, or not answer the question at all.[105] With regard to the latter, non-response rates to the total personal income question are high compared with other census questions and have increased over time, from 5.5 percent in 1986 to 9.7 percent in 2013 (Errington et al., 2008; Stats NZ, 2021c). Such relatively high non-response can potentially introduce bias if non-response rates differ between groups being analysed.

In this context, it is pertinent to consider Suei's (2016) comparison of income data from the census with income data from administrative tax records (neither of which is free from measurement error). She compares 2013 census data on total personal income and sources of personal income with estimates of the same derived from Inland Revenue tax data available in Stats NZ's Integrated Data Infrastructure (IDI) as at May 2015, with the caveat that some incomes sources will be captured in the census but not in the tax data, and vice versa.[106] She uses the tax data to derive an estimate of individuals' total personal income that is as analogous as possible to that captured by the census, then aggregates these estimates into the bands used in the 2013 census, in order to compare the two. In her individual-level comparison, she links usual residents aged 15 years and over in the 2013 census with their administrative records (if any) in the IDI and then compares the income information between the two. Suei (2016) finds that 41 percent of individuals have the same income band in both census and tax data, and 75 percent are within one band, with discrepancies most common for income from student allowances, accident compensation

[105] See Statistics New Zealand (2006, 2009) for a discussion of the main reasons that the total personal income question is not well answered, including social aversion (particularly a perceived connection to government departments monitoring individuals), embarrassment due to low income or reticence about revealing high incomes, lack of involvement with the management of income in the home or not knowing the gross amount received from a benefit or pension, and difficulty assessing one's total annual income from all sources due to multiple income sources or variability in income during the previous year.

[106] Sources captured in the census but not in tax data include non-taxable income sources (e.g., non-taxable benefits like the accommodation supplement), investment income from interest and dividends (collected by Inland Revenue but not available in the IDI at the time of Suei's analysis), support payments from people who do not live in the respondent's household, income from private superannuation, and income earned overseas and not taxed in New Zealand. The IDI tax dataset will therefore not contain anyone for whom these sources are their only source of income. Furthermore, the census captures individuals with no sources of income or zero income, many of whom will have no requirement to file tax returns and hence will not appear in the tax data.

payments, and self-employment. While some people over-report their income in the census compared to administrative figures, and others under-report it, overall the discrepancies "are weighted towards negative differences, where tax income is higher than census income" which "may be partly due to census respon[dents] incorrectly reporting net instead of gross income" (Suei, 2016, p. 23).[107] Suei (2016, p. 30) concludes that:

> Comparisons of income bands between the census and an equivalent administrative population show a reasonably consistent distribution, except for a large discrepancy for the $15,001 to $20,000 income band. Most tax-derived income in this band is from government benefits and New Zealand Superannuation, and the discrepancy appears to reflect lack of coverage in the tax data of additional income sources for these groups. A comparison of income distribution restricted to only wages and salaries shows very close agreement between the census and the tax data. The results, for the most part, are encouraging for the quality of income data provided by census respondents.

What effect might misreporting and non-reporting of income in the census, and any potential inaccuracies vis-à-vis administrative sources, have on estimates of intergenerational persistence? These data limitations are only likely to bias estimates of the IGE if non-response and misreporting differ between parents and offspring in a way that systematically skews the true intergenerational income association. While non-response rates have increased over time, it is not currently known whether these systematically differ by parental income, and similarly there is no evidence for differences in misreporting by parental income. Accordingly, the direction of any bias is unclear and this is an area in need of further research.

[107] Suei (2016) notes that this underreporting of income in census data is seen across all income bands and has been documented in other countries as well.

## Chapter 4: The Christchurch Health and Development Study

In this chapter, data from the Christchurch Health and Development Study are used to estimate intergenerational income persistence and mobility in New Zealand. Section 4.1 describes the data source and income variables used in the analysis. Section 4.2 describes the methods used to construct the long-run income variables and the selection of the samples analysed. Section 4.3 presents descriptive statistics of the sociodemographic and income characteristics of the samples and assesses their representativeness. Section 4.4 presents results of estimates of the intergenerational income elasticity and estimates of persistence and mobility in income rank. Section 4.5 compares these results with those from other countries, and section 4.6 compares the results to those from the NZLC analysis in chapter 3. Section 4.7 discusses the limitations of the analyses.

### 4.1 Data source and income variables

#### *4.1.1 Description of data source*

The Christchurch Health and Development Study (CHDS) is a longitudinal study tracking the health and development of a birth cohort of 1,265 children (635 sons, 630 daughters, which includes 14 pairs of twins) born in Christchurch, New Zealand in the four-month period 15 April to 5 August 1977. The cohort was recruited by contacting mothers giving birth in all maternity units, both public and private, in the Christchurch urban region and inviting them to participate. Of the 1,310 mothers giving birth to live-born children during this period, 1,251 agreed to participate in the research, accounting for 97 percent of children. At the time of writing, this cohort had been studied on 24 occasions, namely: at birth, 4 months postpartum, at annual intervals between ages 1 and 16 years, then 18, 21, 25, 30, 35, and most recently at 40 years, when 904 cohort members were studied. Data has been collected from multiple sources, including from the children's parents, their teachers, the children themselves, and from administrative records, as detailed below:

- Parent interviews: At birth, four months, and annually from ages 1 to 16 years, parent interviews were conducted at the child's home with the child's mother (or mother figure), or with the child's father (or father figure) if the father was the sole parent in the family. Where there was a change in a child's parents or caregivers since the previous interview, the *main custodial caregiver at that age* was chosen for interview. Interviews typically lasted for one to two hours and addressed a range of health outcomes relevant to the child's stage of development as well as parental income and family socio-demographic information.

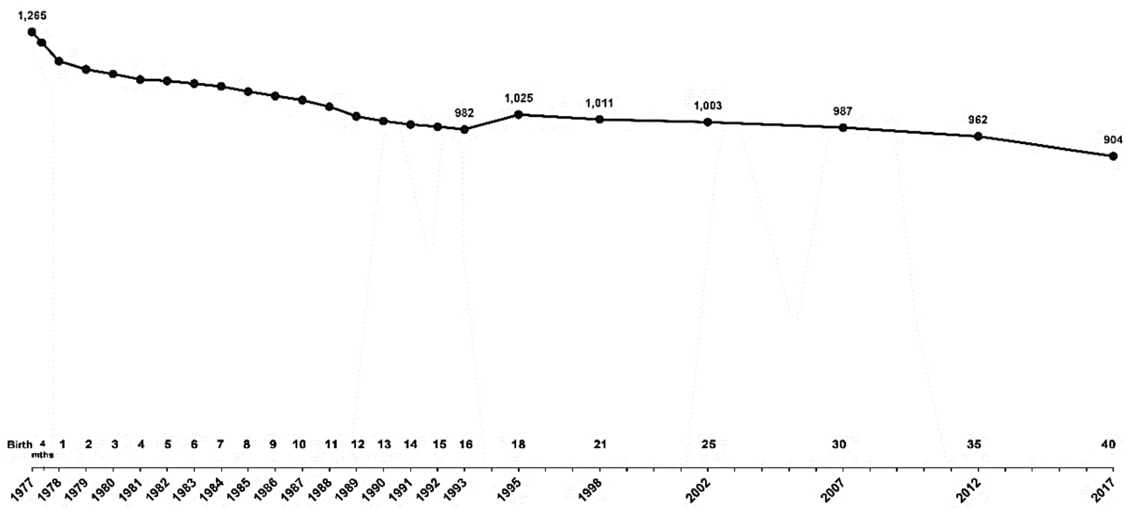
- Cohort member assessments during childhood and interviews in adulthood: From the age of eight to 18 years, children were interviewed about topics relevant to their developmental stage and completed standardised tests of intelligence, reading, mathematical and scholastic ability. From age 21, cohort members were interviewed about a wide range of demographic, economic, social, and health and lifestyle issues.

- Teacher questionnaires: From the age of six to 13 years, teacher reports on child academic achievement and social adjustment were obtained annually by supplying all class teachers of the children with a set of standardised questionnaires. Compliance was very high with completed questionnaires being obtained for over 98 percent of children in any given year.

- Administrative records of cohort members were obtained including police records, hospital records and information from the school dental service.[108]

Figure 25 below displays the years at which the cohort have been followed-up (corresponding to data collection waves), the age of the cohort at each of these timepoints, and the number of cohort members followed-up at each timepoint (the black line). The black line in Figure 25 shows the extent of sample attrition, such that by age 40, 71 percent of the original cohort were followed-up. By international standards, retention in the CHDS study is very respectable compared to other panel studies of similar length. For example, in Great Britain, the National Child Development Study and British Cohort Study achieved retention rates of 62 percent and 57 percent, respectively, at their 42-year follow-up sweeps (that is, 42 years after their samples were initially interviewed), while in the US, the National Longitudinal Study of Youth 1979 achieved a 69 percent retention rate 39 years after beginning (Center for Human Resource Research, 2018; Centre for Longitudinal Studies, 2021).

**Figure 25**. *Years and ages of follow-up, and sample attrition, of the CHDS cohort.*




**Source:** Sample attrition figures sourced from Fergusson et al. (1989, p. 306), Fergusson and Horwood (2001, p. 288), and Boden, Newton-Howes, Foulds, Spittlehouse and Cook (2019, p. 19).

[108] All phases of the study have been subject to ethical approval by the regional Health and Disability Ethics Committee. Prior to age 14 all data were collected on the basis of signed parental consent only; from age 14 onwards, signed consent was also obtained from the cohort members (Fergusson & Horwood, 2001, 2013; Fergusson, Horwood, Shannon, & Lawton, 1989).

As detailed in Fergusson and Horwood (2001), the majority of sample losses occurred in the early years of the study and the reasons for loss to follow-up have been: (1) emigration from New Zealand, because up to the age of 16 years cohort members who migrated outside of New Zealand were treated as being no longer eligible for interview; (2) refusal to participate; (3) mortality (that is, cohort members who have died), and; (4) failure to trace (that is, cohort members who could not be tracked down). Note that cohort members who moved overseas *in adulthood* are eligible for interview, and these interviews are administered online (via videoconferencing), or by telephone, or for some cohort members living in Australia, face-to-face (interviewers are flown to Australia to conduct interviews in person).

Furthermore, there has been differential attrition over time, that is, biases in which cohort members are lost to follow-up. This issue of differential attrition is discussed further in section 4.7, but broadly the CHDS researchers have found small but statistically significant tendencies for the subset of the cohort who are retained in the study to under-represent individuals from socially disadvantaged backgrounds characterised by low parental education, low socioeconomic status, and single parenthood (Fergusson & Horwood, 2001; Fergusson, Horwood, & Boden, 2008).

Table 14 breaks down the cohort members followed-up at ages 30, 35, and 40 by sex and country of residence at those ages. The table shows that of the 987 cohort members interviewed at age 30, 478 were sons (of whom 112 were living overseas) and 509 were daughters (of whom 116 were living overseas). By age 40 years, 904 cohort members were followed up, consisting of 431 sons (87 living overseas) and 473 daughters (89 living overseas). The cohort members followed-up at any given wave are almost entirely a subset of those followed-up at the previous wave (for example, 98 percent of the sons and 99 percent of the daughters interviewed at age 40 were also interviewed at age 35). The samples of offspring used in this analysis are drawn from the CHDS cohort members enumerated in Table 14.

**Table 14***. Cohort members followed-up at ages 30, 35, and 40 years, by sex and country of residence.*

| | Country of residence | 30-years | 35-years | 40-years |
|---|---|---|---|---|
| **Sons** | New Zealand | 366 | 352 | 344 |
| | Overseas | 112 | 111 | 87 |
| | Total | 478 | 463 | 431 |
| **Daughters** | New Zealand | 393 | 388 | 384 |
| | Overseas | 116 | 111 | 89 |
| | Total | 509 | 499 | 473 |
| **Grand total** | | **987** | **962** | **904** |

**Source**: Christchurch Health and Development Study dataset.

### *4.1.2 Income concepts and measurement*

The CHDS has collected various measures or components of income for cohort members (offspring) and parents. The income concepts captured by these measures vary by whether they are before-tax (gross) or after-tax (net), by income source (from employment, from all market sources, or from government transfers), and by whether they were collected from offspring or parents or both. Parents reported their income, and that of their spouse, when their child was aged 1 to 14 years. Specifically, parents reported their weekly gross market income in bands (numbering between 24 and 32 across data collection waves, with an open-ended top category) and their weekly net social welfare benefit income in dollars per week. Since most parent interviews were conducted with mothers, fathers' incomes are almost completely mother-reported. Mothers were asked: "*Looking at the card, what is the code number corresponding to your average gross weekly income (over the last 3 months)?*" and "*What is the code number corresponding to your husband's average gross weekly income (over the last 3 months)?*". Mothers were then asked, "*Does your husband receive any Social Welfare Benefit?*" and "*What is his weekly benefit payment?*", followed by "*Do you receive any Social Welfare Benefits other than Family Benefit?*" and "*What is your weekly benefit payment?*".[109] Appendix Table 6 shows the weekly gross market income bands used at each parent interview from child ages 1 to 14 years.

Offspring reported their own incomes in adulthood (at ages 30, 35, and 40 years) and those of their cohabitating partner/spouse (if they had one).[110] Specifically, they reported their weekly net employment earnings, their weekly net benefit income, and their weekly income from other sources, all in dollars per week. Offspring were asked "*How much per week do you currently receive from paid employment, after tax? (Include all current jobs, if earnings vary record average weekly earnings over past month)*, "*How much do you receive in benefit payments per week?*", and "*Do you currently receive income from any other source, e.g. investments, donations from parents, etc?*".[111] The same questions were then repeated for the offspring's partner (hence partners' incomes are offspring-reported). Figure 26 shows the years and ages at which parents' and offspring's (and offspring's partners') incomes were collected.

These income data are used to quantify two long-run economic concepts: lifetime average earnings and family income. The former pertains to the labour income an individual generates over his or her lifetime. The latter pertains to all sources of income received by all members of an individual's family (measured in this analysis in both the parental generation during the offspring's childhood and in the offspring generation during their own adulthood, pooling across spouses in both generations) and captures "the family's realized economic standing rather than the family's underlying capacity to generate income" (Mitnik & Grusky, 2020, p. 51). Family income has been argued to be a preferable measure of parental income than the commonly-used fathers' earnings because it "incorporates the income of mothers and thus better indexes the full complement of economic resources available to invest in children", and "reflects the ability of families to draw on income sources other than earnings in response to transitory earnings shocks" and "avoids any

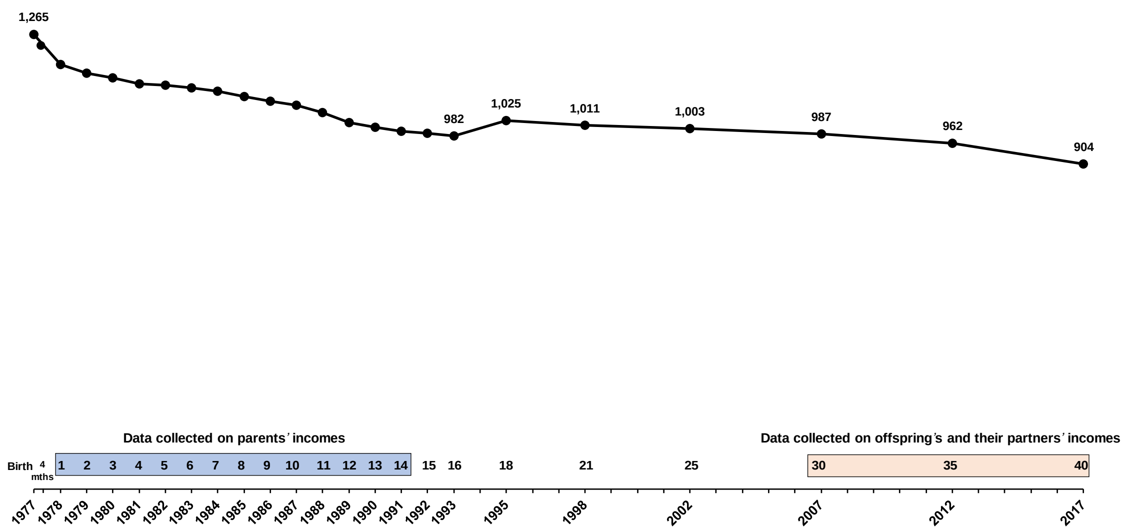
[109] Respondents were instructed to include National Superannuation, New Zealand's universal public pension payment, as part of any social welfare benefit income received. From 1987 (the 10-year follow-up), respondents were also instructed to "[i]nclude Family Support payments only if paid as part of DPB [Domestic Purposes Benefit], etc." Family Support was a tax credit granted to working families and beneficiaries, the latter receiving it as a cash payment added to their main benefit.

[110] Offspring's own income was also collected at ages 16, 18, 21, and 25, but this analysis only uses income data collected at ages 30, 35, and 40, for reasons outlined in section 4.2. For offspring living overseas, the CHDS researchers converted their incomes to New Zealand dollars using purchasing power parity conversion.

[111] Two other income measures were collected – offspring's hourly gross earnings from their main job and offspring's annual gross total personal income – but these are not used in this analysis.

selection bias that may result from omitting children with absent fathers” (Mitnik et al., 2015, p. 40).[112]

**Figure 26**. *Years and ages at which income information was collected from parents of CHDS cohort members over childhood and from cohort members themselves in adulthood.*




## 4.2 Method

### *4.2.1 Construction of income variables*

The CHDS income data are combined in various ways to derive the three income variables used in this analysis: two dependent variables measuring offspring income (offspring's lifetime average earnings, offspring's family income over adulthood) and one independent variable measuring parental income (offspring's family income over childhood).

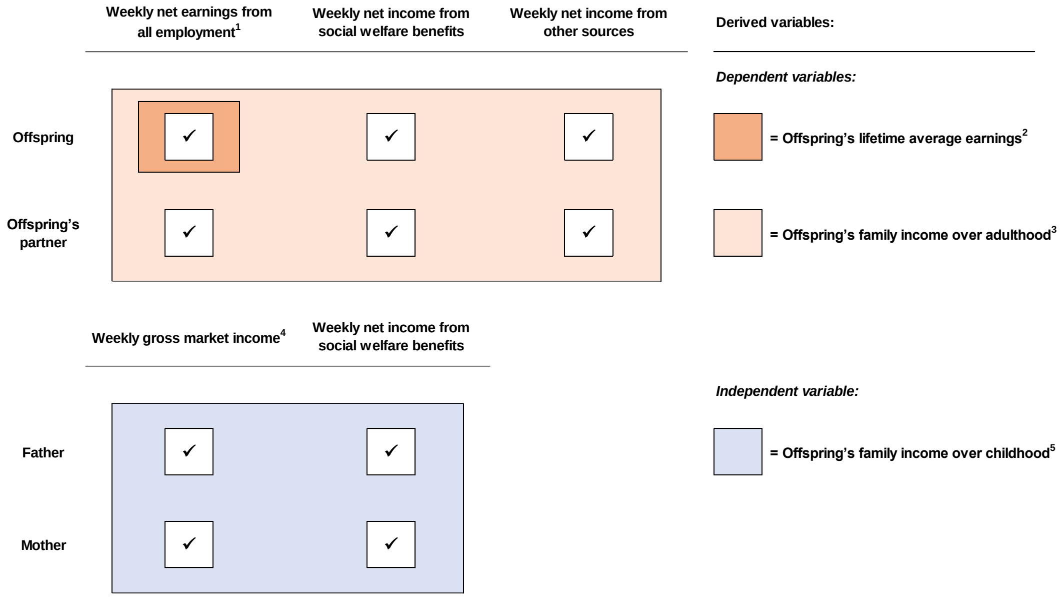
Figure 27 summarises the income concepts and variables collected in the CHDS, the derivation of the two dependent variables and one independent variable used in the current analysis, and the types of analyses to which each derived income variable is put. This is followed in the text by a more detailed description of how each of the three income variables was constructed. For this analysis, all incomes were deflated to September 2012 dollars using the consumer price index.

***Dependent variable 1*: offspring's lifetime average earnings**

To construct a proxy for offspring's lifetime average weekly earnings (the term 'proxy earnings' is used here interchangeably), Mazumder's (2016) method is once again employed, namely, taking a multiyear average of income observations centred on an age at which current income is known to be representative of lifetime average income (the 'proxy age'). As with the analysis of NZLC data in chapter 3, a proxy age of 35 years for sons and 40 years for daughters is chosen, based on the international empirical evidence listed in Table 1 in section 2.3.5 of chapter 2.

[112] In addition, measures of family income – even short-run measures – "rarely equal zero", so their use largely avoids the selection bias that can arise from dropping individuals with zero income who often have weaker labour force attachment (Mitnik & Grusky, 2020, p. 52).

**Figure 27**. *Income information collected from CHDS cohort members in adulthood and their parents in childhood and income variables derived from these.*




Notes:

[1] Where earnings varied from week to week, respondents were asked for their average weekly earnings over the past month.

[2] Offspring's weekly net earnings from all employment, averaged over ages 30, 35, and 40 years for sons where possible (with the average centred on age 35, else the 35-year observation is used) and at age 40 years only for daughters.

[3] Combined weekly net total income of offspring and partner, averaged over ages 30, 35, and 40 years where possible (else a single year is used).

[4] Respondents were asked for their average weekly market income over the past three months.

[5] Combined weekly gross total income of mother and father, averaged over at least eight years of the offspring's childhood (ages 1 to 14 years), where the CHDS researchers have scaled net (after-tax) weekly benefit income by a factorof 1.2 to convert it to an approximate gross benefit income.

As illustrated in Figure 28 below, proxy earnings are constructed following these steps:

1. For sons, a multiyear average is taken of weekly earnings at ages 30, 35, and 40, including years of zero earnings (an average which is centred on – or has a median age of – 35 years).

2. If that is not possible (for example, due to missing earnings observations), an average is taken of earnings at ages 30 and 40 (so that, once again, the average is centred on age 35).

3. If that is not possible, the single 35-year observation is used for the proxy. Note that because the proxy age for daughters is 40 (and multiyear averages centred on age 40 are not possible), only daughters' single-year earnings at age 40 is used as their proxy, and hence only step 3 applies to daughters.

**Figure 28**. *Method for construction of a proxy for lifetime average weekly earnings of CHDS cohort members, using sons as an example.*

| | | Year of follow-up | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Year son born | | 1993 | 1995 | 1998 | 2002 | 2007 | 2012 | 2017 |
| 1977 | Son's age | 16 | 18 | 21 | 25 | 30 | 35 | 40 |

| Proxy earnings: | Earnings observations used: | Median age: |
|---|---|---|
| Multiyear average | 2007, 2012, 2017 / 30, 35, 40 | 35.0 years |
| | else | |
| | 2007, 2012, 2017 / 30, ., 40 | 35.0 years |
| else | else | |
| Single year observation | 2012 / 35 | 35.0 years |

**Notes:** Daughters' lifetime average earnings are proxied by their single-year earnings observation at age 40, equivalent to the final row in Figure 28 for sons.

Calculating proxy earnings in this fashion is not possible for all offspring. A small number of offspring have missing earnings observations at timepoints that are needed for the construction of a proxy. For example, in the case of a son who was not interviewed at ages 30 and 35 but *was* interviewed at age 40, a proxy for his lifetime average weekly earnings *centred on age 35* cannot be constructed. Similarly, a proxy for lifetime average earnings cannot be constructed for daughters who have missing responses to the earnings question at age 40. In addition, some sons and daughters have proxy earnings of zero, and these offspring are necessarily dropped from the sample due to the log-log specification (recall that individual years of zero earnings *are* permitted for sons, but sons' *proxy earnings* must be positive to be included in the analysis). However, a robustness check is performed in which weekly earnings of zero are all coded as $1 and the proxy for lifetime average earnings then reconstructed.

***Dependent variable 2:*** **Offspring's family income over adulthood**

To construct a proxy for offspring's family income over adulthood, variables measuring offspring's weekly net total income (income from all three sources collected in the CHDS), and that of their partner (if they have one), at ages 30, 35, and 40, are used. Weekly net total incomes are derived by the CHDS researchers by summing net employment earnings, net social welfare benefit income, and net income from other sources. As a proxy for offspring's family income over adulthood, the weekly net total incomes of the offspring and their partner at each age are summed,[113] and then an average of the (non-missing) combined incomes is taken over adulthood (ages 30, 35, and 40). If a multiyear average cannot be taken due to there being only one valid family income observation, that single observation is used.

[113] If the offspring has no partner in any given year, the offspring's own weekly total income is used as their family income in that year.

*Independent variable:* **Offspring's family income over childhood**

To construct a proxy for offspring's family income over childhood, the weekly gross family income variables at each age from 1 to 14 years are used, derived by the CHDS researchers as the sum of the weekly gross market incomes and weekly net social welfare benefit incomes of the offspring's mother and father.[114] An average of weekly gross family income is taken over ages 1 to 14, subject to a minimum of eight valid observations (i.e., more than half of the 14 family income observations over childhood must be non-missing). However, a robustness check is performed to test the sensitivity of the results to different minima of non-missing family income observations.

### *4.2.2 Selection of offspring samples*

For the analysis of offspring's lifetime average earnings, sons in the CHDS must be followed-up until at least age 35, while daughters must be followed-up at age 40, and both must meet other selection criteria shown in Figure 29, resulting in sample sizes of 413 sons and 358 daughters. These sons and daughters are pooled into an 'all offspring' sample (n=771) and separate analyses are performed for all offspring (controlling for sex), sons only, and daughters only.[115] For the analysis of offspring's family income over adulthood, sons and daughters need only be followed-up until age 30 and meet other selection rules shown in Figure 30, resulting in final sample sizes of 450 sons and 476 daughters. Once again, separate analyses are performed for all offspring (n=926), sons only (n=450), and daughters only (n=476).

## 4.3 Descriptive statistics of samples

### *4.3.1 Sociodemographic characteristics of offspring and parents*

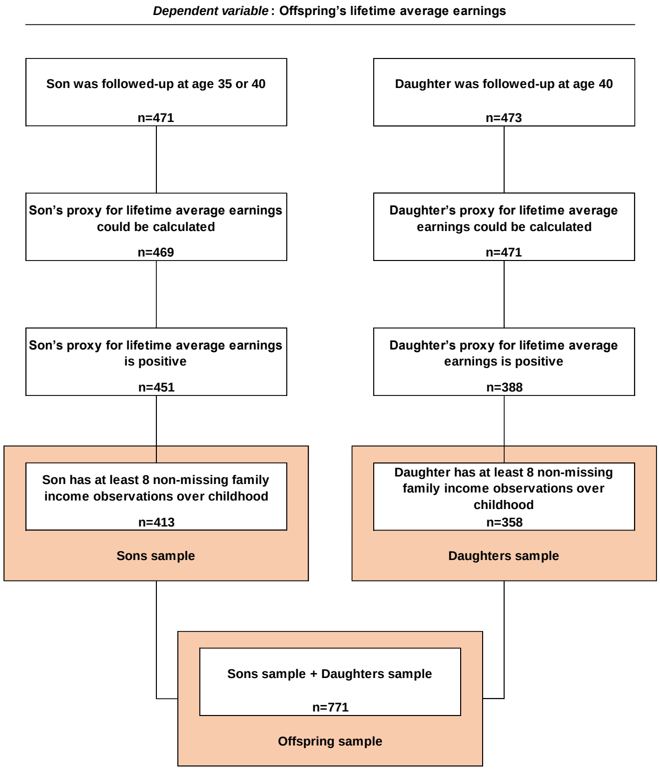
Table 15 presents, for each dependent variable, the sociodemographic characteristics of the offspring, their parents, and their partners in each of the three samples (all offspring, sons, daughters). At age four months, when parents' relationship to the cohort child was first established by the CHDS researchers, 96 percent of mothers and 90 percent of fathers were birth parents. Six percent of children had no father figure present in the household. This is reflected in family type, where 94 percent of children were living in two-parent families and six percent were living in single-parent families. Approximately 70 percent of mothers were aged between 20 to 29 years at the birth of their child, with only about one percent aged 40 years or over. Fathers were older on average, being mostly aged between 25 to 34 at the birth of their child. Almost all mothers and fathers (93 percent and 90 percent, respectively) were New Zealand European (or from other non-Māori non-Pacific ethnic groups), five percent of mothers and eight percent of fathers were Māori, and about two percent of mothers and fathers were of a Pacific ethnicity.

---

[114] In deriving the weekly gross family income variable, the CHDS researchers assigned parents the mid-point of their weekly market income band, while the open-ended top category was coded at the lower bound of the band (e.g., "Over $300" was coded at $300). They also scaled net (after-tax) weekly benefit income by a factor of 1.2 to convert it to an approximate *gross* value.

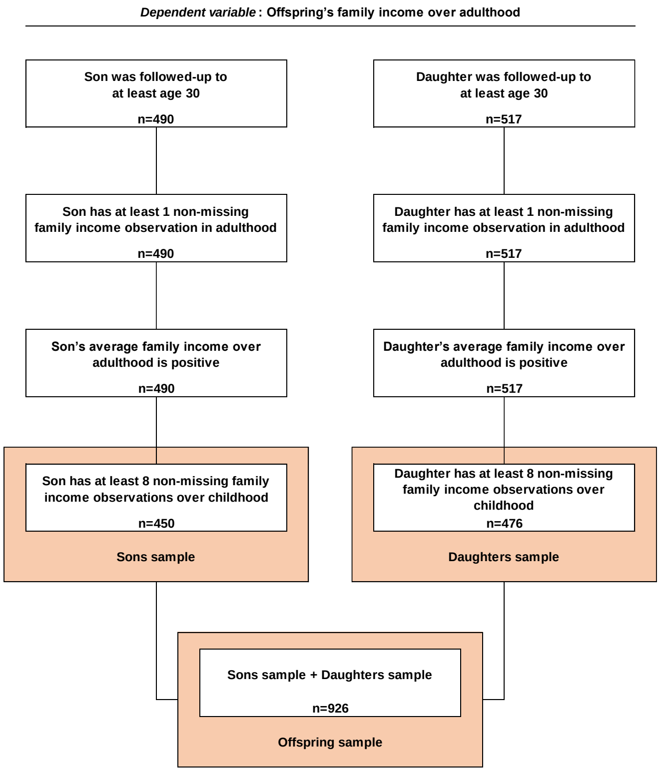
[115] Mitnik et al. (2015, p. 40) argue that sex-specific models should be favoured over models that pool sons and daughters, on the grounds that pooling could "run the risk of conflating gender-based advantage with parental-income advantage", a risk "especially relevant for earnings IGEs given that employment rates differ by gender and parental income". This is less of a problem for the models with offspring's family income in adulthood as the dependent variable since most offspring are partnered and nearly all partnerships are male-female couples.

**Figure 29**. *Flowchart showing how samples of offspring from the CHDS cohort were selected for analysis with dependent variable lifetime average earnings.*




**Source**: Christchurch Health and Development Study dataset.

**Figure 30**. *Flowchart showing how samples of offspring from the CHDS cohort were selected for analysis with dependent variable family income over adulthood.*




**Source**: Christchurch Health and Development Study dataset.

**Table 15**. *Sociodemographic characteristics of CHDS offspring and parents in each sample for each dependent variable.*

| Variable | Category | *Dependent variable*: Offspring's lifetime average earnings | | | *Dependent variable*: Offspring's family income over adulthood | | |
|---|---|---|---|---|---|---|---|
| | | Sample | | | Sample | | |
| | | All offspring (%) | Sons (%) | Daughters (%) | All offspring (%) | Sons (%) | Daughters (%) |
| Mother's relationship to offspring at age four months | Birth parent | 96.6 | 96.4 | 96.9 | 95.8 | 95.3 | 96.2 |
| | Adoptive parent | 2.6 | 3.1 | 2.0 | 3.1 | 3.8 | 2.5 |
| | Foster parent | 0.1 | 0.0 | 0.3 | 0.2 | 0.0 | 0.4 |
| | Other parent figure | 0.0 | 0.0 | 0.0 | 0.1 | 0.2 | 0.0 |
| | Missing | 0.6 | 0.5 | 0.8 | 0.8 | 0.7 | 0.8 |
| Father's relationship to offspring at age four months | Birth parent | 90.1 | 89.8 | 90.5 | 88.9 | 88.7 | 89.1 |
| | Adoptive parent | 2.6 | 3.1 | 2.0 | 3.1 | 3.8 | 2.5 |
| | Foster parent | 0.1 | 0.0 | 0.3 | 0.2 | 0.0 | 0.4 |
| | Other parent figure | 0.6 | 0.7 | 0.6 | 0.8 | 0.9 | 0.6 |
| | No parent | 5.8 | 5.8 | 5.9 | 6.3 | 6.0 | 6.5 |
| | Missing | 0.6 | 0.5 | 0.8 | 0.8 | 0.7 | 0.8 |
| Offspring's family type at offspring's birth | Two-parent family | 93.8 | 94.4 | 93.0 | 93.6 | 87.3 | 88.9 |
| | Single-parent family | 6.2 | 5.6 | 7.0 | 6.4 | 12.7 | 11.1 |
| Mother's age at offspring's birth | 19 years and under | 8.7 | 8.5 | 8.9 | 9.2 | 8.4 | 9.9 |
| | 20-24 years | 29.7 | 32.4 | 26.5 | 30.7 | 33.6 | 27.9 |
| | 25-29 years | 40.6 | 38.3 | 43.3 | 40.4 | 37.6 | 43.1 |
| | 30-34 years | 16.1 | 14.5 | 17.9 | 15.3 | 14.7 | 16.0 |
| | 35-39 years | 3.8 | 4.4 | 3.1 | 3.5 | 4.0 | 2.9 |
| | 40 years and over | 1.2 | 1.9 | 0.3 | 1.0 | 1.8 | 0.2 |
| Father's age at offspring's birth | 19 years and under | 2.5 | 2.4 | 2.5 | 2.6 | 2.2 | 2.9 |
| | 20-24 years | 20.6 | 21.8 | 19.3 | 21.4 | 22.7 | 20.2 |
| | 25-29 years | 37.6 | 37.3 | 38.0 | 38.2 | 36.9 | 39.5 |
| | 30-34 years | 26.1 | 25.9 | 26.3 | 24.9 | 25.3 | 24.6 |
| | 35-39 years | 9.2 | 8.0 | 10.6 | 8.6 | 8.2 | 9.0 |
| | 40 years and over | 3.9 | 4.6 | 3.1 | 4.0 | 4.4 | 3.6 |
| | Missing | 0.1 | 0.0 | 0.3 | 0.2 | 0.2 | 0.2 |
| Mother's ethnicity at offspring's birth | New Zealand European and other[1] | 93.6 | 92.7 | 94.7 | 93.4 | 92.2 | 94.5 |
| | Māori[2] | 4.8 | 5.1 | 4.5 | 4.8 | 5.3 | 4.2 |
| | Pacific | 1.6 | 2.2 | 0.8 | 1.8 | 2.4 | 1.3 |
| Father's ethnicity at offspring's birth | New Zealand European and other[1] | 90.3 | 88.9 | 91.9 | 90.0 | 88.0 | 91.8 |
| | Māori[2] | 7.5 | 7.7 | 7.3 | 7.6 | 8.0 | 7.1 |
| | Pacific | 2.1 | 3.4 | 0.6 | 2.4 | 4.0 | 0.8 |
| | Missing | 0.1 | 0.0 | 0.3 | 0.1 | 0.0 | 0.2 |
| Mother's highest qualification at offspring's birth | No formal qualification | 48.6 | 47.0 | 50.6 | 49.8 | 48.2 | 51.3 |
| | Secondary school qualification | 30.2 | 33.9 | 26.0 | 30.0 | 33.1 | 27.1 |
| | Tertiary qualification | 21.1 | 19.1 | 23.5 | 20.2 | 18.7 | 21.6 |
| Father's highest qualification at offspring's birth | No formal qualification | 44.2 | 44.9 | 44.4 | 46.1 | 44.7 | 47.5 |
| | Secondary school qualification | 33.3 | 34.8 | 32.4 | 31.7 | 33.3 | 30.3 |
| | Tertiary qualification | 20.1 | 20.2 | 20.4 | 19.3 | 19.3 | 19.3 |
| | Missing | 2.3 | 0.0 | 2.8 | 2.8 | 2.7 | 2.9 |
| Father's weekly annualised market income at 1 year | Nil | 2.2 | 2.4 | 2.0 | 2.2 | 2.7 | 1.7 |
| | $1-$1,499 | 0.5 | 0.7 | 0.3 | 0.4 | 0.7 | 0.2 |
| | $1,500-$2,999 | 0.3 | 0.5 | 0.0 | 0.3 | 0.4 | 0.2 |
| | $3,000-$4,999 | 2.7 | 2.7 | 2.8 | 2.7 | 2.7 | 2.7 |
| | $5,000-$6,999 | 9.1 | 10.4 | 7.5 | 10.4 | 11.1 | 9.7 |
| | $7,000-$9,999 | 41.0 | 40.0 | 42.2 | 39.7 | 39.6 | 39.9 |
| | $10,000 and over | 29.7 | 29.1 | 30.4 | 29.7 | 28.4 | 30.9 |
| | No father | 7.4 | 7.7 | 7.0 | 7.9 | 8.2 | 7.6 |
| | Missing | 7.1 | 6.5 | 7.8 | 6.7 | 6.2 | 7.1 |
| Offspring's country of residence at 40 years[3] | New Zealand | 79.2 | 76.8 | 82.1 | 78.6 | 77.8 | 79.4 |
| | Australia | 14.1 | 14.5 | 13.7 | 14.3 | 13.6 | 14.9 |
| | United Kingdom | 3.2 | 4.1 | 2.2 | 2.9 | 3.8 | 2.1 |
| | Europe | 0.6 | 0.5 | 0.8 | 0.8 | 0.4 | 1.1 |
| | USA/Canada | 1.2 | 1.7 | 0.6 | 1.5 | 2.0 | 1.1 |
| | Asia | 1.3 | 2.2 | 0.3 | 1.4 | 2.0 | 0.8 |
| | Other | 0.3 | 0.2 | 0.3 | 0.5 | 0.4 | 0.6 |
| Offspring's highest qualification by 40 years[3] | No formal qualification | 4.5 | 5.6 | 3.4 | 6.2 | 6.9 | 5.5 |
| | Secondary or tertiary level 4 qualification | 48.6 | 55.0 | 41.3 | 50.4 | 56.2 | 45.0 |
| | Tertiary level 5 or 6 qualification | 12.2 | 9.9 | 14.8 | 11.7 | 9.3 | 13.9 |
| | Bachelors degree | 26.8 | 23.2 | 31.0 | 24.8 | 21.8 | 27.7 |
| | Postgraduate degree | 7.8 | 6.3 | 9.5 | 6.9 | 5.8 | 8.0 |
| Offspring's main occupation at 40 years[3,4] | Managers | 24.1 | 28.6 | 19.0 | 20.9 | 26.8 | 15.3 |
| | Professionals | 26.7 | 23.0 | 31.0 | 25.0 | 21.7 | 28.2 |
| | Technicians and trades workers | 13.9 | 20.3 | 6.4 | 12.9 | 20.3 | 5.9 |
| | Community and personal service workers | 8.8 | 6.3 | 11.7 | 8.9 | 5.8 | 11.8 |
| | Clerical and administrative workers | 10.4 | 2.2 | 19.8 | 11.0 | 2.0 | 19.5 |
| | Sales workers | 3.6 | 2.7 | 4.7 | 3.5 | 2.5 | 4.4 |
| | Machinery operators and drivers | 5.2 | 8.5 | 1.4 | 4.6 | 8.0 | 1.3 |
| | Labourers | 7.1 | 8.5 | 5.6 | 7.4 | 8.5 | 6.3 |
| | Not in paid employment | 0.1 | 0.0 | 0.3 | 6.0 | 4.5 | 7.4 |
| Offspring's partnership status at 40 years | Partnered | 75.2 | 74.1 | 76.5 | 69.4 | 68.4 | 70.4 |
| | Not partnered | 20.8 | 18.4 | 23.5 | 20.5 | 23.2 | 21.4 |
| | Missing | 4.0 | 7.5 | 0.0 | 10.0 | 14.2 | 8.2 |
| Partner's highest qualification by 40 years[3] | No formal qualification | 4.7 | 3.4 | 6.1 | 5.4 | 4.0 | 6.7 |
| | Secondary school qualification | 24.8 | 26.6 | 22.6 | 24.2 | 24.9 | 23.5 |
| | Tertiary qualification below degree | 25.8 | 19.1 | 33.5 | 24.6 | 18.0 | 30.9 |
| | Bachelors degree | 27.0 | 32.0 | 21.2 | 25.2 | 30.2 | 20.4 |
| | Postgraduate degree | 6.7 | 7.5 | 5.9 | 6.6 | 7.1 | 6.1 |
| | No partner or missing | 11.0 | 11.4 | 10.6 | 14.0 | 15.8 | 12.4 |
| Partner's main occupation at 40 years[3,4] | Managers | 14.4 | 9.4 | 20.1 | 14.6 | 9.1 | 19.7 |
| | Professionals | 22.6 | 23.5 | 21.5 | 21.7 | 22.2 | 21.2 |
| | Technicians and trades workers | 11.0 | 3.6 | 19.6 | 11.2 | 3.6 | 18.5 |
| | Community and personal service workers | 5.6 | 7.5 | 3.4 | 5.4 | 7.3 | 3.6 |
| | Clerical and administrative workers | 9.7 | 15.5 | 3.1 | 8.6 | 14.4 | 3.2 |
| | Sales workers | 4.0 | 5.1 | 2.8 | 3.7 | 4.7 | 2.7 |
| | Machinery operators and drivers | 3.6 | 0.2 | 7.5 | 3.8 | 0.2 | 7.1 |
| | Labourers | 5.8 | 3.4 | 8.7 | 5.2 | 3.1 | 7.1 |
| | Not in paid employment | 12.1 | 20.1 | 2.8 | 11.7 | 19.3 | 4.4 |
| | No partner or missing | 11.2 | 11.6 | 10.6 | 14.1 | 16.0 | 12.4 |
| Total number of observations | | 771 | 413 | 358 | 926 | 450 | 476 |

**Notes**:
[1] 'Other' refers to ethnic groups that are not Māori, Pacific, or New Zealand European.
[2] The Māori category includes all parents who reported some Māori ancestry.
[3] If responses to these variables are missing at age 40 years, responses at age 35, or else 30, are used instead.
[4] Occupations are coded to the major group level of the Australian and New Zealand Standard Classification of Occupations 2006.

**Source**: Christchurch Health and Development Study dataset.

Approximately half of mothers, and just under half (45 percent) of fathers, had no school qualifications at the time of their child's birth; about one-third of mothers and fathers had a secondary school qualification; and about 20 percent had a tertiary qualification. The distribution of parental characteristics between the samples of sons and daughters are similar, with the following exceptions. Compared to sons, daughters were, on average, born to older mothers, were slightly more likely to have European fathers, were less likely to have Pacific fathers, were more likely to have mothers with either no qualifications or a tertiary qualification and less likely to have mothers with secondary qualifications.

Turning to the characteristics of the offspring (all offspring pooled), at age 40 years about 20 percent were living overseas (mostly in Australia and the United Kingdom). By age 40, about 62 percent of offspring had a secondary or level 4 tertiary qualification as their highest qualification, about 25 percent had a Bachelor degree, and seven percent had a postgraduate degree. Just under half (44 percent) were in managerial or professional occupations, with the next most common occupations being technicians and trades workers (12 percent) and clerical and administrative workers (10 percent). About 14 percent were not in paid employment.[116] Again, some differences can be seen between the samples of sons and daughters. Compared to sons, daughters at age 40 were slightly more likely to be living in New Zealand, to be tertiary qualified, and be working in professional, community and personal service, and (especially) clerical occupations, or not be in the labour force, and were less likely to be managers, technicians and trades workers, and machinery operators and drivers.

About 70 percent of offspring had a partner at age 40. The distribution of partners' educational qualifications, and the distribution of partners' occupation, are both very similar to those of offspring.[117] Approximately five percent of partners had no formal qualification, about one-quarter had a Bachelor degree, and about seven percent had a postgraduate degree. Approximately 42 percent of partners were in managerial or professional occupations and 14 percent were not in paid employment. Differences in partner characteristics between the samples of sons and daughters are apparent. Compared to sons, daughters were more likely to have a partner with a pre-degree tertiary qualification and a partner who is a manager, technician or trades worker, machine operator, or labourer (and less likely to have a partner who is not in paid employment).

*4.3.2 Assessing sample representativeness*

How representative of the population of New Zealand (and indeed of Christchurch) are the parents and offspring in the samples? To assess the representativeness of the samples, one would need to compare the sociodemographic and other characteristics of the CHDS samples with those of the target population of interest – the population about whom the analysis is trying to make

[116] This proportion refers to the sample with 'offspring's family income over adulthood' as the dependent variable. Not surprisingly, in the sample with 'offspring's lifetime average earnings' as the dependent variable, the proportion who were not in paid employment at age 40 is considerably lower (three percent). Compared to the former sample, the latter sample also contains higher proportions of offspring in managerial and professional occupations (about three percent higher in each category) and of partnered status (75 percent versus 69 percent). All other differences in the proportions of sociodemographic characteristics between the two dependent variables are small in magnitude (less than three percent).

[117] This is the case once the 'No partner or missing' category is excluded, and, in the case of highest educational qualification, once differences between offspring and partners in the coding of this variable are taken into account (for offspring, level 4 tertiary qualifications are categorised together with secondary school qualifications, whereas for partners they are coded as 'tertiary qualification below degree').

inferences – which is, for the analysis in this chapter, *offspring born in New Zealand around the late 1970s and their parents*. For the parent-related variables in Table 15, which were mostly collected in 1977, such a comparison could be made with analogous statistics from the 1976 census (where published). However, problems with such a comparison present themselves immediately. First, New Zealand's census publications generally do not report sociodemographic statistics specifically for *parents* (as a subset of all adults or those aged 15 years and over), yet it is parents who are the target population. Therefore, the characteristics of mothers and fathers in the CHDS could generally only be compared with those of the total female and male populations in 1976 (sometimes it may be possible to restrict further to the total female and male populations aged 15 years and over), which can distort the comparison. For example, when mothers' market incomes were first collected in the CHDS in 1978, 75 percent reported zero market income. By comparison, among the total female population aged 15 years and over in the 1976 census, 41 percent reported 'Nil' income (from all sources excluding social welfare benefits) (Department of Statistics, 1980). Presumably, a large part of the difference is due to the mothers in the CHDS *being mothers* (at home raising their one-year-old child, and possibly other children, hence having zero market income), whereas the total female population includes many non-parent females in the labour force. Second, sociodemographic variables collected in the CHDS may not have analogues in the 1976 census, or if they do, they may be coded/categorised differently. For example, the 1976 census did not collect data on highest educational qualification, rather it collected individuals' levels of education attended. While the latter can be ranked into 'highest education level attended' (and categorised as 'Primary school', 'Secondary school', and 'Tertiary institution' to aid comparison with the CHDS highest qualification variable), the comparison is inevitably distorted because, for example, many individuals who attended secondary school (classified in the 1976 census as 'Secondary school') did not achieve a secondary school qualification (classified in the CHDS as 'No formal qualification' rather than 'Secondary school qualification').

Similarly, for the offspring-related variables in Table 15, which were mostly collected in 2017, a comparison could be made with analogous statistics from the 2018 census. Yet the relevant statistics from the 2018 census are generally not reported specifically for the *New Zealand-born* population (the target population), and furthermore, the census represents all individuals present in New Zealand at the time of the census, whereas the CHDS offspring include cohort members who were living overseas in 2017 (and may have been for many years prior).[118] In short, comparison with published census statistics can provide only a crude assessment of sample representativeness and the aforementioned limitations of such a comparison should be borne in mind when assessing representativeness.

With that caution noted, Table 16 presents sociodemographic characteristics of the total population of Christchurch and of New Zealand drawn from published 1976 census statistics, alongside analogous characteristics of the parents in the CHDS 'all offspring' sample (dependent variable: offspring's family income over adulthood, this being the largest of all the samples). The

[118] For this reason, the CHDS offspring are not compared to the Christchurch total population in 2018, only to the New Zealand total population.

table shows that the parents in the sample come very close to accurately representing the proportion of single-parent families in the total population of New Zealand (6.4 percent in the sample compared to 7.5 percent in the New Zealand population), but that they (both mothers and fathers) are less likely to be in the youngest (15 to 19) or oldest (40 years and over) age groups than the female and male populations aged 15 to 49 years of both Christchurch and New Zealand. For example, the proportion of mothers in the sample aged 19 years and under in 1977 was nine percent, compared to about 20 percent in the total female population in 1976 in both Christchurch and New Zealand, a discrepancy which arises because, plainly, the sample has an age profile reflective of when females tend to have children (primarily in their twenties and thirties, less so in their teenage years or in their forties). The ethnic composition of the sample includes a higher proportion of Māori (five percent of mothers and eight percent of fathers) than in Christchurch as a whole (two percent of females and three percent of males), but the difference will be partly due to the CHDS researchers classifying mothers and fathers who report any Māori heritage as 'Māori', whereas the 1976 census classified respondents as 'Māori' if they met a 'half or more' proportion of Māori descent criterion. However, both the sample and the Christchurch total population under-represent Māori relative to the total New Zealand population (nine percent of whom were Māori in 1976, defined by the proportion of descent criterion mentioned), reflecting the relative ethnic homogeneity of Christchurch vis-à-vis the country as a whole.

**Table 16**. *Sociodemographic characteristics of Christchurch and New Zealand total populations in 1976, for comparison with CHDS parental characteristics.*

| 1977 CHDS variable[1] | Category | *Dependent variable*: Offspring's family income over adulthood |
|---|---|---|
| | | All offspring sample (%) |
| Offspring's family type at offspring's birth | Two-parent family | 93.6 |
| | Single-parent family | 6.4 |
| Mother's age at offspring's birth | 19 years and under | 9.2 |
| | 20-24 years | 30.7 |
| | 25-29 years | 40.4 |
| | 30-34 years | 15.3 |
| | 35-39 years | 3.5 |
| | 40 years and over | 1.0 |
| Father's age at offspring's birth | 19 years and under | 2.6 |
| | 20-24 years | 21.4 |
| | 25-29 years | 38.2 |
| | 30-34 years | 24.9 |
| | 35-39 years | 8.6 |
| | 40 years and over | 4.0 |
| | Missing | 0.2 |
| Mother's ethnicity at offspring's birth | New Zealand European and other[4] | 93.4 |
| | Māori[5] | 4.8 |
| | Pacific | 1.8 |
| Father's ethnicity at offspring's birth | New Zealand European and other[4] | 90.0 |
| | Māori[5] | 7.6 |
| | Pacific | 2.4 |
| | Missing | 0.1 |
| Mother's highest qualification at offspring's birth | No formal qualification | 49.8 |
| | Secondary school qualification | 30.0 |
| | Tertiary qualification | 20.2 |
| Father's highest qualification at offspring's birth | No formal qualification | 46.1 |
| | Secondary school qualification | 31.7 |
| | Tertiary qualification | 19.3 |
| | Missing | 2.8 |
| Father's annualised weekly market income at 1 year | Nil | 2.2 |
| | $1-$1,499 | 0.4 |
| | $1,500-$2,999 | 0.3 |
| | $3,000-$4,999 | 2.7 |
| | $5,000-$6,999 | 10.4 |
| | $7,000-$9,999 | 39.7 |
| | $10,000 and over | 29.7 |
| | No father | 7.9 |
| | Missing | 6.7 |

| 1976 census variable | Category | *1976 census* | |
|---|---|---|---|
| | | Christchurch total population[2] (%) | New Zealand total population (%) |
| Household type (one-family households only)[3] | Two-parent family | | 92.5 |
| | Single-parent family | | 7.5 |
| Female age group (15 to 49 years only) | 15-19 | 20.7 | 20.0 |
| | 20-24 | 18.4 | 17.3 |
| | 25-29 | 16.4 | 16.7 |
| | 30-34 | 12.4 | 13.4 |
| | 35-39 | 11.2 | 11.8 |
| | 40-44 | 9.9 | 10.1 |
| | 45-49 | 11.1 | 10.7 |
| Male age group (15 to 49 years only) | 15-19 | 20.9 | 20.1 |
| | 20-24 | 19.2 | 17.2 |
| | 25-29 | 16.4 | 16.5 |
| | 30-34 | 12.2 | 13.2 |
| | 35-39 | 10.6 | 11.8 |
| | 40-44 | 9.6 | 10.2 |
| | 45-49 | 11.1 | 11.0 |
| Female ethnic origin | New Zealand European and other[4] | 97.4 | 89.5 |
| | Māori | 2.0 | 8.6 |
| | Pacific | 0.6 | 1.9 |
| Male ethnic origin | New Zealand European and other[4] | 96.7 | 89.3 |
| | Māori | 2.6 | 8.7 |
| | Pacific | 0.7 | 2.0 |
| Female highest education level attended (15 years and over only) | Primary school[6] | 17.8 | 23.4 |
| | Secondary school | 65.8 | 63.4 |
| | Tertiary institution[7] | 16.4 | 13.2 |
| Male highest education level attended (15 years and over only) | Primary school[6] | 14.6 | 19.5 |
| | Secondary school | 53.8 | 55.5 |
| | Tertiary institution[7] | 31.6 | 25.0 |
| Male income group (15 years and over only) | Nil | 10.8 | 10.8 |
| | $1-$1,499 | 9.7 | 8.4 |
| | $1,500-$2,999 | 8.1 | 7.8 |
| | $3,000-$4,999 | 20.8 | 21.0 |
| | $5,000-$6,999 | 26.7 | 25.3 |
| | $7,000-$9,999 | 15.3 | 16.3 |
| | $10,000 and over | 7.9 | 9.0 |
| | Not specified | 0.7 | 1.5 |

**Notes**:
[1] The father's annualised weekly income variable was collected in 1978.
[2] This column refers to the total population in the Christchurch statistical division in 1976.
[3] A breakdown of one-family households into two-parent and single-parent families was not published for the 1976 Christchurch total population.
[4] 'Other' refers to ethnic origins that are not Māori, Pacific, or New Zealand European.
[5] The Māori category includes all parents who reported some Māori ancestry.
[6] This category includes 'Primary school', 'Nil', and 'Not specified'.
[7] This category includes 'University', 'Training College', and 'Technical Institute'.

**Sources**: Census statistics are from Department of Statistics (1977, 1979, 1980a, 1980c, 1980d, 1980e, 1981). CHDS statistics are the author's calculations.

As mentioned, the 1976 census collected data on highest education level attended (not highest qualification achieved), ranging from 'Nil/Primary school' up to 'University'. For comparison with the samples, these categories were collapsed down to 'Primary school', 'Secondary school', and 'Tertiary institution' (see notes to Table 16 for further detail) and proportions derived for the female and male populations aged 15 years and over in both Christchurch and New Zealand. The differences between the education measures used in the census and CHDS likely explain why roughly half of mothers and fathers had no formal qualification in 1977, yet over half of males and about three-fifths of females aged 15 years and over in the Christchurch and New Zealand populations had attended secondary school in 1976, and about one-fifth in the New Zealand population (15 to 18 percent in Christchurch) had not made it past primary school. Finally, when fathers' weekly gross market income in 1978 is annualised and categorised analogously to the 1976 census income group variable, the comparison shows that fathers in the sample have higher annual incomes, on average, than the total male populations of Christchurch and New Zealand, with considerably smaller proportions in the lowest three income groups. The differences are likely to be due, in large part, to CHDS fathers having an age profile reflective of parenthood which coincides with the prime working years.

Table 17 presents sociodemographic characteristics of the New Zealand total population aged 35 to 44 years drawn from 2018 census statistics alongside analogous characteristics of offspring in the same sample used in the parent comparison. As mentioned, the offspring are not compared to the Christchurch total population in 2018 because the sample includes a non-trivial proportion residing overseas (21 percent at age 40 for the sample under consideration). Compared to the total population, both daughters and sons in the sample are concentrated to a greater extent in the three middle categories of highest qualification and have smaller proportions with postgraduate degrees and no qualifications. For example, eight percent of daughters and six percent of sons have postgraduate degrees, compared to 18 percent of females and 14 percent of males in the total population. Part of the difference may be due to the presence of immigrants in the census (and their absence from the CHDS), as immigrants are (a) more highly qualified, on average, than native-born New Zealanders due to New Zealand's skills-skewed immigration settings (Alimi, Maré, & Poot, 2018a; Poot & Stillman, 2016) and (b) attracted to Auckland to a far greater extent than to Christchurch (Maré, Morten, & Stillman, 2007; Masgoret, Merwood, & Tausi, 2009). The distribution of daughters' and sons' occupations are similar to the female and male total populations, with the exception of two categories: both daughters and sons are more likely to be managers and less likely to be without paid employment than their respective populations. Once again, these differences may be partly attributable to immigrants, as the 2018 census shows that overseas-born individuals aged 35 to 44 were slightly more likely to be unemployed or not in the labour force than New Zealand-born individuals of the same age (20 percent versus 18 percent, respectively), and 2013 census data shows that overseas-born individuals were less likely to be managers than New Zealand-born individuals (17 percent versus 19 percent, respectively) (Stats NZ, 2021b). Remaining differences could be due to age differences between the sample (occupation at age 40) and the census (occupation for the population aged 35 to 44). Partnership status among the sample represents the population very well.

**Table 17**. Sociodemographic characteristics of New Zealand total population aged 35 to 44 years in 2018, for comparison with CHDS offspring characteristics.

| 2017 CHDS variable[1] | Category | *Dependent variable*: Offspring's family income over adulthood | 2018 census variable | Category | *2018 census* |
|---|---|---|---|---|---|
| | | All offspring sample (%) | | | New Zealand total population (%) |
| Daughter's highest qualification by 40 years[2] | No formal qualification | 5.5 | Female highest educational qualification (35 to 44 years only) | No formal qualification | 9.0 |
| | Secondary or tertiary level 4 qualification | 45.0 | | Secondary or tertiary level 4 qualification | 38.9 |
| | Tertiary level 5 or 6 qualification | 13.9 | | Tertiary level 5 or 6 qualification | 10.5 |
| | Bachelor degree | 27.7 | | Bachelor degree | 23.7 |
| | Postgraduate degree | 8.0 | | Postgraduate degree | 17.9 |
| Son's highest qualification by 40 years[2] | No formal qualification | 6.9 | Male highest educational qualification (35 to 44 years only) | No formal qualification | 12.5 |
| | Secondary or tertiary level 4 qualification | 56.2 | | Secondary or tertiary level 4 qualification | 46.5 |
| | Tertiary level 5 or 6 qualification | 9.3 | | Tertiary level 5 or 6 qualification | 10.3 |
| | Bachelor degree | 21.8 | | Bachelor degree | 16.9 |
| | Postgraduate degree | 5.8 | | Postgraduate degree | 13.8 |
| Daughter's main occupation at 40 years[2, 3] | Managers | 16.3 | Female occupation (35 to 44 years only) | Managers | 12.8 |
| | Professionals | 25.6 | | Professionals | 24.5 |
| | Technicians and trades workers | 5.3 | | Technicians and trades workers | 3.7 |
| | Community and personal service workers | 9.6 | | Community and personal service workers | 8.4 |
| | Clerical and administrative workers | 16.3 | | Clerical and administrative workers | 13.4 |
| | Sales workers | 3.9 | | Sales workers | 6.2 |
| | Machinery operators and drivers | 1.1 | | Machinery operators and drivers | 1.0 |
| | Labourers | 4.4 | | Labourers | 5.4 |
| | Not in paid employment | 17.6 | | Unemployed or not in the labour force[4] | 24.6 |
| Son's main occupation at 40 years[2, 3] | Managers | 28.3 | Male occupation (35 to 44 years only) | Managers | 21.6 |
| | Professionals | 22.0 | | Professionals | 20.1 |
| | Technicians and trades workers | 18.9 | | Technicians and trades workers | 16.6 |
| | Community and personal service workers | 5.3 | | Community and personal service workers | 4.1 |
| | Clerical and administrative workers | 1.5 | | Clerical and administrative workers | 4.0 |
| | Sales workers | 2.5 | | Sales workers | 4.3 |
| | Machinery operators and drivers | 7.6 | | Machinery operators and drivers | 7.5 |
| | Labourers | 5.8 | | Labourers | 8.7 |
| | Not in paid employment | 8.1 | | Unemployed or not in the labour force[4] | 13.3 |
| Daughter's partnership status at 40 years[5] | Partnered | 76.7 | Female partnership status (35 to 44 years only) | Partnered | 75.8 |
| | Not partnered | 23.3 | | Not partnered | 24.2 |
| Son's partnership status at 40 years[5] | Partnered | 77.8 | Male partnership status (35 to 44 years only) | Partnered | 78.9 |
| | Not partnered | 22.2 | | Not partnered | 21.1 |

**Notes**:
[1] Some variables are from 2012 (aged 35) or 2007 (age 30) if data from 2017 (age 40) was missing (occupation at age 40 is not one of these variables; see note 2 below).
[2] The distribution of offspring's occupation differs from that in Table 12 as the latter assigns offspring their occupation at age 35 or 30 if they have missing occupation data at age 40, whereas this table only uses occupation at age 40 to aid comparison with the census statistics (both distributions being cross-sectional snapshots).
[3] Occupations are coded to the major group level of the Australian and New Zealand Standard Classification of Occupations 2006.
[4] This category is taken from a separate census variable, 'Work and labour force status'.
[5] Missing responses to partnership status at age 40 have been excluded here to aid comparison with the census statistics.

**Source**: Census statistics are from Stats NZ (2021b). CHDS statistics are the author's calculations.

In summary, taking the parent and offspring comparisons together, the sample appears to do a reasonable job of representing the sociodemographic characteristics of the target population *once the first generation's status as parents, and the second generation's status as New Zealand-born, are taken into account*. Were it possible to make such an 'apples-with-apples' comparison, the sample would likely come even closer to faithfully representing the target population than indicated in Tables 16 and 17, as evidenced by some of the variables which are less affected by parenthood status (such as mother's and father's ethnicity) and native-born status (such as offspring's partnership status) where the proportions come close to exact representativeness. Because the sample appears to be reasonably representative, and because cohort members have been geographically mobile in adulthood, the results from the CHDS presented in this chapter are useful for understanding intergenerational income mobility among New Zealanders. The issue of representativeness is discussed further in section 4.7 addressing bias and limitations.

### *4.3.3 Descriptive statistics of offspring's earnings and family incomes over adulthood and childhood*

Table 18 presents, for each dependent variable, descriptive statistics of offspring's lifetime average earnings, offspring's family income over adulthood, and offspring's family income over childhood, by each of the three samples (all offspring, sons, daughters). All incomes in Table 18 are real weekly incomes in both September 2012 dollars and natural logarithms.

**Table 18**. *Descriptive statistics of CHDS offspring's lifetime average earnings and family incomes over childhood and adulthood.*

| Dependent variable | Sample | Income measure | n | Weekly income in 2012 Q3 NZD ($) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Minimum | | Median | | Maximum | | Mean | | SD | | Mean number of observations used in average |
| | | | | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | |
| Offspring's lifetime average earnings | All offspring | Offspring's lifetime average earnings[1] | 771 | 5 | 1.5 | 850 | 6.7 | 9,549 | 9.2 | 1,017 | 6.7 | 904 | 0.8 | 2.0 |
| | | Offspring's family income over childhood[2] | | 447 | 6.1 | 1,277 | 7.2 | 3,505 | 8.2 | 1,307 | 7.1 | 424 | 0.3 | 12.8 |
| | Sons | Son's lifetime average earnings | 413 | 5 | 1.5 | 938 | 6.8 | 9,471 | 9.2 | 1,155 | 6.8 | 917 | 0.7 | 2.8 |
| | | Son's family income over childhood | | 489 | 6.2 | 1,269 | 7.1 | 3,505 | 8.2 | 1,293 | 7.1 | 431 | 0.3 | 12.9 |
| | Daughters | Daughter's lifetime average earnings | 358 | 48 | 3.9 | 690 | 6.5 | 9,549 | 9.2 | 856 | 6.5 | 863 | 0.8 | 1.0 |
| | | Daughter's family income over childhood | | 447 | 6.1 | 1,298 | 7.2 | 2,601 | 7.9 | 1,324 | 7.1 | 416 | 0.3 | 12.8 |
| Offspring's family income over adulthood | All offspring | Offspring's family income over adulthood[3] | 926 | 115 | 4.7 | 1,406 | 7.2 | 20,014 | 9.9 | 1,623 | 7.2 | 1,289 | 0.6 | 2.8 |
| | | Offspring's family income over childhood | | 447 | 6.1 | 1,250 | 7.1 | 3,505 | 8.2 | 1,284 | 7.1 | 424 | 0.3 | 12.9 |
| | Sons | Son's family income over adulthood | 450 | 115 | 4.7 | 1,394 | 7.2 | 12,915 | 9.5 | 1,619 | 7.2 | 1,244 | 0.7 | 2.8 |
| | | Son's family income over childhood | | 489 | 6.2 | 1,249 | 7.1 | 3,505 | 8.2 | 1,277 | 7.1 | 429 | 0.3 | 12.9 |
| | Daughters | Daughter's family income over adulthood | 476 | 186 | 5.2 | 1,418 | 7.3 | 20,014 | 9.9 | 1,626 | 7.2 | 1,333 | 0.6 | 2.9 |
| | | Daughter's family income over childhood | | 447 | 6.1 | 1,250 | 7.1 | 2,601 | 7.9 | 1,290 | 7.1 | 419 | 0.3 | 12.8 |

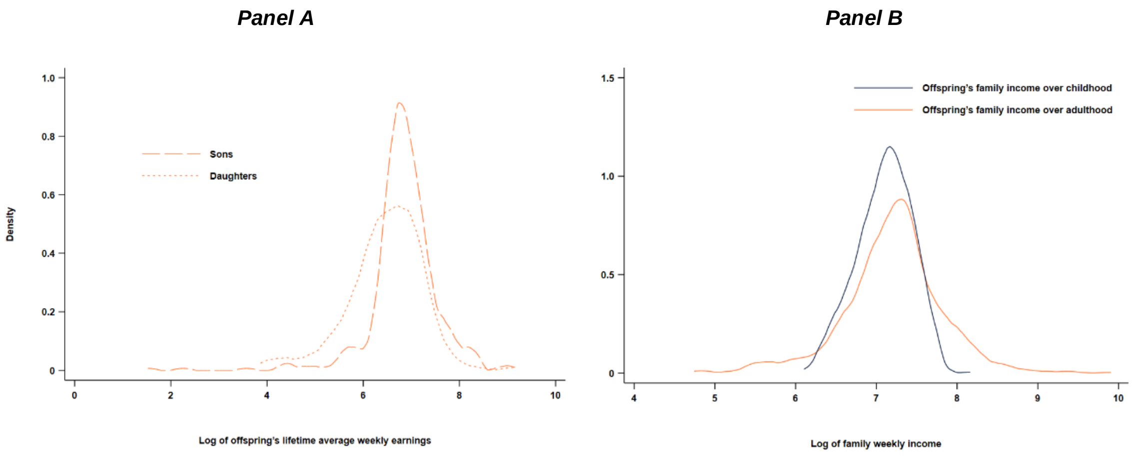
**Notes:**
[1] Offspring's weekly net earnings from all employment, averaged over ages 30, 35, and 40 years for sons where possible (centred on age 35, else the 35-year observation is used) and at age 40 years only for daughters.
[2] Combined weekly gross total income of mother and father, averaged over at least eight years of the offspring's childhood (ages 1 to 14 years).
[3] Combined weekly net total income of offspring and partner, averaged over ages 30, 35, and 40 years where possible, else a single year is used.

**Source**: Author's calculations from Christchurch Health and Development Study dataset.

Considering the 'all offspring' sample with offspring's family income over adulthood as the dependent variable, the CHDS cohort came from families with real weekly incomes over childhood spanning $447 up to $3,505, with a fairly symmetrical distribution between these extremes, indicated by the closeness of the median and mean family incomes ($1,250 and $1,284, respectively). On average, family income over childhood is averaged over 12.9 observations. As adults, sons and daughters from the cohort had an estimated mean lifetime average earnings of $1,155 and $856 per week, respectively. Most had formed their own partnerships/families with mean real family income of $1,623 per week, representing average growth in family incomes across generations of 26 percent. However, 42 percent of offspring had family incomes in adulthood that were, in real terms, lower than their family incomes in childhood, representing downward absolute mobility. The dispersion of family incomes increased substantially across generations, from a standard deviation of $424 in childhood to $1,289 in adulthood. Offspring's lifetime average earnings are averaged over two observations (comprised of an average of 2.8 for sons and 1.0 – the single observation at age 40 – for daughters), while offspring's family income over adulthood is averaged over 2.8 observations, on average.

Figure 31 displays kernel density plots showing the probability density function of log lifetime average earnings for sons and daughters (Panel A) and log family weekly income over childhood and adulthood for all offspring pooled (Panel B). Panel A shows that sons' log earnings distribution sits to the right of the daughters' distribution, reflecting higher average lifetime earnings among men compared to women. The sons' distribution is left-skewed whereas daughters' is symmetrical about its mean. This difference in the shapes of the distributions is partly due to the presence of time-averaging among sons (and its absence among daughters) and the dropping from the sample of offspring with proxy earnings of zero. For example, sons with low weekly earnings at age 30 and zero weekly earnings at ages 35 and 40 (e.g., due to unemployment) have their proxy earnings calculated as an average over all three years (including the two years with zero earnings, since all three observations are non-missing), resulting in a very low proxy of lifetime average earnings. This ‘time-averaging including years with zero earnings’ does not occur among daughters, who are simply assigned their weekly earnings at age 40 as a proxy for their lifetime average earnings (and if it is zero, they are dropped from the sample). This results in the sons’ distribution being pulled to the left: the lowest proxy earnings observed among daughters in the sample are the smallest positive weekly earnings at age 40, which are greater than the lowest proxy earnings among sons who are subject to ‘averaging with zeroes’.

**Figure 31**. *Kernel density plots of CHDS sons’ and daughters’ log lifetime average earnings and offspring’s log family incomes over childhood and adulthood.*




**Source**: Author’s calculations from Christchurch Health and Development Study dataset.

Panel B of Figure 31 shows the rightward shift (growth) in the distribution of family incomes across generations, however, incomes among the second generation are much more dispersed. This is likely due to a combination of greater smoothing (averaging) of parental incomes, top-coding of parental incomes, and increasing income inequality over the period of study (Perry, 2019), recalling that parental incomes are measured over 1978 to 1991 and offspring incomes over 2007, 2012, and 2017.[119]

## 4.4 Results

### *4.4.1 Estimates of the intergenerational income elasticity*

To estimate the IGE, ordinary least squares is applied to equation (1):

$$Y_{ij}^{\text{offspring}} = \alpha + \beta Y_i^{\text{parent}} + \varepsilon_i, \quad (1)$$

where $Y_{ij}^{\text{offspring}}$ is the natural logarithm of lifetime average earnings or family income over adulthood of offspring $i$ in family $j$, $Y_i^{\text{parent}}$ is log family income of the parents in family $j$ (over the

[119] Based on Perry (2019, p. 100), the Gini index of household income inequality in New Zealand (before housing costs) was 28.5 on average over the period 1982 to 1992 (measured biennially), and 32.9 on average over 2007, 2012, and 2017.

offspring's childhood), $\varepsilon_i$ is an error term capturing factors orthogonal to parental income, and $\beta$ is the intergenerational income elasticity. When analysing 'all offspring', offspring's sex is included as a covariate (1=male, 2= female) when the dependent variable is lifetime average earnings but excluded when the dependent variable is offspring's family income over adulthood (given that most offspring are partnered and nearly all partnerships are male-female couples). Controls for parental age are not included on the grounds that none of the models attempt to estimate parents' lifetime average income, the focus instead being on the family resources that were actually available to the offspring during childhood, irrespective of how old their parents were at the time. Controls for offspring's age are unnecessary given that the CHDS cohort are all born within a four-month window in the same year. The regression results are presented in Table 19.[120]

**Table 19**. *Estimates of the intergenerational income elasticity for offspring from the CHDS cohort, by each sample for each dependent variable.*

| | ***Dependent variable*: Offspring's lifetime average earnings** | | | ***Dependent variable*: Offspring's family income over adulthood** | | |
|---|---|---|---|---|---|---|
| | **Sample** | | | **Sample** | | |
| | **All offspring** | **Sons** | **Daughters** | **All offspring** | **Sons** | **Daughters** |
| Offspring's family income over childhood (IGE) | 0.479 *** (0.079) | 0.533 *** (0.100) | 0.412 ** (0.125) | 0.432 *** (0.058) | 0.449 *** (0.087) | 0.414 *** (0.077) |
| Sex | -0.388 *** (0.053) | | | | | |
| Constant | 3.818 *** (0.564) | 3.048 *** (0.711) | 3.526 *** (0.895) | 4.132 *** (0.411) | 3.999 *** (0.616) | 4.273 *** (0.548) |
| R-squared | 0.101 | 0.065 | 0.029 | 0.057 | 0.056 | 0.057 |
| Observations | 771 | 413 | 358 | 926 | 450 | 476 |

**Symbols:** ****p*<0.001 ***p*<0.01 **p*<0.05

**Notes:** Robust standard errors in parentheses. Sex is coded as 1=male, 2=female.

**Source**: Author's calculations from Christchurch Health and Development Study dataset.

The IGE of offspring's lifetime average earnings with respect to their family income over childhood is 0.479 (0.533 for sons, 0.412 for daughters). Conventionally interpreted, this means that a 10 percent increase in real family income over childhood is associated with, on average, a 4.79 percent increase in offspring's earnings averaged over a lifetime. Put another way, if one family has twice the income of another, then the offspring of the higher-income family will, as an adult, earn roughly 48 percent more income over their lifetime than the offspring of the lower-income family (53 percent more if they are a son, 41 percent more if they are a daughter, indicating stronger intergenerational income persistence among sons than daughters). Recall that if family income had no influence on offspring's earnings, the IGE would be zero (or close to zero). Offspring's sex is significantly associated with lifetime average earnings, with daughters having lower log earnings than sons, holding log childhood family income constant. All IGE estimates are statistically significant at the one percent level. The IGE point estimates and their 95 percent confidence intervals are presented diagrammatically (together with robustness checks) in section 4.4.2.

[120] Standard errors are not clustered at the family level because there are only eight sets of twins in the 'all offspring' sample with offspring's lifetime average earnings as the dependent variable, and nine sets of twins in the 'all offspring' sample with offspring's family income over adulthood as the dependent variable. When standard errors *are* clustered, the IGE estimates are, in both cases, the same as the estimates without clustering to seven decimal places.

The earnings elasticity is 29 percent higher among sons than among daughters. Why would childhood family income be more strongly linked to the earnings of sons than daughters? As Mitnik et al. (2015) conjecture, this may be because the earnings distribution for women is less dispersed than that for men, while the parental income distribution is essentially the same for women and men, which would have the mechanical effect of lowering the IGE for women relative to men, all else equal. However, in the CHDS samples, the standard deviation of lifetime average earnings is only marginally smaller for daughters than for sons, while the standard deviation of parental incomes (family income over childhood) is similar between the two (see Table 18 in section 4.3.3), so this 'mechanical' effect is not likely to play a significant role in the gender difference in earnings elasticities found here.

To understand gender differences in intergenerational income persistence, Mitnik et al. (2015) outline an 'accounting framework' in which income can be secured either via one's own earnings in the labour market or via a spouse in the marriage market. It follows that intergenerational income persistence will occur whenever higher parental income is associated with: (1) increases in offspring's own earnings, or; (2) increases in the chances of marrying, thereby increasing offspring's family income via income-sharing with their spouse, or; (3) conditional on being married, increases in the chances of marrying a higher-earning spouse.[121] Mitnik et al. (2015) draw on the literature to demonstrate that these channels affect men and women differently which in turn produce gender differences in intergenerational income persistence. With respect to the first 'direct' channel in which higher parental income increases offspring's earnings in the labour market, this will differentially affect men and women if women (to a greater extent than men) are, for a variety of reasons, impeded from realising their earnings potential. In particular, if women from high-income backgrounds with high earnings capacity are not able to fully realise this capacity in the labour market, then any link between their parental income advantage and their own earnings advantage is weakened, thereby reducing the intergenerational association between parental income and daughters' earnings. One mechanism that may prevent women from fully realising their earnings potential is occupational sex segregation, wherein "[b]y virtue of socialization, discrimination, and other social and economic forces, women tend to work in a constrained set of occupations, and their earnings potential may not, as a result, be as reliably realized as men's" (Mitnik et al., 2015, pp. 61-62). Another potential mechanism is unequal division of labour within households, whereby a disproportionate share of unpaid domestic work is borne by women which may induce them to opt for part-time or low-paying jobs that do not fully realise their earnings potential.[122] These sex-based differences reduce the ability of women from high-income backgrounds to capitalise on their advantages in the labour market, and we should therefore expect weaker intergenerational persistence with respect to daughters' own earnings than for son's own earnings.

[121] Mitnik et al. (2015) consider the first channel to be 'direct' and the second and third channels to be 'indirect' because the latter are spouse-mediated. As they put it, their framework "distinguishes the "direct pathway" (via own earnings) and the "indirect pathway" (via marriage) through which inequality in [family] total income is transmitted across generations" (Mitnik et al., 2015, p. 38). Choi, Chung and Breen (2020) distinguish the same three channels in their model of intergenerational persistence of family income in the US.

[122] Other mechanisms include "gender-specific effects of parental earnings on skill formation" and "wage discrimination by gender" (Raaum et al., 2007, p. 12).

Mitnik and colleagues' second channel via the marriage market (in which higher parental income increases the chances of marrying) will differentially affect men and women if the chances of marrying are more strongly related to parental income among women than among men (or vice versa, but they find the former holds). If women from high-income backgrounds are more likely to marry than men from high-income backgrounds, then a greater proportion of women secure increases to their *family income* (via income-sharing between spouses, referred to by Mitnik et al. (2015, p. 64) as the "marriage-probability payoff") while a greater proportion of men remain single, forgoing gains from spousal income. Therefore, when the focus is on persistence of *family income* across generations, we should expect stronger intergenerational persistence among daughters' family income than among sons' family income, provided the link between marriage chances and parental income is stronger among daughters than sons.

Drawing on the work of Raaum et al. (2007), Mitnik and colleagues' third channel (via the marriage market, in which higher parental income increases the chances of marrying a higher-earning spouse) will differentially affect men and women if women from high-income backgrounds with higher earnings potential marry higher-earning partners (that is, marital sorting or assortative mating) and as a consequence choose to work fewer hours or not at all – especially where they have young children – thereby lowering their own earnings (that is, family labour supply responses) (Raaum et al., 2007). The joint effect of assortative mating and the negative response of female labour supply to spousal income leads to marriages in which women from high-income backgrounds with high earnings potential do not fully realise this potential. By contrast, "married men are expected to be breadwinners (although this normative presumption is likely weakening), which means that their labor supply is largely unresponsive to their spouses' earnings potential" (Mitnik et al., 2015, p. 61).[123] Therefore, when the focus is on the persistence of *earnings* across generations, we should expect weaker intergenerational persistence among daughters than among sons as marital sorting and labour supply responses affect the realised earnings potential of the former to a greater extent than the latter. But when the focus is on the persistence of *family income*, we should not necessarily expect marked gender differences and, to the extent there are any, their size and direction will depend on the combination of the three channels which may have offsetting effects.

Mitnik et al. use this framework to separate out the contributions that each of the channels makes to the IGE of offspring's expected family income, thereby illuminating gender differences in the intergenerational transmission of income: if the IGE of offspring's expected family income is similar for women and men, then the framework can show "whether that similarity nonetheless conceals differences in the way in which it is generated", whereas if the IGE differs between women and men, the framework can "specify whether that difference is attributable to gender differences in the direct or indirect forms of transmission" (Mitnik et al., 2015, p. 39). Using US data on offspring's earnings, their spouses' earnings, and offspring's marriage probability, Mitnik et al. (2015, pp. 71, 63) find that parental income matters more for men's earnings than for women's earnings: "Although both men and women secure an earnings payoff from being raised

[123] Raaum et al. (2007, p. 8) remark that "several studies suggest that the labor supply of married men is quite inelastic and affected by neither the own nor the partner's wage".

in higher-income families, men have a much higher payoff than do women", with the payoff being lower for women because "when they are married, they tend to reduce their supply of labor as their spouses' earnings increase (while married men do not)", suggesting that "assortative mating and the negative income elasticity of labor supply are indeed generating a lower earnings [IGE] for women than for men". But with respect to family income, Mitnik et al. (2015, p. 67) find that the IGE "takes on a similar size for women and for men", yet "the processes through which it is generated differ markedly across genders", with the direct pathway (via own earnings) accounting for 61 percent of the family income IGE for sons, whereas the indirect pathway (via marriage and earnings conditional on marriage) accounted for 71 percent of the IGE for daughters.

In the New Zealand context, Alimi, Maré, and Poot (2018b) use census microdata to estimate the degree of educational assortative mating – partnering of people with similar education levels – among male-female couples working full-time in New Zealand and find that it increased over the period 1986 to 2013[124] and accounted for about 20 percent of income inequality observed among the couples they study. With respect to labour supply, Mercante and Mok (2014a, p. 28) find that for partnered women in New Zealand in the period 2006 to 2011, "having a university-educated husband has a negative effect on the wife's preferences for work", an effect which is "slightly larger than the effect of the wife's own education on preferences for work", hence "the more educated the husband, the greater the negative impact on the wife's preferences for working". Kalb and Scutella (2003, p. 18) similarly find that in New Zealand "if the partner's education level is higher, then a woman's preference for work is reduced to some extent". These findings for New Zealand imply that the earnings payoff from having high-income parents will be lower among daughters than sons in the CHDS – indeed, lower among women than men in New Zealand more generally – which would explain the lower earnings IGE found here among daughters.

The IGE of offspring's family income over adulthood with respect to their family income over childhood is 0.432 (0.449 among sons, 0.414 among daughters). This means that a 10 percent increase in family income over childhood is associated with, on average, a 4.32 percent increase in offspring's family income over adulthood.[125] If one family has twice the income of another, then the offspring of the higher-income family will, as an adult, have a family income that is roughly 43 percent higher than that of the offspring of the lower-income family (45 percent more if they are a son, 41 percent more if they are a daughter, indicating similar degrees of intergenerational income persistence among sons and daughters). All IGE estimates are statistically significant at the one percent level. The IGE point estimates and their 95 percent confidence intervals are presented diagrammatically (together with robustness checks) in section 4.4.2.

[124] However, this increase was driven entirely by matching among people with below-degree-level educational qualifications – assortative mating decreased among individuals with at least a Bachelor degree and among those with no formal qualifications. Alimi et al. (2018b) also find that occupational assortative mating – the partnering of people with similar occupations – decreased over the period of study.

[125] As previously noted, Mitnik and Grusky (2020) argue that the correct interpretation of these results is as follows: a 10 percent increase in family income over childhood is associated with a 4.79 percent increase in *the geometric mean* of offspring's lifetime average earnings, and a 4.32 percent increase *in the geometric mean* of offspring's family income in adulthood.

In contrast to the earnings IGE, the family income IGE is fairly similar between sons and daughters, differing only by eight percent. This is consistent with results from Raaum et al. (2007) who find that the same shift – switching the measure of offspring income from own earnings to combined earnings of offspring and spouse – results in substantial reductions in gender differences in IGE estimates across five countries.[126] It is also consistent with results from Chadwick and Solon (2002) and Mitnik et al. (2015). Why is the gender difference observed in the earnings elasticities reduced substantially in the elasticities of family income? Mitnik et al. (2015) look to their second channel for answers, using their US data to empirically test whether daughters are more likely to marry than sons as parental income increases. Mitnik et al. find a very strong relationship between parental income and marriage chances among both sons and daughters, but the relationship is stronger among daughters, leading them to conclude that a "higher "marriage-probability payoff" explains, in part, why the [family income] elasticity for women is nearly as large as that for men, even though their earnings elasticity is so much smaller" (Mitnik et al., 2015, p. 65). In other words, daughters' higher marriage chances compensate for their smaller earnings IGE, lifting daughters' IGE of family income closer to that of sons. Thus, the similarity in family income IGEs observed between sons and daughters in the CHDS cohort may be because the latter obtain a higher 'marriage-probability payoff'.

Comparing the elasticities from the two dependent variables, the IGE for lifetime average earnings among all offspring pooled is 11 percent higher than the IGE for family income, consistent with analogous results from Rohenkohl (2019). However, the difference is driven entirely by sons, for whom the earnings IGE is 19 percent higher than the family income IGE, whereas for daughters the two IGEs are very similar (in fact, daughters' family income IGE is marginally higher than their earnings IGE). Consistent with this latter result, Deutscher and Mazumder (2020) find that switching between earnings and family income barely changes the IGE among daughters (from 0.166 for earnings to 0.181 for family income, a nine percent increase). Inconsistent with these results, Mitnik et al. (2015) find that the IGE of sons' family income is 23 percent higher than the IGE of sons' earnings (a reversal of the pattern found here), Deutscher and Mazumder (2020) find no change in the IGE between the two measures of sons' income (as they found with daughters), and Chadwick and Solon (2002) find little change in the IGE between the two measures of sons' income.

Thus, the empirical literature gives mixed results on the effect of changing the dependent variable on IGE estimates. What should we expect in theory? Recall from the discussion in chapter 2 that the IGE increases with the standard deviation of offspring's log income and decreases with the standard deviation of parents' log income. Helsø (2021) argues that this feature means that when offspring's income is defined as their *individual* income and parents' income is defined as *family* income (pooled across parents), there are more close-to-zero income observations among the former than among the latter, which increases the dispersion of offspring's income relative to parents' income, which in turn mechanically increases the IGE. This would explain why changing

[126] For example, in Finland the IGE of own earnings is 44 percent higher among sons compared to daughters, but the difference halves to 22 percent when the dependent variable is the combined earnings of offspring and their spouse (among samples pooling single and married offspring). The reduction is even larger in Norway, Great Britain, and the US (Raaum et al., 2007).

the dependent variable to offspring's *family* income (pooled across offspring and partner) – thus eliminating most close-to-zero incomes – has the effect of reducing IGE estimates. The fact that this mechanical effect is observed among sons but not daughters in the CHDS cohort may be because among the latter the effect is offset by some of the gender-based mechanisms affecting women's earnings described above, such as assortative mating and the negative cross-wage elasticity of labour supply, that move in the opposite direction (increasing daughters' family income IGE relative to their earnings IGE).

*4.4.2 Robustness checks of the intergenerational income elasticity estimates*

This section reports checks of the robustness of the IGE estimates to alternative specifications, first for dependent variable 1 and then for dependent variable 2. Figure 32 displays the IGE point estimate and 95 percent confidence interval from the 'all offspring' model with offspring's lifetime average earnings as the dependent variable ("Model 1", the left-most estimate), followed by three alternative specifications:

Model 2. Family income over childhood includes in its derivation social welfare benefits which have been scaled by an (arbitrary) factor of 1.2, in order to convert them from after-tax to an approximate before-tax value (see Figure 27 and footnote 114 in section 4.2.1). To test the sensitivity of the IGE estimate to this treatment of parental benefit income, Model 2 excludes social welfare benefits from family income over childhood. The resulting measure – family income excluding transfers – is comparable to that used in Mazumder (2016) and Chetty et al. (2014). This specification drops from the sample some offspring whose family income over childhood derived only from benefits.

Model 3. Weekly family incomes of zero at each year of childhood,[127] and offspring weekly earnings of zero at ages 30, 35, and 40, are all coded as $1, and then offspring's proxy for lifetime average earnings and their family income averaged over childhood are both re-derived. This specification increases the sample size to n=865.

Model 4. Mitnik's alternative estimator – the IGE of *expected* income – is estimated using Poisson pseudo-maximum likelihood (PPML) estimation. As in model 2, the sample size increases to n=865.

These checks are repeated for the samples of sons and daughters and displayed in Figure 33, with one additional check performed only for sons:

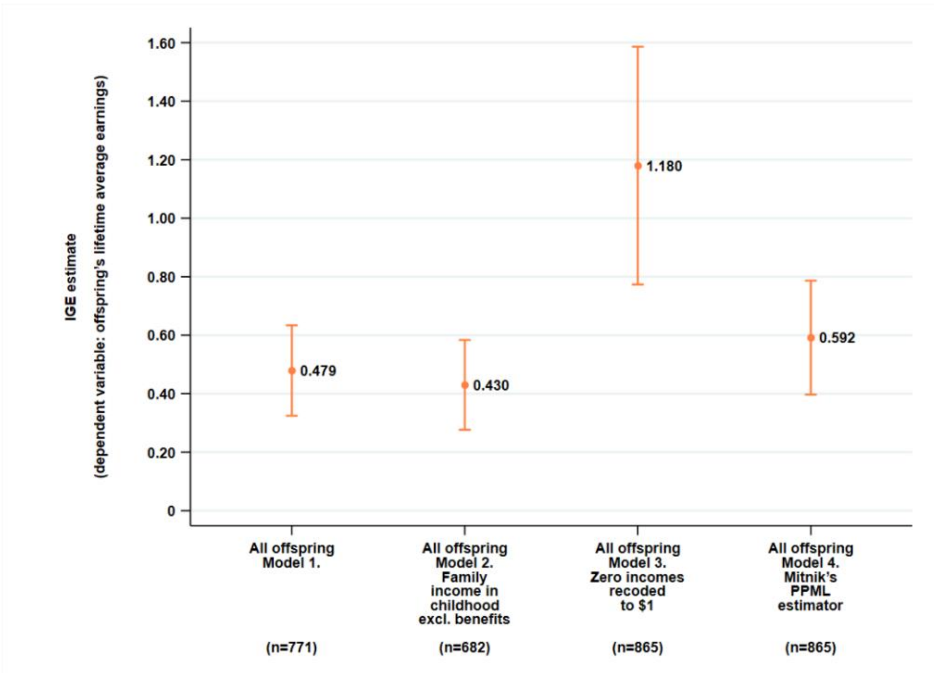
Model 5. Sons' proxy earnings are measured at age 40 (rather than averaged around age 35, meaning no time-averaging of earnings is used in this specification, as with daughters' proxy earnings). Age 40 is a common age used to proxy men's incomes in empirical studies (e.g., Aaronson & Mazumder, 2008; Mazumder,

[127] At some specific ages of childhood (1 to 14 years), there are a very small number of cohort members (never more than four for dependent variable 1) who had weekly family incomes of zero. These families were usually either dependent on someone else for financial support or were in some form of transition (for example, living off savings between jobs) at that particular age. However, no cohort members have an *average* family income over childhood of zero.

2016) and falls within empirically-estimated age windows for minimising life-cycle bias (see Table 1 in section 2.3.5).

Figure 32 shows that, when social welfare benefits are excluded from family income over childhood, the IGE decreases by 10 percent to 0.430, but the confidence interval overlaps considerably with that for model 1. When incomes of zero (on both sides of the equation) are coded as $1, the IGE increases by 146 percent to 1.180, indicating that the IGE is highly sensitive to the treatment of zero incomes.[128] This result is driven by the addition to the sample of 94 cohort members with proxy earnings of zero (which are coded as $1). Of these 94 cohort members, almost all (98 percent) were not in paid employment at the age their proxy earnings were measured, and 82 percent were female. In model 4 (Mitnik's PPML estimator), despite the sample size and composition being the same as for model 3 (incorporating those with proxy earnings of zero), the PPML estimator reaches a very different estimate of the IGE of 0.592 (owing to its treatment of zero incomes in a very different way to the standard log-log specification). This is an increase in the IGE of 24 percent compared to model 1.

**Figure 32**. *Robustness checks of IGE estimate (dependent variable: offspring's lifetime average earnings) for sample of all offspring.*



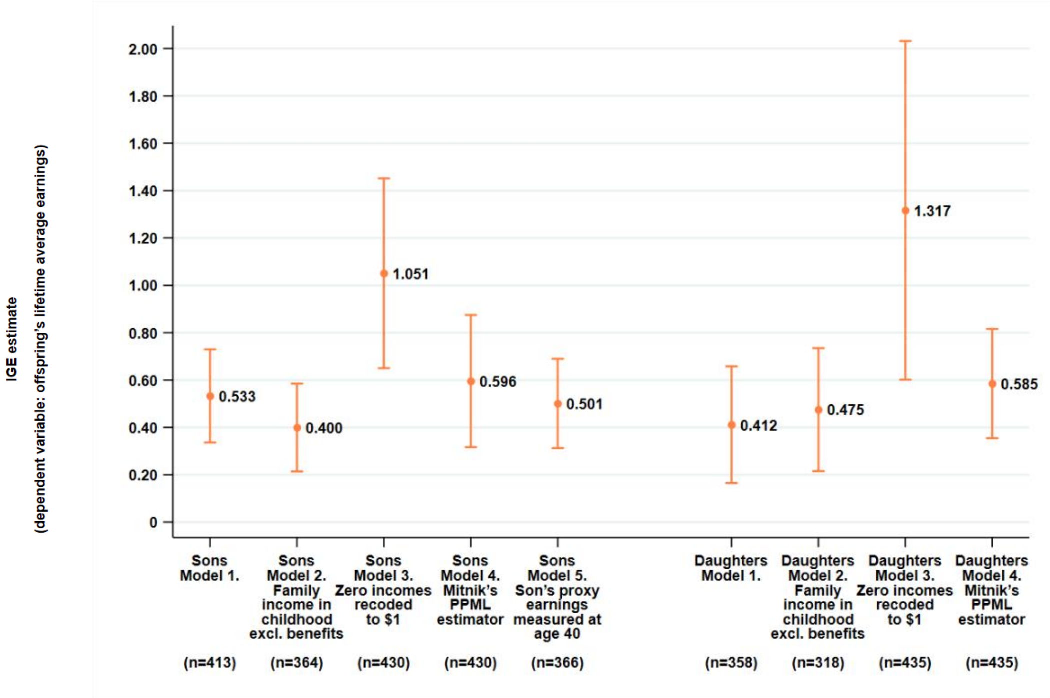

**Source**: Author's calculations from Christchurch Health and Development Study dataset.

Continuing with the analysis of dependent variable 1, Figure 33 displays the IGE point estimates and 95 percent confidence intervals for the samples of sons and daughters, along with results from robustness checks. For the most part, the same patterns are observed among each sex: coding zero incomes as $1 substantially increases the IGE estimates and Mitnik's PPML estimator

[128] Mitnik et al. (2015) use tax and other administrative data to estimate the IGE of total family income for son-parent and daughter-parent pairs in the US. When those with missing tax and administrative income data are excluded, their estimate of the conventional IGE is 0.43 for sons and 0.36 for daughters. When they impute missing incomes at $1, their estimates rise to 1.09 for sons (a 140 percent increase) and to 1.03 for daughters (a 186 percent increase). When Chetty et al. (2014) and Landersø and Heckman (2017) run similar sensitivity checks – imputing zero incomes with $1 – their estimates of the IGE (sons and daughters pooled) increase by 80 percent and 70 percent, respectively.

increases the estimates for both sexes but especially for daughters, where the IGE increases by 42 percent to 0.585. Excluding benefits from family income in childhood decreases the IGE among sons but increases it by 15 percent among daughters. Finally, proxying sons' lifetime average earnings at age 40 rather than age 35 has little effect on the IGE, decreasing it by six percent.

**Figure 33***. Robustness checks of IGE estimates (dependent variable: offspring's lifetime average earnings) for samples of sons and daughters.*




**Source**: Author's calculations from Christchurch Health and Development Study dataset.

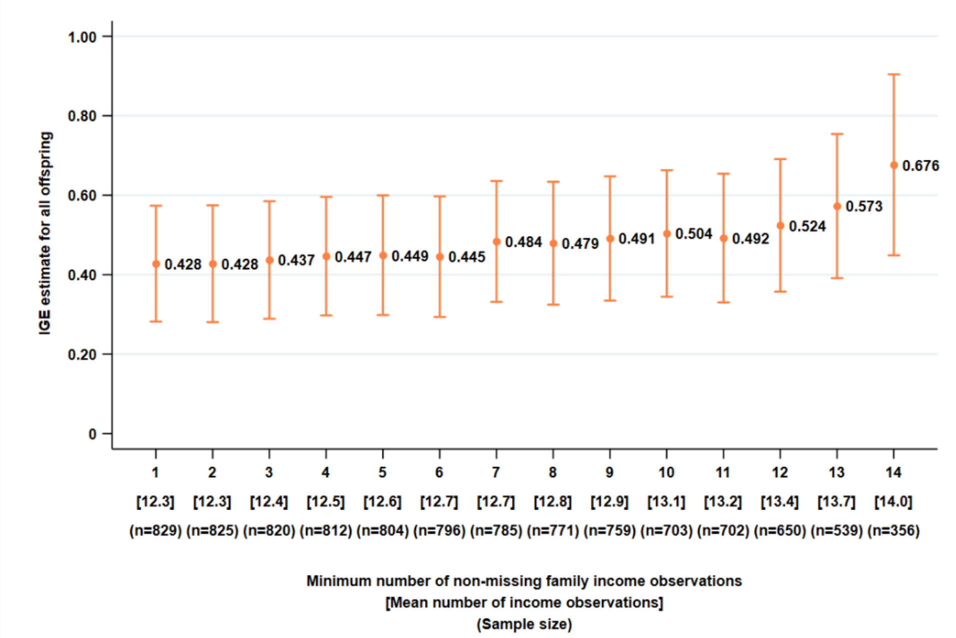
A final robustness check is performed testing the sensitivity of the IGE estimate to different minima of non-missing family income observations over childhood. Recall that family income over childhood is calculated by averaging over at least eight valid observations of family income across ages 1 to 14 years. Figure 34 displays, for the 'all offspring' sample,[129] the IGE estimate by different minima of valid family income observations, ranging from a minimum of one (that is, only one family income observation is required to be included in the sample) to a minimum of 14 (that is, family income must be observed at each of the 14 years of childhood). With a minimum of one, the sample size is 829 and, owing to relatively high rates of item response across childhood, the average number of valid family income observations is 12.3. With a minimum of 14, the sample size declines to 356.

A 'three-step' pattern can be discerned, in which the IGE is fairly stable across minima of one to six, increases to around 0.480-0.500 for minima of seven to 11, and increases again over each of the remaining three minima (hence the 'third step' is more of an upward-sloping curve), with a particularly large increase at minimum 14 to 0.676. This pattern is consistent with evidence that IGE estimates tend to increase in the number of parental income observations used on the right-

[129] The same pattern of IGE estimates by different minima of family income observations is found among the samples of sons only and daughters only (not shown here).

hand side (Mazumder, 2005; Solon, 1992). However, it is likely that sample composition also plays a part in the increasing IGE across minima. For example, it is likely that those families in the CHDS who were the most compliant – agreeing to take part at all 14 waves across childhood and reporting their family income at every one of those waves – exhibit the highest degree of intergenerational income persistence.

**Figure 34**. *Check of sensitivity of IGE estimate (dependent variable: offspring's lifetime average earnings) to different minima of non-missing family income observations over childhood of CHDS cohort.*




**Source**: Author's calculations from Christchurch Health and Development Study dataset.

Turning to dependent variable 2 (offspring's family income over adulthood), Figure 35 displays the IGE point estimates and 95 percent confidence intervals for each sample ("All offspring Model 1", "Sons Model 1", "Daughters Model 1"), each followed by two alternative specifications:

Model 2. Weekly family incomes of zero at each year of childhood (ages 1 to 14 years), and at each year of adulthood (ages 30, 35, and 40), are all coded as $1, and then offspring's family income averaged over childhood and adulthood are re-derived. The sample remains the same.

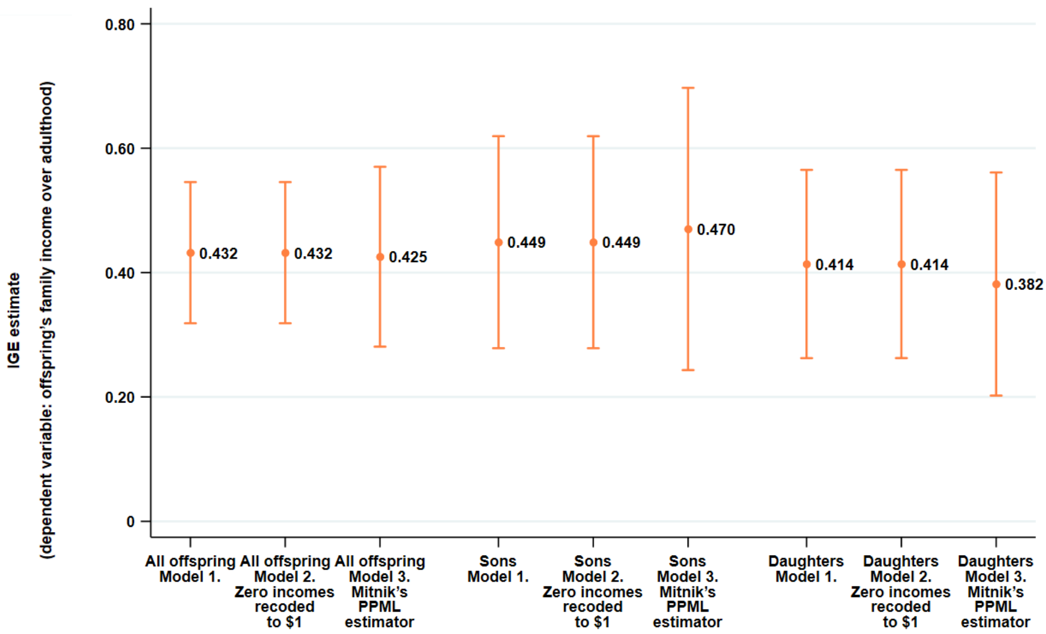
Model 3. Mitnik's alternative estimator – the IGE of *expected* income – is estimated using PPML estimation. The sample remains the same.

Figure 35 shows that, in all samples, coding zero incomes as $1 does not change the IGE estimates. Note that there are very few offspring with family incomes of zero in any given year of childhood or adulthood.[130] Mitnik's PPML estimator increases the IGE among sons by five percent

[130] For dependent variable 2, never more than five offspring have family incomes of zero across ages 1 to 14 years, and at ages 30, 35, and 40, the number of offspring with family incomes of zero is 14, nine, and five, respectively. See footnote 127 in section 4.4.2 for why, at any given data collection wave, some families report no weekly income.

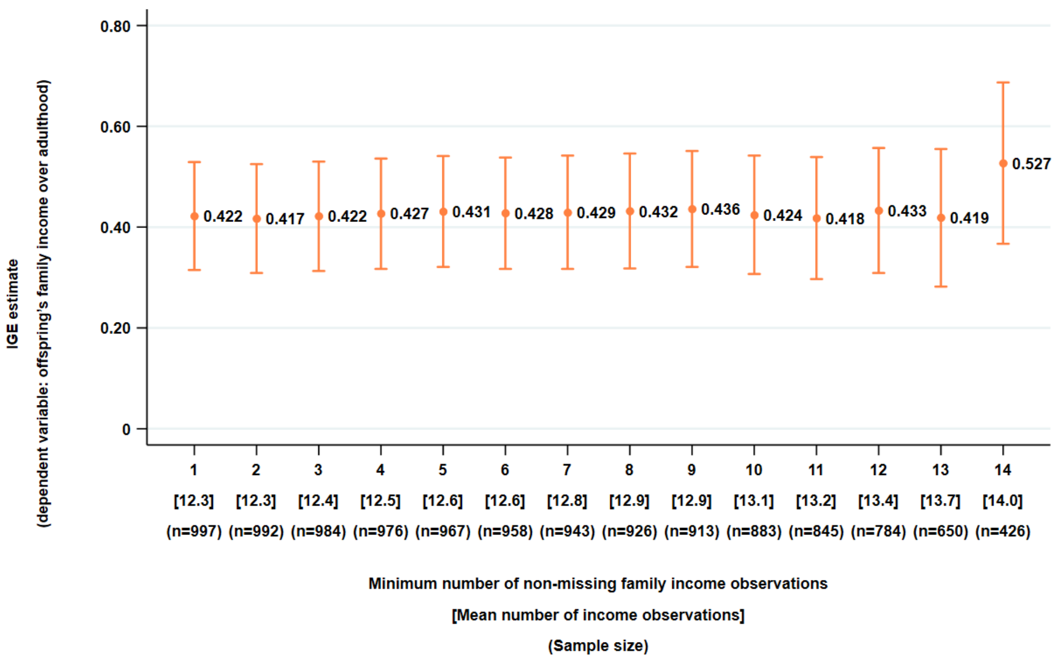
(to 0.470) and decreases it among daughters by eight percent (to 0.382), with the net effect among all offspring being a negligible decrease.

**Figure 35**. *Robustness checks of IGE estimate (dependent variable: offspring's family income over adulthood) for samples of all offspring, sons, and daughters.*




**Source**: Author's calculations from Christchurch Health and Development Study dataset.

**Figure 36**. *Check of sensitivity of IGE estimate (dependent variable: offspring's family income over adulthood) to different minima of non-missing family income observations over childhood of CHDS cohort.*



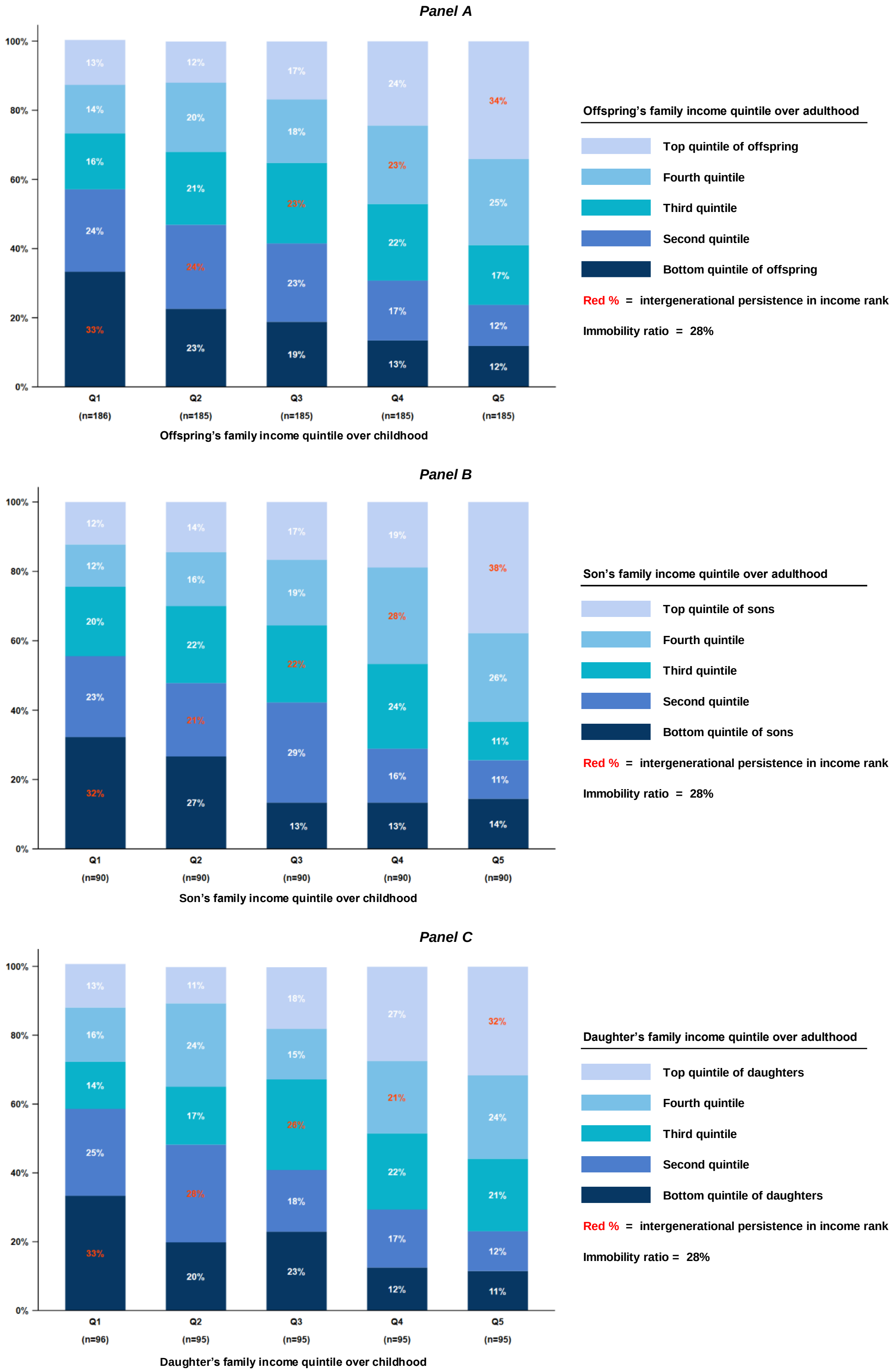

**Source**: Author's calculations from Christchurch Health and Development Study dataset.

As with dependent variable 1, a further robustness check is performed testing the sensitivity of the IGE estimate to different minima of non-missing family income observations over childhood. Figure 36 displays, for the 'all offspring' sample, the IGE estimate by different minima of valid family income observations over childhood. With a minimum of one, the sample size is 997 and the average number of valid family income observations is 12.3. With a minimum of 14, the sample size declines to 426. The IGE is fairly stable across minima of one to thirteen. Only at a minimum of 14 is there a noticeable change in the IGE, increasing to 0.527, a 22 percent increase compared to the estimate from model 1 (minimum of eight). If the estimates were changing in response to increased smoothing of income over childhood, we would expect to see a gradual increase over the range of minima. Instead, a sudden shift is observed at minimum 14, implying a sample composition effect on the IGE estimate, namely that intergenerational income persistence is higher among the (more compliant) subset of cohort families who participated in every annual data collection wave from ages 1 to 14 and were willing to disclose their income at every one of those timepoints.

### *4.4.3 Transition matrices of intergenerational persistence in income rank*

The analysis now shifts to an examination of intergenerational persistence in income rank, sometimes referred to as intergenerational 'rank mobility'. This analysis uses offspring's family income over adulthood as the dependent variable. In order to estimate the correlation between family income rank in each generation, the distributions of offspring's family income over childhood and offspring's family income over adulthood are each divided into quintiles (however, a robustness check in section 4.4.5 uses the more-conventional percentiles). The two quintile-ranked income variables are then cross-tabulated, producing a transition matrix with transition probabilities of intergenerational mobility. Figure 37 displays transition matrices in graph form, for all offspring (Panel A), sons (Panel B), and daughters (Panel C). For each quintile of family income during childhood along the bottom of the graph (quintile of origin), a bar displays the distribution of offspring across quintiles of family income in adulthood (quintile of destination). For example, looking at Panel A (all offspring), of the offspring from the poorest families in childhood (Q1), 33 percent remained in the bottom quintile of family income by adulthood, 24 percent had risen to quintile 2, 16 percent had risen to quintile 3, 14 percent had risen to quintile 4, and 13 percent had risen to the top quintile. Percentages highlighted in red represent intergenerational persistence in family income rank. An immobility ratio can be computed for each transition matrix, which is simply the percentage of offspring who remain in their parents' quantile of income (the proportion along the leading diagonal). If there is complete immobility, the immobility ratio is 100 percent. If there is 'origin independence' (equal probability of ending up in each destination quintile, for each origin quintile), the immobility ratio is 20 percent. As reported in Figure 37, the immobility ratio is 28 percent for all samples, meaning 72 percent of offspring changed quintile (either upward or downward) across generations.

**Figure 37**. *Graphed transition matrices of offspring's family income rank over childhood by family income rank over adulthood, for samples of all offspring, sons, and daughters.*




**Source**: Author's calculations from Christchurch Health and Development Study dataset.

This degree of immobility is similar to the degree of *socioeconomic* immobility found among the Dunedin Study birth cohort. Belsky et al. (2018) use data from the Dunedin Study to estimate intergenerational socioeconomic persistence by comparing the socioeconomic ranking of 831 cohort members' occupations at age 38 with the socioeconomic ranking of their parents' occupations over the cohort members' childhoods.[131] These socioeconomic scores are then divided into quintiles and cross-tabulated into a transition matrix. They find rates of intergenerational socioeconomic persistence of 41 percent for parental quintile 1 (that is, roughly four-in-ten cohort members whose parents occupied the lowest socioeconomic group were in that same group themselves in adulthood), 24 percent for parental quintile 2, 31 percent for parental quintile 3, 23 percent for parental quintile 4, and 30 percent for parental quintile 5. These figures imply an overall level of persistence of about 30 percent, that is, just under one-third of cohort members at age 38 where in the same socioeconomic group as their parents, meaning about 70 percent had changed socioeconomic quintile.

Yet despite the relatively high level of mobility overall, some intergenerational persistence can be observed in Figure 37, especially at the top and bottom of the income distribution. Among offspring from the highest-income families, about 60 percent remain in the top two quintiles in adulthood. A similar proportion of offspring from the lowest-income families (57 percent) remain in the bottom two quintiles in adulthood. This is consistent with other studies that find greater 'stickiness' at the top and bottom of the income distribution when analysed with transition matrices (e.g., Dearden, Machin, & Reed, 1997; Hirvonen, 2008; Jäntti et al., 2006). Long-range mobility – upward movement from the bottom to the top quintile, or downward movement in the opposite direction – is relatively uncommon. Only 13 percent of offspring from the lowest-income families experience a 'rags to riches' rise to the top quintile, which compares with eight percent in the US (Chetty et al., 2014), 11 percent in Canada (Corak, 2020a), and 12 percent in Australia (Deutscher & Mazumder, 2020).[132] Thus, long-range upward mobility in New Zealand appears to be relatively high compared to other countries.[133] With respect to 'riches to rags' downward mobility, 12 percent of offspring from the highest-income families drop to the bottom quintile, which is similar to an estimated 11 percent in the US (Chetty et al., 2014) but lower than an estimated 14 percent in Australia (Deutscher & Mazumder, 2020). The pattern of transitions is similar between sons and daughters.

[131] Occupations are assigned a socioeconomic score based on the *New Zealand Socio-economic Index 2006* for cohort members and Elley and Irving's (1976) *Revised Socioeconomic Index* for parents, where the highest occupational status of either parent is averaged over eight time-points spanning the cohort member's childhood from birth to age 15. Both indices score occupations based on income and education levels of persons with that occupation in the New Zealand census.

[132] All studies cited use quintile transition matrices of family or household income, as in the current analysis. Jäntti et al. (2006) estimate rags to riches mobility based on quintiles of *individual earnings* between son-father and daughter-father pairs to be between 11 and 14 percent for sons and between 14 and 16 percent for daughters in Denmark, Finland, Norway, Sweden, and Great Britain, with only the US standing out as having relatively little long-distance upward mobility at 8 percent for sons and 10 percent for daughters.

[133] Long-range upward *educational* mobility – tertiary-educated offspring who came from families with lower-secondary-educated parents – is higher in New Zealand than in all 27 other OECD countries analysed in OECD (2018).

*4.4.4 Estimates of the intergenerational rank correlation*

To estimate the rank correlation, ordinary least squares is applied to equation (2):

$$R_{ij}^{\text{offspring}} = \alpha + \rho R_{i}^{\text{parent}} + \varepsilon_i, \qquad (2)$$

where $R_{ij}^{\text{offspring}}$ is the rank in the distribution of lifetime average earnings or family income over adulthood of offspring $i$ in family $j$, $R_{i}^{\text{parent}}$ is the rank in the distribution of family income of the parents in family $j$ (over the offspring's childhood), $\varepsilon_i$ is an error term capturing factors orthogonal to parental rank, and $\rho$ – Spearman's rho – is the intergenerational rank correlation (also known as the rank-rank slope). As with the IGE for this dependent variable, controls for offspring's sex, offspring's age, and parents' age are not included in the model. The regression results are presented in Table 20.

**Table 20**. *Estimates of the intergenerational rank correlation for offspring from the CHDS cohort.*

| | ***Dependent variable*: Offspring's family income quintile over adulthood** | | |
|---|---|---|---|
| | **Sample** | | |
| | **All offspring** | **Sons** | **Daughters** |
| Offspring's family income quintile over childhood (rank correlation) | 0.269 *** | 0.278 *** | 0.264 *** |
| | (0.032) | (0.045) | (0.044) |
| Constant | 2.193 *** | 2.167 *** | 2.204 *** |
| | (0.105) | (0.151) | (0.147) |
| R-squared | 0.072 | 0.077 | 0.070 |
| Observations | 926 | 450 | 476 |

**Symbols:** ****p*<0.001 ***p*<0.01 **p*<0.05

**Notes:** Robust standard errors in parentheses.

**Source**: Author's calculations from Christchurch Health and Development Study dataset.

The intergenerational rank correlation is 0.269 (0.278 for sons, 0.264 for daughters). This means moving up one quintile in the parental income distribution is associated with, on average, an increase in offspring's income rank of 0.269 of a quintile, or just over one-quarter of a quintile (0.278 of a quintile if they are a son, 0.264 of a quintile if they are a daughter, indicating marginally stronger intergenerational rank persistence among sons than daughters). If parents' family income had no influence on offspring's family income, the rank correlation would be (at or near) zero. All rank correlation estimates are statistically significant at the one percent level. The rank correlation point estimates and their 95 percent confidence intervals are presented diagrammatically (together with robustness checks) in the next section.

The closeness of the estimates for sons and daughters is consistent with very small differences in rank correlations observed between sons and daughters in Denmark (Helsø, 2021), Australia (Deutscher & Mazumder, 2020), and the US (Chetty et al., 2014) and the same gendered channels of marriage-probability payoffs and assortative mating discussed in relation to the IGE results may also apply here.

*4.4.5 Robustness checks of intergenerational rank correlation estimates*

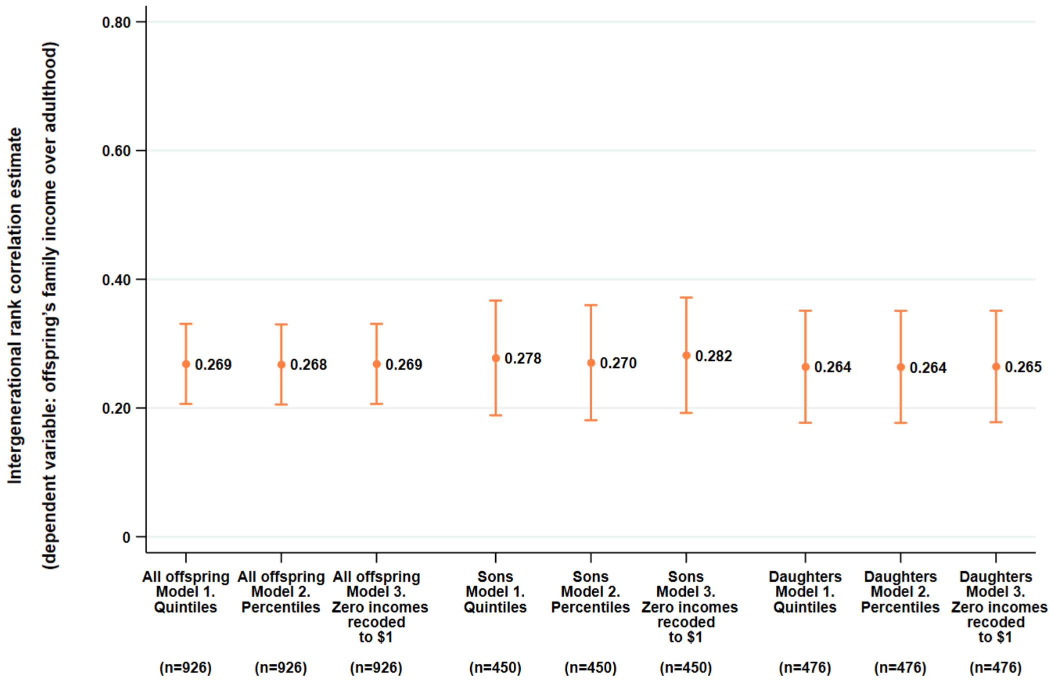
This section reports checks of the robustness of the rank correlation estimates to alternative specifications. Figure 38 displays the rank correlation point estimate and 95 percent confidence intervals for each sample (“All offspring Model 1”, “Sons Model 1”, “Daughters Model 1”), each followed by two alternative specifications:

Model 2. Family incomes are divided into the more-conventional percentiles. In the all offspring sample there are only nine or 10 offspring in each percentile, while in the sons and daughters samples there are only four or five in each. The sample remains the same.

Model 3. Weekly family incomes of zero at each year of childhood (ages 1 to 14 years), and at each year of adulthood (ages 30, 35, and 40), are all coded as $1, and then offspring’s family income averaged over childhood and adulthood are re-derived. The sample remains the same.

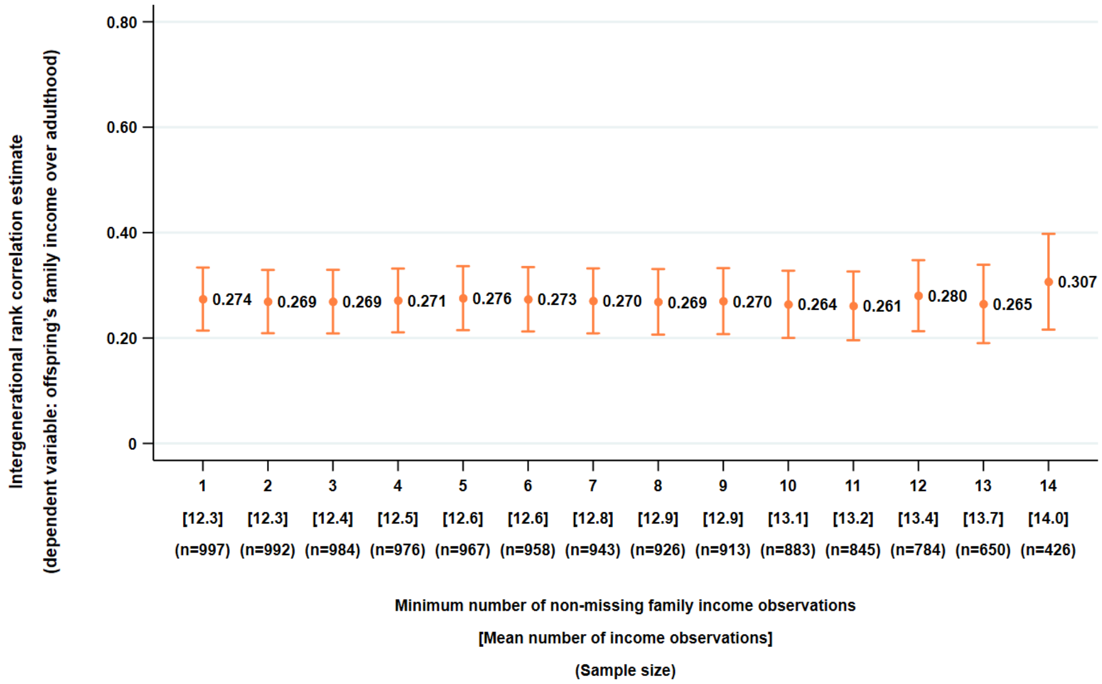
Figure 38 shows that, in all samples, using percentile rank instead of quintile rank does not materially alter the rank correlation estimates. Coding zero incomes as $1 also has a negligible effect on the estimates, an expected result given that coding the lowest incomes does not change ranks in the distribution.

**Figure 38**. *Robustness checks of intergenerational rank correlation between family incomes in childhood and adulthood for samples of all offspring, sons, and daughters.*




**Source**: Author’s calculations from Christchurch Health and Development Study dataset.

**Figure 39**. *Check of sensitivity of intergenerational rank correlation estimate to different minima of non-missing family income observations over childhood of CHDS cohort.*




**Source**: Author's calculations from Christchurch Health and Development Study dataset.

A further robustness check is performed testing the sensitivity of the rank correlation estimate to different minima of non-missing family income observations over childhood. Figure 39 displays, for the 'all offspring' sample, the rank correlation estimate by different minima of valid family income observations. With a minimum of one, the sample size is 997 and the average number of valid family income observations is 12.3. With a minimum of 14, the sample size declines to 426. As with the IGE of offspring's family income over adulthood, the rank correlation is fairly stable across minima of one to thirteen, and only at a minimum of 14 is there a noticeable change, increasing to 0.307, a 14 percent increase compared to the estimate from model 1 (minimum of eight).[134] Again, this is likely due to a sample composition effect, namely that intergenerational persistence in income rank is higher among the more compliant subset of cohort families who participated in (and disclosed their income at) every annual data collection wave from ages 1 to 14.

### 4.5 Cross-national comparison

In this section, the CHDS estimates of the IGE and rank correlation are compared to estimates from the two other existing New Zealand studies and from studies of other countries. Estimates from comparator studies have been chosen to be as analogous as possible by selecting estimates that are (a) for the same offspring-parent pair (i.e., CHDS estimates for sons and daughters are only compared to son-specific and daughter-specific estimates from other studies, and CHDS estimates for all offspring are only compared to estimates that also pool sons and daughters), and; (b) for a similar income concept for both offspring and parents (e.g., CHDS estimates for

[134] The same pattern is observed among the samples of sons and daughters (not shown here).

offspring's earnings are only compared to estimates from other studies that also use offspring's individual earnings, or failing that, individual total income, and similarly for parents), and; (c) where more than one study uses a comparable offspring-parent pair and comparable income concept, choosing the study that uses offspring birth cohorts and/or ages at which offspring income is observed that are closest to the CHDS birth cohort year (1977) and ages of income observation (in the case of earnings, 35 for sons and 40 for daughters).[135] As previously discussed in section 3.5, even similar income concepts are defined slightly differently in each country/study considered here and there are always a range of differences in income definitions, sample selection criteria, time periods covered, and methodology between any two or more studies. Consequently, precise rankings of countries are fraught with difficulty and should be interpreted with caution.

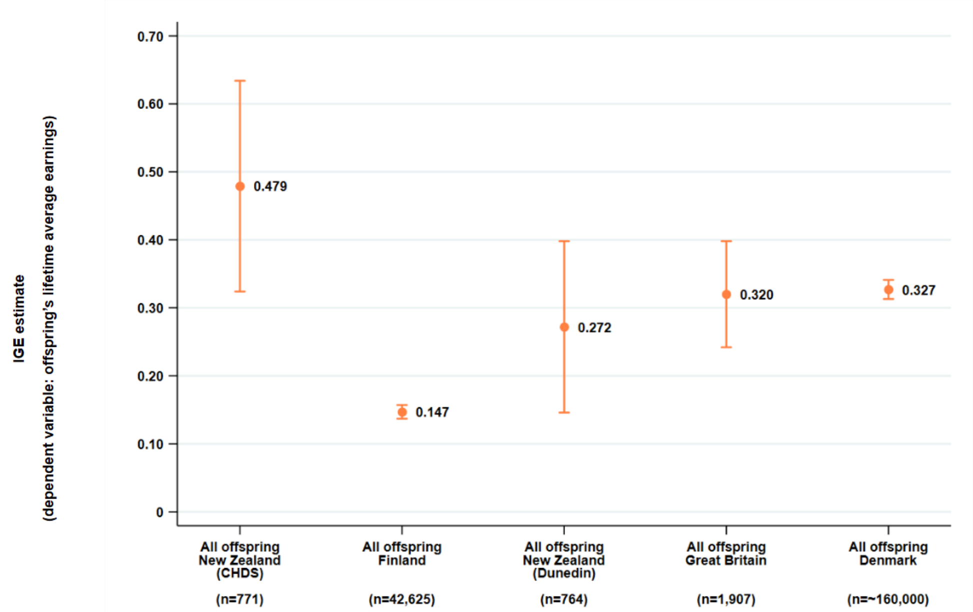
In Figures 40 to 47 below, the IGE and rank correlation estimates are plotted alongside comparable estimates from studies of New Zealand, Finland, Denmark, Norway, Sweden, Australia, Germany, Great Britain, Canada, and the US (for each dependent variable and each sample). Included beneath these figures are tables providing further details of the studies from which the estimates are taken, including the type of data source (whether the data was from an administrative or survey source), the income concept used for offspring (broadly defined), the particular birth cohorts studied (years that offspring were born), and the ages at which offspring's income is observed. For some countries, estimates plotted in one figure may come from a different source to estimates in another figure for the same country. For example, for Great Britain, Rohenkohl (2019) estimates the IGE of offspring's earnings with respect to parents' family income for all offspring (sons and daughters pooled) but does not report separate estimates by sex, so a British son-specific estimate from Gregg, Macmillan, et al. (2017) is used for comparison with the CHDS son-specific estimate.

Österbacka (2001) uses data on 42,625 offspring drawn from Finland's 1970 census who were born between 1950 and 1960 and whose annual earnings are observed at the 1985, 1990, and 1995 censuses spanning ages 25 to 45 and estimates the IGE of offspring's log individual earnings with respect to their parents' log family earnings to be 0.147. Gibbons' (2010) study uses data on 764 offspring drawn from the longitudinal Dunedin Study born in 1972-1973 whose annual incomes are observed at age 32 and estimates the IGE of offspring's log individual total income with respect to their parents' log family income to be 0.272. Rohenkohl (2019) uses data on 1,907 offspring from the British Household Panel Survey (BHPS)[136] born between 1973 and 1991 whose incomes are observed over windows of varying lengths across ages 25 to 43 (average of 6.1 observations, observed at an average age of 30 years) and estimates the IGE of offspring's log

[135] As noted in section 3.5, comparisons are restricted to estimates from OLS estimation, as estimates from other approaches such as TSIV are upward-biased both in theory and in practice (see footnote 82 in section 3.5), thus making estimates from the two approaches not directly comparable.

[136] The BHPS was subsequently incorporated into the UK Household Longitudinal Study (or 'Understanding Society'). As part of BHPS wave 18, participants were invited to join the new, larger, and more wide-ranging Understanding Society survey, which began in 2009 as a household panel conducting annual interviews with a nationally representative sample of approximately 100,000 individuals from 40,000 households from across the United Kingdom. The first interviews with BHPS participants in Understanding Society were carried out in wave 2 in 2010-2011. Rohenkohl (2019) makes use of Understanding Society data up to wave 7 (2015-2016) which has been integrated and harmonised with the earlier BHPS data back to 1991.

**Figure 40**. *Comparison with other studies of intergenerational income elasticity estimate (dependent variable: offspring's lifetime average earnings) for sample of all offspring.*




| *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|
| NZ (CHDS) | Survey | | Individual labour earnings | 1977 | 35 (sons), 40 (daughters) |
| Finland | Administrative | Österbacka (2001, p. 474) | Individual labour earnings | 1950-1960 | 25-45 |
| NZ (Dunedin) | Survey | Gibbons (2010, p. 20) | Individual total income | 1972-1973 | 32 |
| Great Britain | Survey | Rohenkohl (2019, p. 31) | Individual labour earnings | 1973-1991 | 30 (average) |
| Denmark | Administrative | Helsø (2021, p. 528) | Individual labour earnings | 1973-1975 | 35-39 |

individual gross monthly labour earnings with respect to their parents' log household gross monthly income to be 0.320. Helsø (2021) uses data on approximately 160,000 Danish offspring drawn from the full population administrative register born between 1973 and 1975 whose incomes are observed over three-year windows across ages 35 to 39 and estimates the IGE of log individual 'wage earnings' with respect to their parents' log family income to be 0.327.[137] Figure 40 plots these IGE estimates from the Dunedin Study, Great Britain and Denmark alongside the CHDS 'all offspring' estimate. On these estimates, intergenerational persistence between the earnings of offspring (of both sexes pooled) and their parents' incomes is highest in the New Zealand city of Christchurch, lowest in Finland, with Dunedin, Great Britain, and Denmark in between. The latter two countries display similar degrees of persistence to each other; however, all confidence intervals are mutually overlapping with the exception of Finland which only overlaps marginally with the lower limit of the Dunedin point estimate.

The large disparity between the estimates of intergenerational persistence for Christchurch and Dunedin – both cohorts born in the 1970s in cities in the South Island of New Zealand with similar demographic make-up – is likely to be due to differences in data rather than genuine differences in intergenerational mobility. First, the difference in the number of parental income observations (two available in the Dunedin Study, 14 available in the CHDS) is likely to have increased the

[137] Helsø (2021) includes self-employment income in her definition of individual earnings, which she shows has the effect of substantially increasing the IGE compared to estimates based on a definition of earnings that does not include self-employment income (only wages and salaries), as in Landersø and Heckman (2017), who estimate the IGE of offspring's wage earnings (excluding self-employment income) with respect to family income in Denmark to be 0.083. Helsø (2021, p. 527) argues that "[t]he wage-earnings IGE coefficients are particularly low [in Landersø and Heckman (2017)] because wage earnings do not include self-employment income…the wage-earnings IGEs increase by a factor larger than 3 once self-employment income is counted as earnings".

CHDS estimate relative to the Dunedin estimate, as IGE estimates are known to increase in the number of income observations over which parental income is averaged.[138] Second, the difference in the age at which offspring incomes are observed (age 32 in Gibbons' study, around age 35 for sons and 40 for daughters in the current analysis) may have increased the CHDS estimate relative to the Dunedin, since IGE estimates tend to increase with the age at which offspring's income is observed, at least among sons (Bratberg et al., 2005; Chen et al., 2017; Grawe, 2006; Gregg, Macmillan, et al., 2017; Nybom & Stuhler, 2017).[139] Third, the difference in the number of offspring earnings observations (one used in Gibbons' study, an average of two used in the current analysis) may increase the CHDS estimate relative to the Dunedin estimate, as Nybom and Stuhler (2016) show a tendency for IGE estimates to be higher the greater the number of observations over which offspring's income is averaged.[140] Finally, the age of offspring when their parents' incomes were observed (ages 13 and 15 in Dunedin, ages 1 to 14 in the CHDS, hence a fuller coverage of childhood in the latter) may have increased the CHDS estimate relative to the Dunedin estimate if Muller's (2010) finding that parental income measured in early childhood results in higher IGE estimates than when measured in the teenage years (among father-son pairs in the US) is applicable to these New Zealand data. Using CHDS data from birth to age 21 to analyse the effects of cohort members' family income during three different stages of childhood (ages 1 to 5, 6 to 10, and 11 to 14) on having no educational qualifications by age 21 (and four other detrimental life outcomes), Maloney (2004) finds that family income between ages 1 to 5 is significantly associated with educational failure and has the strongest effects of all three stages of childhood, a link that holds (albeit with reduced effect size) even when controlling for a range of potential confounders including parents' educational qualifications (family income between ages 6 to 10 and 11 to 14 are no longer associated with having no qualifications with the addition of the confounders).[141] Maloney's (2004) results suggest family income *in early childhood* is particularly important for children's educational attainment and consequently may be more strongly associated with children's incomes in adulthood than parents' family income in middle or late childhood.[142]

[138] Gibbons' (2010, p. 26) estimate for all offspring increases to 0.337 when he excludes six offspring-parent pairs with extremely low incomes in either generation. This higher estimate may be more comparable to the CHDS estimate if it better reflects the IGE estimate that would have been obtained had more extensive time-averaging of Dunedin parental and offspring incomes, akin to that in the current analysis, been possible (i.e., the low incomes of these six cases may have been 'pulled up' with greater smoothing).

[139] Bratberg et al. (2005) find a less clear-cut pattern among daughters in Norway (when paired to their fathers); if anything, the earnings IGE estimates among daughters tend to decrease after age 30 and then fluctuate slightly through to age 45.

[140] Nybom and Stuhler (2016) analyse Swedish sons whose annual total incomes, averaged over up to seven years centred on ages 30 to 40, are regressed on their fathers' near-lifetime average total incomes (thus, left-side measurement error is present while right-side measurement error is virtually absent). As discussed in section 2.3.4 of chapter 2, the reason why the IGE increases in the number of offspring income observations used appears to be a reduction in lifecycle bias that arises because multiyear averages on the left-hand side "counteract the disproportionate influence of occasional low-income episodes" (Nybom & Stuhler, 2016, p. 266).

[141] The link between family income from ages 1 to 5 and educational failure continues to remain statistically significant at the 10 percent level when two mediating variables – child IQ and conduct problems – are controlled for in addition to the confounders, again with reduced effect size.

[142] An additional difference is the income concept (offspring's individual total income in the Dunedin Study, offspring's individual earnings in the CHDS). Empirical studies that estimate the IGE of both income concepts with the same sample have generated mixed results: Björklund et al. (2012), Munk et al. (2016), Landersø and Heckman (2017), and Chen et al. (2017) find the elasticity of total income to be higher than that for earnings, whereas Murray, Clark, Mendolia and Siminski (2018), Rohenkohl (2019), and some specifications in Helsø (2021) find the opposite result.

Turning to son-specific and daughter-specific estimates, Bratsberg et al. (2007) use administrative data on 27,364 Norwegian sons drawn from the full-population administrative register born in 1958 whose earnings are observed at ages 34 and 41, 18,706 Danish sons drawn from the full-population administrative register born in 1958 whose earnings are observed at ages 40 and 42, and 9,937 Finnish sons drawn from quinquennial census panel data born between 1956 and 1960 whose earnings are observed at ages 38 to 42 and 41 to 45. Sons in each country are matched to their parents' family income, computed as the sum of father's and mother's earnings. The IGE of sons' log earnings with respect to parents' log family income is estimated to be 0.138 in Norway, 0.155 in Denmark, and 0.231 in Finland.[143] Björklund et al. (2017) use administrative data on Swedish sons and daughters drawn from the population register born in 1969 whose earnings are observed at ages 30, 34, and 38 and survey data on sons and daughters drawn from the British Cohort Study born in 1970 whose earnings are also observed at ages 30, 34, and 38, and estimate the IGE of offspring's log earnings with respect to their parents' log family income to be 0.181 among sons in Sweden, 0.133 among daughters in Sweden, 0.271 among sons in Great Britain, and 0.329 among daughters in Great Britain. Österbacka (2001), described above, estimates the IGE of daughters' log annual earnings with respect to their parents' log family earnings to be 0.132 among 20,278 Finnish daughters. Deutscher and Mazumder (2020) use administrative data on 1,025,800 Australian sons and daughters drawn from federal income tax records born in Australia between 1978 and 1982 whose individual total incomes and family total incomes are observed over five-year windows across ages 29 and 37 (average 33 years) and estimate the IGE of offspring's log individual total income with respect to their parents' log family income to be 0.181 for sons and 0.166 for daughters. Mitnik et al. (2015) use data on 6,552 sons drawn from a US administrative panel, the Statistics of Income Mobility (SOI-M) Panel,[144] who were born between 1972 and 1975 and whose individual labour earnings are observed in population tax (and other administrative) data at ages 35 to 38, and estimate the IGE of sons' log earnings with respect to their parents' log family income to be 0.350.[145] Choi et al. (2020) use survey data on 651 daughters drawn from the PSID born between 1963 and 1975 whose individual labour earnings are observed over ages 35 to 42, and estimate the IGE of daughters' log earnings with respect to their parents' log family income to be 0.399. Gregg, Macmillan, et al. (2017) use survey data on 4,312 sons drawn from the longitudinal British Cohort Study born in 1970 whose monthly earnings are observed at ages 26, 30, 34, 38, and 42[146] and estimate the IGE of sons' log earnings with respect to their parents' log family income to be

[143] These are Bratsberg and colleagues' (2007) IGE estimates from linear regressions. However, the main aim of their paper is to estimate IGEs using *nonlinear* specifications that add higher-order polynomial terms to the regression equations and estimate the IGE at the 10th, 50th, and 90th percentiles. In this way, they show that the functional form of the relationship between sons' and fathers' log earnings is nonlinear in the Nordic countries (but much closer to being linear in the two other countries they study, Great Britain and the US).

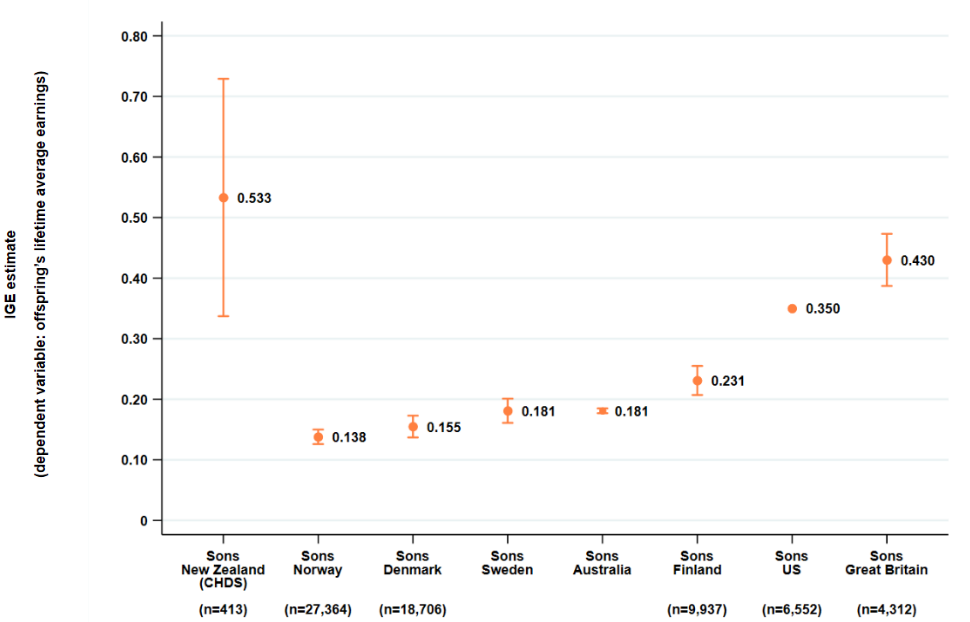
[144] Offspring in the SOI-M Panel are dependents identified in their parents' 1987 tax returns and living in the US in 1987, supplemented with a sample of dependents born in the same years whose parents did not file a tax return in 1987 (due to their income falling below the filing threshold) but who did file in at least one of the subsequent years to 1996. The person or persons who claimed the child as a dependent in their 1987 (or subsequent) returns are defined as the child's parents.

[145] Mitnik and colleagues' (2015) analysis is focused on estimating their alternative estimand – the IGE of offspring's expected income – but they also estimate the conventional IGE, and it is the latter results that are reported here for comparison with the CHDS estimates.

[146] Zero earnings due to spells out of work are replaced by the average social welfare benefit level available at the time of the workless spell (replaced whether benefits were actually claimed or not) and other missing earnings are imputed.

0.430.[147] Using data on 1,528 daughters drawn from the same British Cohort Study who were employed and had positive income at age 30, Blanden, Goodman, Gregg and Machin (2004) estimate the IGE of daughters' log earnings at age 30 with respect to their parents' log family income to be 0.400.[148] Figures 41 and 42 plots these IGE estimates alongside the CHDS estimates for sons (Figure 41) and daughters (Figure 42).

**Figure 41**. *Comparison with other studies of intergenerational income elasticity estimate (dependent variable: offspring's lifetime average earnings) for sample of sons.*



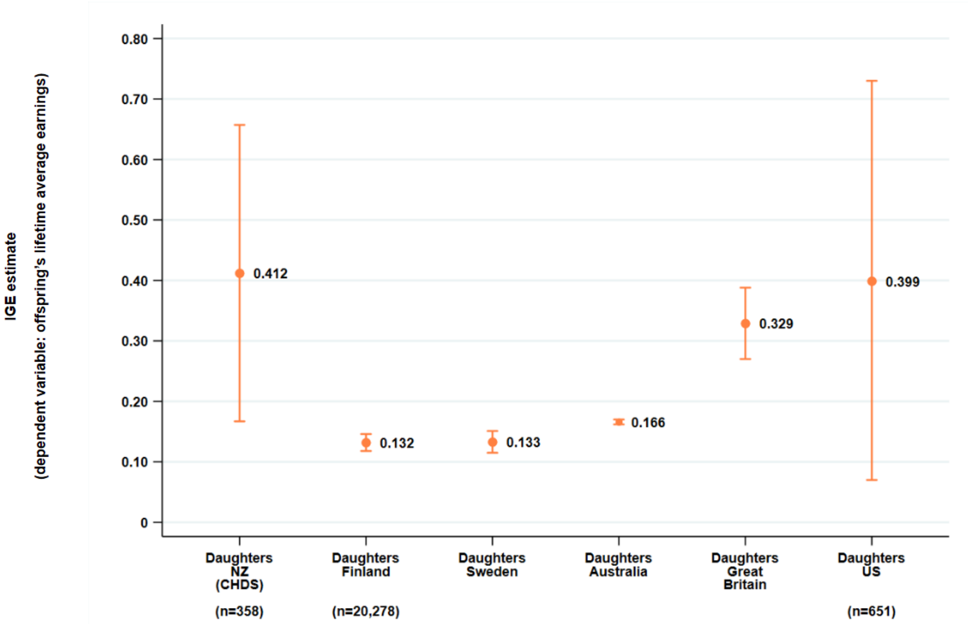

| *Sex* | *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|---|
| Sons | NZ (CHDS) | Survey | | Individual labour earnings | 1977 | 35 (centred) |
| Sons | Norway | Administrative | Bratsberg et al. (2007, p. C85) | Individual labour earnings | 1958 | 34 and 41 |
| Sons | Denmark | Administrative | Bratsberg et al. (2007, p. C85) | Individual labour earnings | 1958 | 41 (average) |
| Sons | Sweden | Administrative | Björklund et al. (2017, appendix p. 1) | Individual labour earnings | 1969 | 30, 34, and 38 |
| Sons | Australia | Administrative | Deutscher & Mazumder (2020) | Individual total income | 1978-1982 | 29-37 |
| Sons | Finland | Survey | Bratsberg et al. (2007, p. C85) | Individual labour earnings | 1956-1960 | 38-42 and 41-45 |
| Sons | US | Administrative | Mitnik et al. (2015, p. 98) | Individual labour earnings | 1972-1975 | 36.5 (average) |
| Sons | Great Britain | Survey | Gregg et al. (2017, p. 96) | Individual labour earnings | 1970 | 34 (centred) |

On these estimates, intergenerational income persistence among sons in New Zealand is the highest of all countries considered - the confidence interval overlapping only with those for the US and Great Britain - and is significantly higher than intergenerational persistence in the four Nordic countries and Australia. Among daughters, intergenerational persistence is also estimated to be substantially higher in New Zealand than in Finland, Sweden, and Australia, moderately higher than in Great Britain, and similar to the US.

[147] This estimate for Great Britain is used in the son-parent cross-national comparison below in preference to Björklund and colleagues' (2017) IGE estimate of 0.271 for British son-parent pairs reported above. Clearly, these estimates – derived from the same data source – differ markedly. The difference may be due to the different treatment of zero incomes and the different ages over which sons' earnings are averaged between the two studies.

[148] Raaum et al. (2007) also estimate the IGE of daughters' earnings (observed at or near age 41) with respect to their parents' combined earnings for daughters in Denmark, Norway, Great Britain (all born 1958), Finland (born 1956 to 1960) and the US (born 1957 to 1964). They use a specification that adds higher-order polynomials in parental income to account for nonlinearities in intergenerational persistence and they evaluate the IGE at the 50th percentile of the parental income distribution. They estimate IGEs for daughters of 0.190 in Denmark, 0.186 in Norway, 0.270 in Great Britain, 0.197 in Finland, and 0.252 in the US. These estimates are not used in the cross-country comparison above due to the different specification used.

**Figure 42**. *Comparison with other studies of intergenerational income elasticity estimate (dependent variable: offspring's lifetime average earnings) for sample of daughters.*




| *Sex* | *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|---|
| Daughters | NZ (CHDS) | Survey | | Individual labour earnings | 1977 | 40 |
| Daughters | Finland | Administrative | Österbacka (2001, p. 474) | Individual labour earnings | 1950-1960 | 25-45 |
| Daughters | Sweden | Administrative | Björklund et al. (2017, appendix p. 1) | Individual labour earnings | 1969 | 30, 34, and 38 |
| Daughters | Australia | Administrative | Deutscher & Mazumder (2020) | Individual total income | 1978-1982 | 29-37 |
| Daughters | Great Britain | Survey | Björklund et al. (2017, appendix p. 1) | Individual labour earnings | 1970 | 30, 34, and 38 |
| Daughters | US | Survey | Choi et al. (2020, p. 366) | Individual labour earnings | 1963-1975 | 35-42 |

Turning to the IGE estimates from the regression with offspring's family income over adulthood as the dependent variable, these are compared to results from studies of Norway, Sweden, Germany, the US, Canada, Australia, and Great Britain. Bratberg et al. (2017) use data on 324,870 Norwegian offspring drawn from the full-population administrative register born between 1957 and 1964 whose family earnings are observed over 11-year windows across ages 32 to 49 years, 251,288 Swedish offspring drawn from the full-population administrative register born between 1957 and 1964 whose household incomes are observed on seven occasions across ages 32 to 50, 1,128 German offspring drawn from the German Socio-Economic Panel born between 1957 and 1976 whose household total incomes are observed over 12-year windows across ages 25 to 55, and 6,298 American offspring drawn from the National Longitudinal Survey of Youth (NLSY79)[149] born between 1957 and 1964 whose total family incomes are observed over windows of varying length across ages 32 to 52 years (minimum of one annual observation, maximum of seven).[150] They estimate the IGE of offspring's family (or household) income with respect to their parents' family (or household) income to be 0.194 in Norway, 0.231 in Sweden, 0.314 in Germany, and 0.432 in the US. Corak (2020a) uses data on over two million Canadian sons and daughters drawn from administrative tax data born between 1963 and 1970 whose total family incomes are observed over five-year windows encompassing ages 34 to 45 (minimum of one annual observation, maximum of five) and estimate the IGE of offspring's log total family

[149] The NLSY79 is a study of a nationally representative sample of youth in the US who were aged 14 to 22 when they were first surveyed in 1979.
[150] The samples for Sweden, Norway, and Germany in Bratberg et al. (2017) were selected to mirror the US sample as closely as possible, for the purposes of cross-country comparison. The IGE is not a focus of their analysis and is estimated merely as a benchmark against the extant literature.

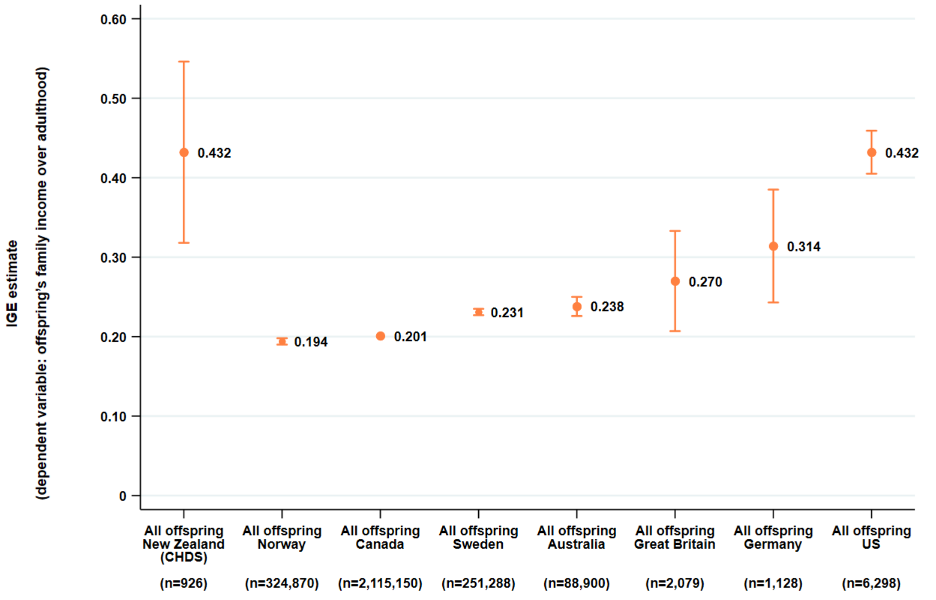
income with respect to parents' log total family income to be 0.201 for all offspring pooled, 0.223 for sons, and 0.172 for daughters.[151] Deutscher and Mazumder's (2020) study, described above, estimates the IGE of Australian offspring's log family income with respect to their parents' log family income to be 0.188 among sons and 0.181 among daughters. Deutscher and Mazumder also perform a sensitivity check in which only the 1978 birth cohort is used (for all offspring pooled, thus no sex-specific estimates) so that offspring incomes are centred on age 36 (instead of age 33 in the baseline analysis), described by Deutscher and Mazumder (2020) as a "conservative" estimate because it "restrict[s] the sample to those children and parents where income is measured closest to mid-life".[152] The IGE estimate from this sensitivity check, equal to 0.238, is the Australian estimate used in the cross-country comparison of 'all offspring', since it comes from a model specification that most closely resembles the CHDS model. Hirvonen (2008) uses administrative data on Swedish sons and daughters drawn from the full-population register born between 1962 and 1965 whose family earnings are observed in 1999 (ages 34 to 37) and estimates the IGE of offspring's log family earnings with respect to their parents' log family earnings to be 0.296 among sons and 0.249 among daughters. Rohenkohl's (2019) study, described above, estimates the IGE of offspring's log family income with respect to parents' log household income to be 0.270 among 2,079 British offspring. Because Rohenkohl (2019) does not report sex-specific estimates, these are taken from Blanden et al. (2004), described above, who estimate the IGE of sons' and daughters' family income at age 30 with respect to their parents' family income to be 0.300 among 2,015 sons and 0.307 among 2,285 daughters from the British Cohort Study. Finally, Mitnik and colleagues' (2015) study, also previously described, estimates the IGE of offspring's log family income with respect to parents' log family income to be 0.430 among sons and 0.360 among daughters in the US.[153] Figure 43 plots these IGE estimates alongside the CHDS all offspring estimate, and Figures 44 and 45 plots the son-specific (Figure 44) and daughter-specific (Figure 45) estimates alongside the equivalent CHDS estimates.

[151] These estimates are for Canada as a whole, but the main aim of Corak (2020a) is to estimate IGEs for regions within Canada. The sample size is over three million offspring-parent pairs when weighted. Family total income is defined for parents as the sum of fathers' and mothers' total incomes (i.e., income from all sources) and for offspring the sum of offspring's and their spouses' total incomes. For both parents and offspring, these family incomes are 'individualised', that is, divided equally between spouses (where partnered), thus producing a measure of "individual permanent income from all sources…assuming equal sharing between the adult partners in the household when more than one is present" (Corak, 2020a, p. 2142). Parental and offspring average family incomes are required to be no less than $500 to be retained in the sample.

[152] This restricted sample is also weighted to account for differences in the probability that a child is successfully linked to their parents and uses only the highest quality links.

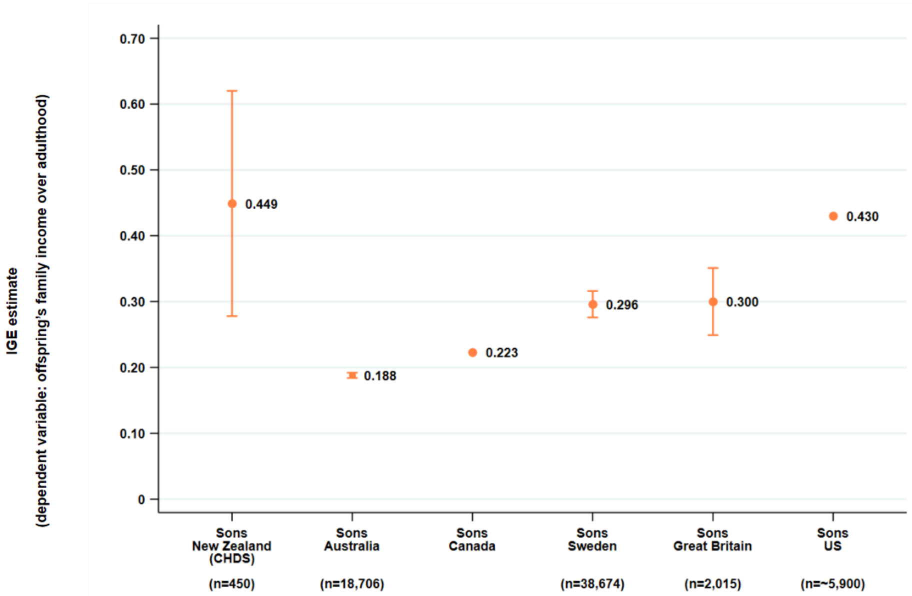
[153] These are Mitnik and colleagues' (2015) IGE estimates when individuals without tax or other administrative income data are dropped.

**Figure 43**. *Comparison with other studies of intergenerational income elasticity estimate (dependent variable: offspring's family income over adulthood) for sample of all offspring.*



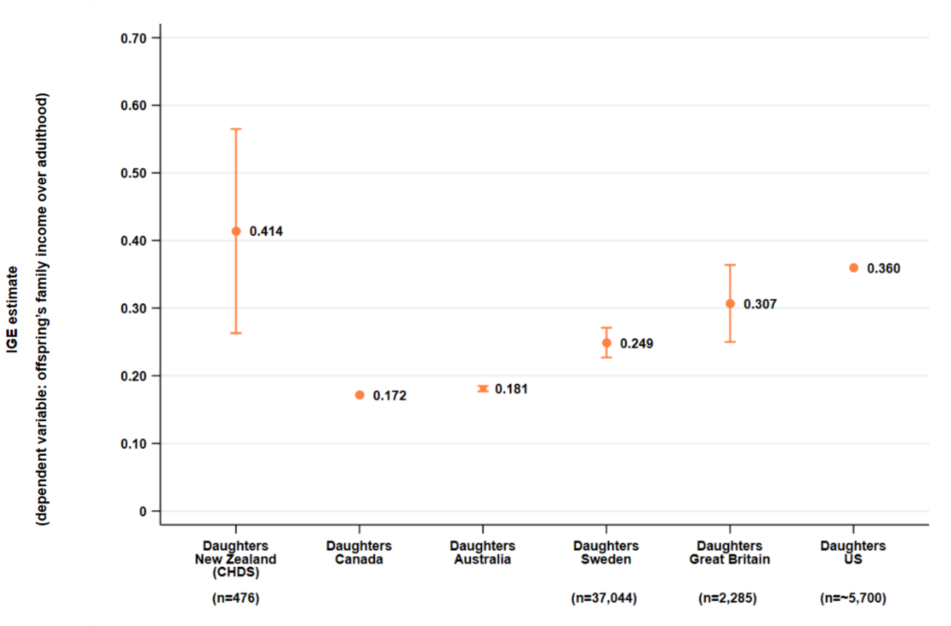

| *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|
| NZ (CHDS) | Survey | | Family income | 1977 | 30, 35, and 40 |
| Norway | Administrative | Bratberg et al. (2017, p. 82) | Family income | 1957-1964 | 32-49 |
| Canada | Administrative | Corak (2020, p. 2143) | Family income | 1963-1970 | 34-45 |
| Sweden | Administrative | Bratberg et al. (2017, p. 82) | Household income | 1957-1964 | 32-50 |
| Australia | Administrative | Deutscher & Mazumder (2020) | Family income | 1978 | 36 (centred) |
| Great Britain | Survey | Rohenkohl (2019, p. 31) | Family income | 1973-1991 | 30 (average) |
| Germany | Survey | Bratberg et al. (2017, p. 82) | Household income | 1957-1976 | 25-55 |
| US | Survey | Bratberg et al. (2017, p. 82) | Family income | 1957-1964 | 32-52 |

**Figure 44.** *Comparison with other studies of intergenerational income elasticity estimate (dependent variable: offspring's family income over adulthood) for sample of sons.*




| *Sex* | *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|---|
| Sons | NZ (CHDS) | Survey | | Family income | 1977 | 30, 35, and 40 |
| Sons | Australia | Administrative | Deutscher & Mazumder (2020) | Family income | 1978-1982 | 29-37 |
| Sons | Canada | Administrative | Corak (2020, p. 2143) | Family income | 1963-1970 | 34-45 |
| Sons | Sweden | Administrative | Hirvonen (2008, p. 792) | Family earnings | 1962-1965 | 34-37 |
| Sons | Great Britain | Survey | Blanden et al. (2004, p. 132) | Family income | 1970 | 30 |
| Sons | US | Administrative | Mitnik et al. (2015, p. 95) | Family income | 1972-1975 | 36.5 (average) |

**Figure 45**. *Comparison with other studies of intergenerational income elasticity estimate (dependent variable: offspring's family income over adulthood) for sample of daughters.*




| *Sex* | *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|---|
| Daughters | NZ (CHDS) | Survey | | Family income | 1977 | 30, 35, and 40 |
| Daughters | Canada | Administrative | Corak (2020, p. 2143) | Family income | 1963-1970 | 34-45 |
| Daughters | Australia | Administrative | Deutscher & Mazumder (2020) | Family income | 1978-1982 | 29-37 |
| Daughters | Sweden | Administrative | Hirvonen (2008, p. 789) | Family earnings | 1962-1965 | 34-37 |
| Daughters | Great Britain | Survey | Blanden et al. (2004, p. 132) | Family income | 1970 | 30 |
| Daughters | US | Administrative | Mitnik et al. (2015, p. 95) | Family income | 1972-1975 | 36.5 (average) |

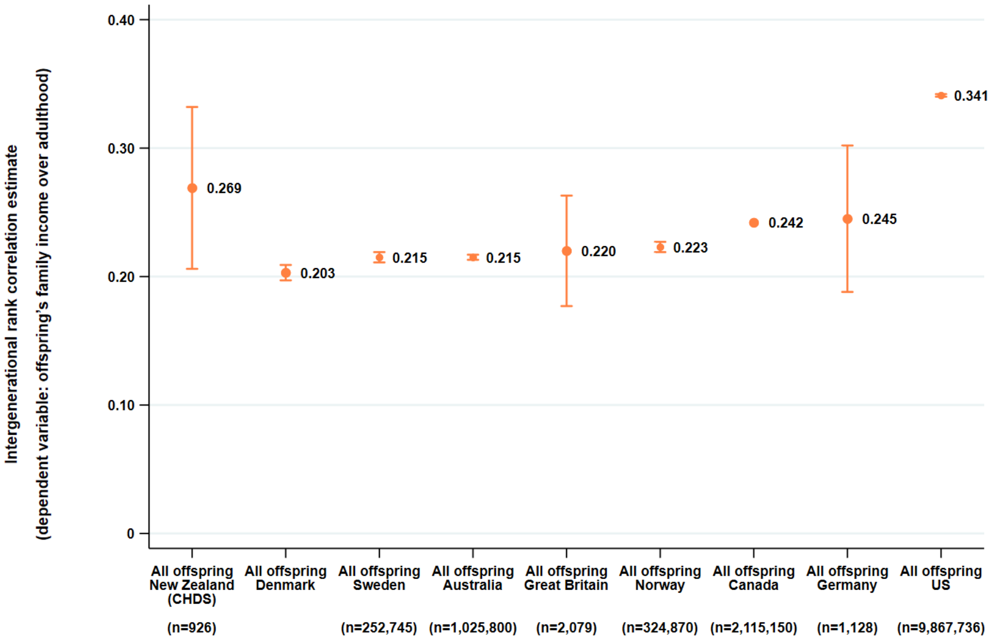
The comparison in Figure 43 positions New Zealand as having similar intergenerational persistence in family incomes to US offspring and much higher persistence than in the Nordic countries, Australia, and Canada, with German and British offspring somewhere in between the two groupings. Figures 44 and 45 similarly position New Zealand sons and daughters as having much higher persistence than their counterparts in Australia, Canada, and Sweden, but now somewhat higher persistence than in the US (note that the sex-specific US estimates are from Mitnik et al. (2015), whereas the pooled offspring estimate for the US is from Bratberg et al. (2017), as Mitnik et al. (2015) do not report an IGE of family income for sons and daughters pooled)[154] and Great Britain again occupying a middle-ground position (except among sons where the estimate is near-identical to the estimate for Swedish sons). Taking all the IGE cross-country comparisons together (across Figures 40 to 45), the picture painted is of New Zealand displaying a high degree of intergenerational income persistence, comparable to that found in the US, and (with very few exceptions) considerably higher than in the Nordic countries, Australia,[155] and Canada (with the British rankings being somewhat unstable).

[154] The US frequently ranks as having greater intergenerational persistence among sons and daughters than New Zealand if survey-based estimates of the IGE are instead used for comparison. For instance, Chadwick and Solon (2002) use data on 395 daughters and 380 sons drawn from the US PSID born between 1951 and 1961 with incomes observed in 1991 (when offspring were aged between 30 and 40) and estimate the IGE of family income to be 0.584 among sons and 0.443 among daughters.

[155] The IGE comparison with Australia is less clear-cut when results from survey data are instead used for Australia. Murray et al. (2018) use data on 489 offspring-parent pairs drawn from the longitudinal Household, Income and Labour Dynamics in Australia (HILDA) survey of a nationally-representative household panel, who were born between 1984 and 1986 with incomes observed at ages 28 to 31 and parental household incomes averaged over five years, and estimate the IGE of household total income in Australia to be 0.282. They find that their IGE estimate increases as the number of years over which parental income is averaged increases; averaging over five years reduces attenuation bias by

Turning now to the rank correlation estimates, these are compared to results from studies of Denmark, Sweden, Norway, Germany, Australia, Great Britain, Canada, and the US. The rank correlation between offspring's family income (or household income) and parents' family income (or household income) is estimated to be 0.203 in Denmark (0.215 among sons, 0.190 among daughters) (Helsø, 2021), 0.215 in Sweden, 0.223 in Norway, 0.245 in Germany (Bratberg et al., 2017), 0.215 in Australia (0.217 among sons, 0.211 among daughters) (Deutscher & Mazumder, 2020), 0.220 in Britain (Rohenkohl, 2019), 0.242 in Canada (Corak, 2020a), and 0.341 in the US (0.336 among sons, 0.346 among daughters) (Chetty et al., 2014). Figure 46 plots these estimates for all offspring alongside the CHDS estimate, and Figure 47 plots the son-specific and daughter-specific estimates alongside the equivalent CHDS estimates. On these estimates, persistence in income rank is higher in New Zealand than in the Nordic countries and Australia[156] (with Great Britain also amongst this grouping), somewhat higher than in Canada and Germany, and lower than in the US, an overall pattern similar to that found for the IGE. However, the confidence intervals of the CHDS estimates overlap with those of all other countries, with the

**Figure 46**. *Comparison with other studies of intergenerational rank correlation estimates (dependent variable: offspring's family income over adulthood) for sample of all offspring.*



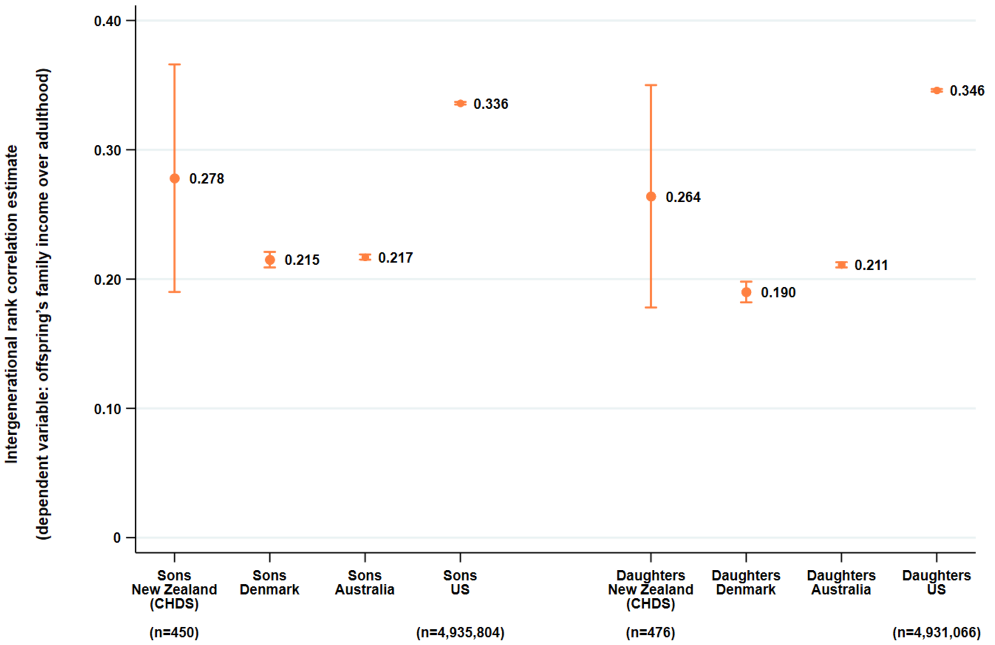

| *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|
| NZ (CHDS) | Survey | | Family income | 1977 | 30, 35, and 40 |
| Denmark | Administrative | Helsø (2021, p. 516) | Family income | 1980-1982 | 29-32 |
| Sweden | Administrative | Bratberg et al. (2017, p. 83) | Household income | 1957-1964 | 32-50 |
| Australia | Administrative | Deutscher & Mazumder (2020) | Family income | 1978-1982 | 29-37 |
| Great Britain | Survey | Rohenkohl (2019, p. 31) | Family income | 1973-1991 | 30 (average) |
| Norway | Administrative | Bratberg et al. (2017, p. 83) | Family income | 1957-1964 | 32-49 |
| Canada | Administrative | Corak (2020, p. 2144) | Family income | 1963-1970 | 34-45 |
| Germany | Survey | Bratberg et al. (2017, p. 83) | Household income | 1957-1976 | 25-55 |
| US | Administrative | Chetty et al. (2014, p. 1571) | Household income | 1980-1982 | 30.5 (average) |

---

approximately 44 percent relative to using one parental income observation. Drawing on this finding, and on empirical estimates of attenuation bias from Mazumder (2005), they argue that their IGE estimate of 0.282 is likely to be attenuated by (at most) 30 percent due to averaging over only five years of parental income, "leading to a bias-adjusted IGE estimate of 0.409" if parental lifetime income were available (Murray et al., 2018, p. 452). Such an adjusted estimate would place Australia closer to the CHDS estimate for New Zealand.

[156] As for the IGE (discussed in footnote 155 above), the rank correlation for Australia is much closer to (indeed higher than) the CHDS estimate for New Zealand if Murray and colleagues' (2018) survey-based estimate for Australia of 0.273 (using household total income for both parents and offspring) is used in the comparison. Preference has been given here to Australian estimates from Deutscher and Mazumder (2020) on the grounds of considerably larger and more representative samples from administrative sources, and because offspring incomes are observed, on average, at an older age than in Murray et al. (2018) (and thus a more comparable age to CHDS).

**Figure 47**. *Comparison with other studies of intergenerational rank correlation estimates (dependent variable: offspring's family income over adulthood) for samples of sons and daughters.*




| *Sex* | *Country* | *Data source* | *Reference* | *Offspring income concept* | *Birth cohorts* | *Age offspring income observed* |
|---|---|---|---|---|---|---|
| Sons | NZ (CHDS) | Survey | | Family income | 1977 | 30, 35, and 40 |
| Sons | Denmark | Administrative | Helsø (2021, p. 516) | Family income | 1980-1982 | 29-32 |
| Sons | Australia | Administrative | Deutscher & Mazumder (2020) | Family income | 1978-1982 | 29-37 |
| Sons | US | Administrative | Chetty et al. (2014, p. 1571) | Household income | 1980-1982 | 30.5 (average) |
| Daughters | NZ (CHDS) | Survey | | Family income | 1977 | 30, 35, and 40 |
| Daughters | Denmark | Administrative | Helsø (2021, p. 516) | Family income | 1980-1982 | 29-32 |
| Daughters | Australia | Administrative | Deutscher & Mazumder (2020) | Family income | 1978-1982 | 29-37 |
| Daughters | US | Administrative | Chetty et al. (2014, p. 1571) | Household income | 1980-1982 | 30.5 (average) |

exception of the US estimate for all offspring which is unambiguously higher than those of all other countries. Recall that the rank correlation is a *positional* measure of income persistence, so the mapping of ranks to actual living standards will differ across these countries. For example, moving up ten percentiles in the income distribution will translate into a greater increase in absolute income in a high-inequality country than in a low-inequality country.

Other analyses of CHDS data examining the relationship between family background and offspring socioeconomic, educational, and health outcomes in adulthood also find significant intergenerational associations. For instance, Gibb, Fergusson and Horwood (2012) use multiple linear regression to test for associations between cohort members' quintile of family income in childhood and their annual personal income at age 30 (and other economic, educational, and health outcomes).[157] Adjustment for a wide range of covariates drawn from the CHDS dataset (including parents' educational attainment and child IQ) proceeded in three stages: a baseline regression model in which family size was the only covariate included in the model, a full model that included all covariates, and a reduced model that included only statistically significant

[157] Family income in childhood was averaged over ages 1 to 10 years and children were then grouped into quintiles of average family income over childhood. Annual personal income at age 30 was measured by cohort members selecting one of 12 income bands spanning "Zero income or loss" up to "$100,001 or more" (with cohort members assigned the mid-point of their band except for the top category which was top-coded at $120,000).

covariates. In all three models, they find a statistically significant association between childhood family income quintile and annual personal income at age 30. In the baseline model, which controls for family size only and thus comes closest to measuring the 'gross' association between the two variables, as per the IGE, they find a statistically significant association between childhood family income and annual personal income (*p*<0.0001) with a coefficient of $5,289 (that is, a one quintile increase in childhood family income is associated with a $5,289 increase in annual personal income at age 30, controlling for family size). In the reduced model containing only statistically significant covariates, the association remained highly significant (*p*<0.0001) and children from the highest quintile of family income earned on average $11,750 more per annum at age 30 than children from the lowest quintile, after adjustment for all significant covariates.[158] Gibb et al. (2012, p. 1984) conclude that "[i]n general these findings clearly suggest an intergenerational transmission of educational and economic privilege in which increasing childhood family income was associated with increasing educational and economic privilege even after adjustment for a wide range of childhood, family and related factors".

Boden, Fergusson and Horwood (2013) use CHDS data to investigate the predictors of cohort members' *family incomes at age 30*, measured as the sum of offspring's and their partner's weekly annualised total incomes at age 30 equivalised for household size and composition. A wide range of predictors were drawn from the CHDS database including cohort members' sex and ethnicity, their parents' educational attainment, their family socioeconomic status at birth, their family's average standard of living over childhood, their own educational attainment by age 30, their experiences of mental disorder and substance abuse in adulthood, and characteristics of the cohort member's own family at age 30 (whether the cohort member was living in New Zealand, number of earners in the family, number of dependent children, early parenthood, single parenthood, family receipt of social welfare benefits). These predictors were subject to stepwise selection of variables to fit a multivariable model. The significant predictors of family income at age 30, adjusted for all other predictors, were mainly characteristics of the cohort member's own family at age 30 (higher family incomes were associated with more earners in the family, smaller family size, and living overseas), all of which had relatively large effect sizes as measured by standardised coefficients. Two other predictors – the cohort member's highest qualification attained by age 30 and the cohort member's average family living standards over their childhood – also exerted significant effects on family income at age 30, but with smaller effect sizes. Boden et al. (2013, p. 129) conclude that family income in adulthood is "strongly linked to earning power [of the family, as manifested in number of earners, number of children, and whether resident in New Zealand], educational achievement, mental health and substance use disorders, and the lingering influence of family living standards in childhood". The latter result on the influence of family living standards in childhood on cohort members' family income in adulthood leads the

[158] Gibb et al. (2012) note that the associations between childhood family income and later adult outcomes tended to be linear, suggesting that the associations do not merely reflect the adverse effects of poverty on later outcomes. Rather, "the evidence suggests a process in which there is an intergenerational transmission of educational achievement and economic achievement in which increasing levels of childhood family income are associated with increasing levels of educational and economic success independently of contextual factors" (Gibb et al., 2012, p. 1985).

authors to state that their results provide evidence for an intergenerational transmission of economic status:

> An interesting finding was that independently of other factors in the regression models (earning power; education; mental health; substance use), exposure to socioeconomic disadvantage during childhood, as measured by average family living standards, was a predictor of economic inequality [differences in family income] at age 30. These findings clearly hint at processes of intergenerational transmission in which, independently of other factors, the individual's level of exposure to economic advantage/disadvantage during childhood carries over into adult life….These findings hint at the presence of a series of non-observed socialisation processes in which advantage begets advantage and disadvantage begets disadvantage. The nature of these processes is unknown but it can be conjectured that childhood family economic conditions are associated with a range of factors including exposure to parental role models, the development of life course expectations and aspiration which lead to an intergenerational transmission of economic advantage and disadvantage (Boden et al., 2013, pp. 126-127).

Maloney et al. (2003) use CHDS data to estimate the intergenerational correlation in social welfare benefit receipt between parents and their offspring in New Zealand. CHDS cohort members' 'benefit propensity' – the proportion of months through ages 16 and 21 that they were in receipt of the unemployment benefit or domestic purposes benefit – is compared with their parents' benefit propensity, the proportion of years over the cohort members' childhood (ages 1 to 14) that either of their parents were in receipt of these benefits. Using both ordinary least squares and instrumental variables estimation to produce lower-bound and upper-bound estimates of the true intergenerational correlation coefficient, they find that the intergenerational correlation coefficient is at least 0.312, which is more than two times higher than estimates found in the Nordic countries and Canada of 0.10 to 0.16 (Moisio et al., 2015). These results are consistent with the results found here for income persistence in that they both suggest New Zealand has comparatively high intergenerational transmission of economic outcomes.

Sociological analyses of social mobility in Christchurch also find evidence of intergenerational persistence in social class. Lauder and Hughes (1990b) surveyed 2,753 students after they left twenty secondary schools in Christchurch in 1982, collecting information on their educational qualifications attained upon completing school, cognitive ability test scores at around age 13, their first occupation after leaving school, and their father's occupation. Students' occupations (their 'social destinations') and their fathers' occupations (their 'social origins') were ranked into a hierarchy of three social classes: the professional class (professional and managerial occupations), the lower middle class (office, sales, and technical occupations), and the working class (skilled, semi-skilled, and unskilled manual occupations). Students who went to university or into professional training (e.g., teachers' college) after secondary school were assumed to end up in professional and managerial occupations. To analyse intergenerational social mobility, Lauder and Hughes cross-tabulate students' social origins with their social destinations. Among students from professional/managerial backgrounds, just under half were employed in, or on track to enter, a professional/managerial occupation themselves, about one-third had experienced downward mobility into the lower middle class, and the remaining 20 percent had entered the working class. Among students from working class backgrounds, about 55 percent had working class jobs themselves, about one-third had experienced upward mobility to the lower middle class, and less than 15 percent had risen into the professional/managerial class. Students from lower middle class backgrounds were fairly evenly spread across the three classes (although a plurality

– just under 40 percent – had lower middle-class occupations themselves). Overall, there was strong intergenerational persistence among students from professional and working class backgrounds, and little long-range mobility either upward or downward. Lauder and Hughes (1990b, pp. 54, 55) conclude that "social class [background] has a major impact on school performance and subsequent occupational destinations" and "New Zealand is not the open socially mobile society the common myth of everyone getting a 'fair go' suggests", yet "there is enough mobility to provide a material basis for the perpetuation of the equal opportunity myth".[159]

Hall et al. (1983) use occupational information from marriage certificates (which record the occupations of groom, bride, and their respective fathers), collected from two Christchurch suburbs spanning marriages between 1866 and 1981, to analyse the degree of intergenerational persistence between the groom's (son's) occupation and his father's occupation (at the time of marriage; n=1,577).[160] Sons' and fathers' occupations are ranked on a seven-point occupational prestige ranking scale and then cross-tabulated in transition matrices. They find that about one-third of sons are found in the same occupational rank as their father, surmising that "there has been considerable occupational recruitment [persistence] from father to son within the same broad occupational category", and that the mobility that *did* occur was "mostly short range movement" from blue-collar to lower white-collar occupations (Hall et al., 1983, pp. 132, 140). Their Christchurch data are re-analysed in Pearson and Thorns (1986), but for the period 1928 to 1970,[161] and similar conclusions are reached. Overall, they conclude that their results on intergenerational mobility, when viewed in conjunction with their results on marital homogamy, present a picture of "a social structure which is clearly stratified and does not have extensive upward or downward mobility" and of a society that "possesses broad constraints to individual opportunity", challenging "popularly-held assumptions" of New Zealand as an open and mobile society (Hall et al., 1983, pp. 130, 128; Pearson & Thorns, 1986, p. 222). However, significant changes have occurred in the labour market between when these studies end (up to the early

[159] Lauder and Hughes (1990b) also find that students' socioeconomic background exerts an influence on their educational attainment and occupational destinations even when measured cognitive ability is controlled for, as students from lower socioeconomic backgrounds, on average, achieved lower levels of qualifications and entered occupations of a lower socioeconomic status than their *equally-able* counterparts from higher socioeconomic backgrounds, meaning these students "leave school vastly under qualified given their potential", constituting a significant "wastage of talent in New Zealand" and contradicting "the meritocratic thesis that those of equal measured ability have equal life chances regardless of their social background" (Hughes & Lauder, 1991, pp. 14, 15; Lauder & Hughes, 1990b, p. 49). In a separate paper comparing the twenty schools in their sample by their average levels of student educational attainment and occupational destinations in both 'unadjusted' terms and adjusted for the gender, cognitive ability, and social class background of their student bodies, Lauder and Hughes (1990a, pp. 55, 56) find that the 'socioeconomic mix' of the school – the ratio of high SES students to low SES students – has "an impact both on pupil achievement and on their destinations when leaving school" holding constant school differences in average cognitive ability, namely that educational attainment is, in effect, 'dragged down' by attending a school that has a predominantly disadvantaged socioeconomic mix (and dragged up by a privileged socioeconomic mix, especially among students of lower cognitive ability), concluding that "[i]n our view the most direct way of boosting the life chances of working-class students is to equalise school mix".

[160] As they also recorded the occupations of brides, Hall et al. (1983, pp. 140, 134) are able to analyse intergenerational occupational mobility between daughter-father pairs, finding "differences in the mobility chances of sons and daughters" with "class recruitment [intergenerational persistence] of daughters…lower than for sons", that is, they observe higher intergenerational mobility among daughters, consistent with the results found here. However, they note that "[o]ne of the problems in examining female mobility, particularly prior to the 1960s, is the small number of brides who held an occupation on marriage" (Hall et al., 1983, p. 130).

[161] In Pearson and Thorns' (1986) re-analysis of the Christchurch data, they exclude all grooms over the age of 35 at the time of marriage to minimise age differences between grooms (thereby reducing 'life-cycle bias' from observing sons at different points in their careers), exclude all sons with farming occupations (which do not fit neatly into traditional occupational hierarchies), and re-code the occupational rankings in order to harmonise them with analogous occupational data from Wellington which coded occupations into social classes based on a socio-economic index (rather than an occupational prestige scale).

1980s) and when the incomes of the CHDS offspring are observed (2007 to 2017), so the opportunities for mobility in Christchurch may have changed over this period.

### 4.6 Comparison with results from the New Zealand Longitudinal Census

The analyses of NZLC data in chapter 3 estimated the intergenerational elasticity of individual total income to be 0.239 among son-father pairs, 0.054 among son-mother pairs, 0.145 among daughter-mother pairs, and 0.135 among daughter-father pairs and, on the basis of these results, concluded that, by international comparison, New Zealand has relatively low to middling intergenerational income persistence (with the exception of daughters in relation to their mothers), hence can be broadly characterised as a country with comparatively average-to-high mobility. Clearly, this conclusion is at odds with the conclusion from the CHDS analyses just presented. Why the apparent discrepancy?

Recall that differences between the two data sources in income concepts, study design, and other data-related and sample-related differences meant that each data source lent itself to a different conceptual approach to analysis. For example, the NZLC dataset contains annual total income from all sources and matches individual offspring to individual parents, whereas the CHDS dataset contains offspring's earnings and family income and parental family income (pooling mothers' and fathers' incomes), all measured in weekly figures. In the NZLC analysis, parental income is measured around a particular age when current income is judged to represent lifetime average income, whereas in the CHDS analysis, parental income is measured over their offspring's childhood (which may not necessarily coincide with the lifetime-representative age). Being based on census records, the NZLC sample covers all geographic areas of the country and multiple birth cohorts, whereas the CHDS is a single birth cohort from a single city with a relatively ethnically-homogenous population. These data differences, and the different conceptual approaches consequently taken to their respective analyses, may in part explain the different results. Divergence in findings between the permanent income and childhood resources approaches has been observed in the literature before. For example, adopting a permanent income approach, Bratsberg et al. (2007) estimate the intergenerational elasticity of individual labour earnings among son-father pairs in Denmark to be 0.121 and Jäntti et al. (2006) estimate the same elasticity among daughter-father pairs in Denmark to be 0.034. In contrast, adopting a childhood resources approach, Helsø (2021) estimates the elasticity of offspring's individual labour earnings with respect to their parents' family income when the offspring were aged 7 to 15 to be 0.327 in Denmark (albeit, for cohorts born approximately 16 years after those studied in Bratsberg et al. and Jäntti et al.).

In the end, is difficult to gauge the extent to which the qualitatively different conclusions reached by the analyses in chapters 3 and 4 reflect real differences in different concepts of mobility or simply differences in the 'who, what, and when' of the intergenerational data. Further research is needed to understand why the two data sources paint disparate pictures of intergenerational mobility in New Zealand.

### 4.7 Bias and limitations

As prefaced in section 4.1.1, a limitation of the CHDS data is that there has been differential attrition in the cohort. For example, CHDS researchers analysing data from the subset of cohort members followed-up at age 30 found that, when compared with those cohort members not followed up at that age, the former was found to "under-represent participants from more socioeconomically disadvantaged backgrounds (low parental education, low SES, low income, single parent family) and males" (Gibb et al., 2012, p. 1983). Tests at age 21 similarly found "small but statistically significant" differential attrition (Fergusson & Horwood, 2001, p. 288). The large statistical imprecision in estimates, arising from the relatively small size of the CHDS sample, is another limitation. However, because the differential attrition observed in the cohort tends to be small in magnitude, and because differential attrition does not necessarily result in selection bias,[162] Fergusson and Horwood (2013, p. 80) argue that tests of "the effects of sample losses on study validity….have invariably shown that while the sample losses have reduced study precision (as a result of reduced sample size) they do not threaten study validity".

The broader issue, however, is that the design of the CHDS study means that the cohort (whether the full original cohort or an attrition-reduced subset) is unlikely to be representative of offspring born elsewhere in New Zealand in the late 1970s (nor their parents), particularly in terms of ethnic composition, as canvassed in section 4.3.2.[163] Does this lack of representativeness imply that the current findings cannot be generalised to the wider population of interest? Addressing this question in relation to the longitudinal Dunedin Study, Poulton et al. (2006, p. 53) draw a distinction between 'representativeness' and 'generalisability':

> Classically, representativeness refers to sampling methods that faithfully represent all members of the target population (in a New Zealand nationally representative study this would mean the whole of the country), whereas generalisability refers to the ability to extrapolate findings to the wider population, despite imperfect representativeness. Deriving a sample that is perfectly representative of the major population groups of interest (in terms of socioeconomic status and geographic location for example) is resource intensive and costly. Moreover, for a longitudinal study, generalisability is more important than representativeness. By its nature, a cohort study cannot remain truly representative of the population of interest. Thus, although the Dunedin Study sample appears to be broadly representative of New Zealand children born in 1972/1973, they will not necessarily be representative of children born in 1992 or 2002. Nevertheless, it is today's New Zealand children that are most likely to benefit from the lessons that we have learned from the Dunedin Study.

Poulton et al. (2006, pp. 45, 54, 53) then conduct an analysis comparing health outcomes among cohort members from the Dunedin Study with those of people of the same age in the nationally-representative New Zealand Health Survey and National Nutrition Survey, and find that "[f]or most outcomes, the Dunedin Study members were very similar to the nationally representative samples" meaning "findings from the Dunedin Study are likely to be broadly generalisable to the

---

[162] As noted in section 3.6.1, differential attrition only leads to selection bias where the exposure-outcome association differs between participants retained in the study and participants lost to follow-up (Carter et al., 2012; Lacey et al., 2014).

[163] A related issue is that the birth cohort design means that findings may not represent the experiences of immigrants to New Zealand. However, this is not a major concern for the current analysis for which the target population of interest is children who are born in New Zealand, as it is their life experiences that best capture the effects that New Zealand's national culture and social institutions have on intergenerational mobility. On the other hand, unlike the NZLC data analysed in chapter 3 (see section 3.6.2), the CHDS does capture the experience of emigrants, albeit a small subsample. This is because cohort members who move overseas in adulthood are retained in the study and followed-up where possible, meaning intergenerational mobility could be analysed and compared between cohort members who remain in New Zealand and those who emigrate (this analysis was not performed as part of the current study as it is not central to the aims of this thesis).

wider New Zealand population" and "by implication similar studies such as the Christchurch Health [and] Development Study". In discussing the generalisability of his own findings on intergenerational income mobility using data from the Dunedin Study, Gibbons (2010, p. 44) also makes the point that, owing to New Zealand's unitary state, "[w]elfare payments, entitlements to public education and health services, and the minimum wage are the same throughout New Zealand". The crucial point is that, for a correlational analysis of the kind presented here, the sample only needs to be representative with respect to *the correlation of interest* for the findings to be generalisable. Epidemiologists Rothman, Greenland, and Lash (2008, p. 146) note that the well-regarded British Physicians' Study of smoking and health and the Nurses' Health Study, "neither of which were remotely representative of the general population with respect to sociodemographic factors", were nevertheless internally valid studies that produced generalisable findings because the bias in the sample was unrelated to the effects being studied.[164]

The income data collected in the CHDS present some drawbacks for a study of intergenerational income persistence. As outlined in section 4.1.2, parents' incomes are computed in gross (before-tax) terms, including scaling after-tax social welfare benefit income by an arbitrary factor to approximate before-tax amounts, whereas offspring incomes are collected in net (after-tax) amounts. Furthermore, parents' market incomes are collected in bands, with an open-ended top category that is top-coded at the category's lower bound (and in theory are bottom-coded at zero as well, since market incomes can be negative yet there was no allowance for this in the income bands used). In addition, all income data are collected in weekly amounts.

The effects of banding on the estimates of persistence, previously discussed in section 3.6.3 in relation to the NZLC data, are likely to be small, based on the empirical evidence covered in that section (Björklund et al., 2017; Gregg, Macmillan, et al., 2017; Raaum et al., 2007; Ueda, 2009). However, in the current study the effects of banding will be partially offset by the averaging of income over multiple years – always for parents and wherever possible for offspring (except for daughters' earnings) – which creates much more variability in the resulting averaged data.

What are the effects of using weekly rather than the (more conventional) annual income data? In their cross-country comparison of intergenerational persistence of annual earnings in Denmark, Finland, Norway, Sweden, Great Britain, and the US, for which the British offspring earnings data are only available in weekly amounts, Jäntti et al. (2006) make use of the fact that their US data source also contains offspring's weekly earnings (in addition to annual earnings) to check the sensitivity of their Great Britain-US gap in persistence to a 'weekly-with-weekly' comparison.[165] They find some sensitivity in the US IGE estimates (the son-father IGE decreases by 11 percent and the daughter-father IGE decreases by 16 percent when switching to weekly earnings, thus shrinking the Great Britain-US difference) but no change in the correlation coefficients. Thus, switching from annual to weekly earnings appears to decrease IGE estimates, so the CHDS

[164] As an example, Rothman et al. (2008, p. 147) note that "although most of the decisive data connecting smoking to lung cancer was derived from observations on men, no one doubted that the strong effects observed would carry over at least approximately to women, for the lungs of men and women appear to be similar if not identical in physiologic detail". Thus, lack of representativeness does not necessarily threaten generalisability if the unrepresentativeness of the sample is unrelated to the correlation being studied.

[165] Weekly incomes cannot be reliably derived for the Nordic countries in the analysis.

estimates found in the current analysis may be biased upwards somewhat, relative to estimates using annual earnings (including those for countries covered in the cross-national comparison).[166]

Finally, as with the NZLC data, all CHDS income data are self-reported (in the case of fathers' and offspring's partners' incomes, they are reported by the mother and offspring, respectively) which potentially introduces recall bias (forgetting how much one receives in income per week, or one's spouse receives) and social desirability bias (e.g., reporting higher earnings than is actually the case to give a favourable impression or to avoid embarrassment).

[166] Pooling New Zealand Household Labour Force Survey respondents over the period 1997 to 2004, Hyslop and Yahanpath (2006) compare respondents' reports of their weekly total income (measured continuously) with reports of their annual total income (measured categorically) by annualising the weekly measure and then categorising it analogously to the annual measure. They find that "the simple correlation coefficient between individuals' weekly and annual income reports is 0.81" and that "the two distributions are reasonably comparable", except for differences between those reporting zero annual income (eight percent) and zero weekly income (16 percent) and those reporting annual income of $1 to $5,000 (nine percent) and the corresponding weekly income (five percent), consistent with the view that "some workers with low annual income are part-year workers and thus have zero income in some weeks" (Hyslop & Yahanpath, 2006, p. 309).

# Chapter 5: Transmission mechanisms of intergenerational income persistence

Identifying the mechanisms that lead to the intergenerational transmission of income is critical to understanding whether it can be influenced by individuals, families, and public policy, and if so, how. This chapter aims to understand how income persists across generations by studying transmission mechanisms that may give rise to the intergenerational association in incomes documented in chapters 3 and 4. Section 5.1 outlines the method used to quantify the role played by transmission mechanisms in generating intergenerational persistence, a method that decomposes the relationship between parental and offspring income into various pathways that mediate that relationship. Section 5.2 presents the results of the decomposition of the NZLC estimates of persistence and section 5.3 presents the results for the CHDS 'all offspring' estimate. Section 5.4 compares these results to those from other studies, and section 5.5 discusses the limitations of the analyses.

## 5.1 Accounting for intergenerational income persistence

Any skill or trait that is correlated across generations and linked to labour market outcomes contributes to intergenerational income persistence. Transmission mechanisms or pathways are therefore those characteristics observed in the offspring that are related to their parents' income *and* related to their own income.

To act as a channel of intergenerational persistence, a transmission mechanism must relate to parents' income and offspring's income *in the same direction*. That is, the pathway must either be *positively* associated with parental income and *positively* associated with offspring's income (in which case the pathway channels the intergenerational transmission of advantage) or be *negatively* associated with parental income and *negatively* associated with offspring's income (in which case the pathway channels the intergenerational transmission of disadvantage). In either case, the pathway gives rise to an intergenerational association (persistence) in incomes. An example of the former is the well-established finding that children's educational attainment is both positively related to parental income (children of richer parents obtain better educational outcomes on average, e.g., Maani and Kalb (2007)) and has a positive return in the labour market (children with higher qualifications earn more on average, e.g., Mercante and Mok (2014b)). An example of the latter is the finding that behavioural and emotional problems in childhood are negatively related to parental income (children of poorer parents exhibit higher rates of behavioural problems on average, e.g., Violato, Petrou, Gray and Redshaw (2011)) and have negative consequences in the labour market (children with certain behavioural and emotional problems earn less on average, e.g., Knapp, King, Healey and Thomas (2011)). In contrast, if a pathway is, say, *positively* associated with parental income but *negatively* associated with offspring's income, it does not act as a channel for intergenerational persistence; rather, the pathway acts as a channel for intergenerational *mobility* because it de-links parental income from offspring income (relating to each in opposing directions).

The analyses in this chapter account for the role that transmission mechanisms play in producing intergenerational persistence by decomposing estimates of persistence, measured here by the IGE, into various pathways that mediate the intergenerational association. The analyses utilise a decomposition approach commonly employed in the literature, e.g., by Österbacka (2001), Blanden, Gregg, and Macmillan (2007), Hirvonen (2010), Mood, Jonsson, and Bihagen (2012), Blanden et al. (2014), Gregg, Jonsson, et al. (2017), and Björklund et al. (2017). These authors all acknowledge that the decomposition approach they employ requires some strong assumptions and does not attempt to establish causal links. As noted in chapter 2, the IGE is a 'catch-all' measure capturing the influence on offspring income of all factors correlated with parent income, including omitted parental characteristics such as parents' education. The IGE therefore overstates (biases upwards) the *causal effect* of parental income on offspring income. When a mechanism such as offspring's education is then explored as a mediator of the intergenerational association in incomes, its mediating effect is also likely to be overstated since offspring's education will similarly be positively correlated with the same omitted parental characteristics that determine offspring's income. Thus, a mediation analysis of a non-causal income association "requires strong assumptions", in particular that any potential transmission mechanism is uncorrelated with omitted determinants of offspring's income (Mendolia & Siminski, 2017, p. 2). For this reason, the commonly-employed decomposition technique used here "is a descriptive and not a causal decomposition, as observed associations may also pick up effects of unmeasured factors" (Gregg, Jonsson, et al., 2017, p. 127) and provides "*suggestive evidence* of how parents with more income produce higher earning [offspring]" (Blanden et al., 2007, p. C58, emphasis added).[167] These limitations are discussed further in section 5.5.

In this decomposition approach, the IGE point estimate, represented by $\beta$, is decomposed into two components:

(a) the extent to which intergenerational persistence is transmitted through a pathway factor (for example, educational attainment), which is the product of two measures:

(i) the strength of the relationship between parental income and the pathway factor, and

(ii) the income pay-off of the pathway factor to the offspring.

(b) a non-pathway effect of parental income on offspring's income, indicating the effect of parental income not mediated through the pathway factor.

Following the example in Blanden et al. (2014), assume two pathway factors, offspring's educational attainment and offspring's occupation, are each measured by a single continuous variable (although in fact both variables are categorical in the current analysis). In the first step, the relationship between offspring's educational attainment and parental income is estimated by regressing offspring's educational attainment $Educ_i^{\text{offspring}}$ on a constant $\alpha_{educ}$, parents' log lifetime average income $Y_i^{\text{parent}}$, and an error term $e_{1i}$ capturing factors orthogonal to parental income:

[167] For a causal or 'structural' decomposition of the intergenerational income elasticity endeavouring to distinguish the effects of father's income from father's human capital on son's income, see Lefgren, Lindquist, and Sims (2012).

$$Educ_i^{\text{offspring}} = \alpha_{educ} + \lambda_{educ} Y_i^{\text{parent}} + e_{1i}, \quad (9)$$

In the second step, the pay-off to educational attainment in the labour market independent of parental income is estimated by regressing offspring's log lifetime average income $Y_i^{\text{offspring}}$ on a constant $\omega_1$, offspring's educational attainment, parents' log lifetime average income, and an error term $v_{1i}$:

$$Y_i^{\text{offspring}} = \omega_1 + \rho_{educ} Educ_i^{\text{offspring}} + \gamma_{inc} Y_i^{\text{parent}} + v_{1i}. \quad (10)$$

The overall $\beta$ is then decomposed into the income return to educational attainment multiplied by the relationship between parental income and education, plus the unexplained persistence in income that is not transmitted through (i.e., not accounted for by) education:

$$\beta = \rho_{educ} \lambda_{educ} + \gamma_{inc} \quad (11)$$

The first term in equation (11), $\rho_{educ}\lambda_{educ}$, is the explained component of $\beta$ (in this case, by education); the second term is the unexplained component of $\beta$.

The decomposition model of Blanden et al. (2014) adds the pathway variables in the order in which they generally occur in the life course, reflecting the fact that a variable such as educational attainment usually occurs chronologically prior to a variable such as occupation.[168] To add offspring's occupation into the model with educational attainment already included, the relationship between occupation $Occup_i^{\text{offspring}}$ and parental income, is first estimated as follows:

$$Occup_i^{\text{offspring}} = \alpha_{occup} + \lambda_{occup} Y_i^{\text{parent}} + e_{2i}, \quad (12)$$

Next, a regression equation is estimated that relates offspring income to both pathway factors, educational attainment and occupation:

$$Y_i^{\text{offspring}} = \omega_2 + \gamma_{educ} Educ_i^{\text{offspring}} + \gamma_{occup} Occup_i^{\text{offspring}} + \gamma_{inc} Y_i^{\text{parent}} + v_i. \quad (13)$$

Equation (13) provides estimates of the returns to each pathway variable, conditional on the other variables. The decomposition then becomes:

$$\beta = \lambda_{educ}\gamma_{educ} + \lambda_{occup}\gamma_{occup} + \gamma_{inc}. \quad (14)$$

The first term of equation (14), $\lambda_{educ}\gamma_{educ}$, is the component of $\beta$ associated with education, while $\lambda_{occup}\gamma_{occup}$ gives the component related to occupation. The difference between the education component from equation (11) and the education component from equation (14) ($\rho_{educ}\lambda_{educ}$ –

[168] As Blanden et al. (2014, p. 435) note, this chronology assumes that "education is determined independently of occupation, but that occupation depends on educational achievement" but "[t]his may not be the case if individuals take education with a specific career goal in mind". However, Blanden et al. (2014, p. 435) believe that their educational and occupational variables are classified at a sufficiently broad level that "this endogeneity problem [is] less likely than if we were using more specific occupations and qualifications such as "law degree" and "lawyer"". The educational and occupational categories used in the current analysis are similarly coded to a very broad level (e.g., 'undergraduate degree', 'professional', 'clerk').

$\lambda_{educ}\gamma_{educ}$) is a measure of the extent to which the influence of education is transmitted through occupation.

In order to identify the coefficients on education ($\gamma_{educ}$) and occupation ($\gamma_{occup}$) in equation (14), it is necessary to assume that the error term $v_i$ in equation (10) is uncorrelated with the error terms in equations (9) and (12), $e_{1i}$ and $e_{2i}$. In other words, to correctly identify the mediating role played by offspring's education and occupation in transmitting income from parent to offspring, it is necessary to make the strong identifying assumption that there are no unobserved factors influencing both parental income and offspring's outcomes. If $v_i$ and $e_{1i}$ are positively correlated, then the importance of education in intergenerational transmission will be overstated, and, furthermore, if $Cov(e_{1i}, v_i) > Cov(e_{2i}, v_i)$, then the importance of education relative to occupation will be overestimated. These identifying assumptions are discussed further in section 5.5.

### 5.2 Decomposing the NZLC estimates of the intergenerational income elasticity

This section and the next (section 5.3) decompose the NZLC and CHDS estimates of the intergenerational income elasticity, respectively. In both sections, the pathway variables used in the decomposition are first outlined, followed by the decomposition results.

#### *5.2.1 Pathway variables used in the decomposition of NZLC intergenerational income elasticity estimates*

The analysis of NZLC data in chapter 3 estimated the elasticity of offspring's individual total income with respect to their parent's individual total income to be 0.239 among son-father pairs, 0.054 among son-mother pairs, 0.145 among daughter-mother pairs, and 0.135 among daughter-father pairs. These IGE estimates are decomposed into two pathway variables drawn from the census dataset – offspring's educational attainment and offspring's occupational attainment. The census collects very few variables that could sensibly be considered transmission mechanisms for intergenerational persistence, with educational attainment and occupation standing out as the clearest candidates (however, a wider set of mechanisms drawn from the much-richer CHDS dataset are included in that decomposition in section 5.3).

Following Blanden et al. (2014) and Björklund et al. (2017), the educational attainment variable (offspring's highest qualification as at their latest census) and the occupational attainment variable (offspring's occupation as at their latest census, see Tables 8 and 9 in section 3.3.2 for the distributions of both variables) are recoded to a series of 'incremental' indicators. Educational attainment is now coded into a hierarchy of six dummy variables: 'At least 5th form/Year 11 or overseas or other school qualification',[169] 'At least 6th form/Year 12', 'At least 7th form/Year 13', 'At least trade certificate', 'At least undergraduate degree', and 'At least postgraduate degree'. The omitted reference category is 'No school qualifications'. Similarly, occupation is now coded into eight dummy variables reflecting the level one (broadest) classification of the *New Zealand*

[169] Note that the categories 'Overseas school qualification' and 'Other school qualification' have been combined with 'Qualification in 5th form/Year 11'. This is because in order to estimate the 'incremental' return of each education level (see explanation in the text), the qualification categories need to be rank ordered. Since 'Overseas' and 'Other' school qualifications have no obvious ranking in relation to the other school-level categories, the decision was taken to combine it with the 5th form/year 11 category, being the lowest category of school qualification (but not the lowest category overall, which is the omitted category 'No school qualifications').

*Standard Classification of Occupations* (NZSCO) which ranks occupations based on the level and specialisation of skills required: 'At least plant or machine operator', 'At least trades worker', 'At least agriculture/fishery worker', 'At least service/sales worker', 'At least clerk', 'At least technician/associate professional', 'At least professional', and 'At least legislator/ administrator/manager', with the omitted reference category being 'Elementary occupations' (which includes labourers and other low-skilled occupations).[170] As Blanden et al. (2014, p. 432) explain, the categorising of variables as 'at least' a particular level of education or occupation does not change the estimation of equation (7) because the other education/occupation dummies are also controlled for.[171] But it *does* change the interpretation of the coefficient estimate, which becomes an estimate of *the incremental return of that education/occupation level compared to the next lower level*, rather than that compared to the omitted education/occupation level.[172]

*5.2.2 Results of the decomposition of the NZLC intergenerational income elasticity estimates*

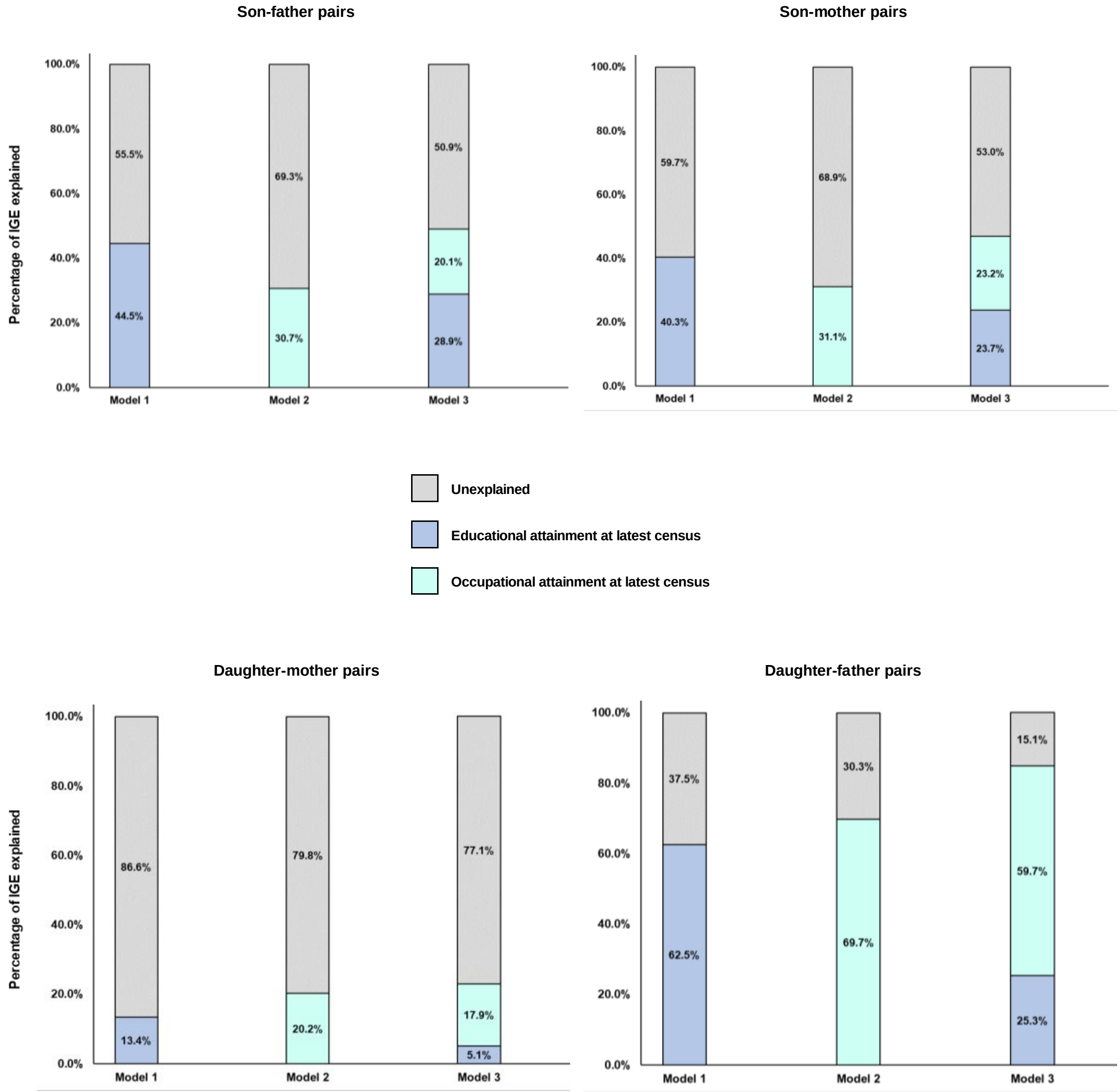
The 'big picture' of the decomposition results is presented first to show the *overall* role played by educational attainment and occupation in the intergenerational persistence of income by summing the mediating effects of the incremental dummies. This is followed by more detailed results breaking down the mediating effects of education and occupation into each of the incremental dummies they comprise. The more detailed results also show the influence of parental income on, and the income returns to, each of the incremental indicators (but only for a 'full' model in which all pathways are considered simultaneously, representing the chronological termination of the presumed sequencing of offspring pathways).

Figure 48 presents the 'big picture', showing the sequential decomposition in diagrammatic form for each offspring-parent pair, in which educational attainment (model 1) and occupation (model 2) are modelled separately as individual pathways and their mediating role is evaluated as in equation (8), and then modelled jointly (model 3) with their mediating roles evaluated as in equation (11). As noted above, in Figure 48 the mediating effects of the incremental dummies have been summed to show the *total* mediating effects of educational attainment and occupation.

Taking the decomposition for sons first, the upper panel of Figure 48 shows that the results are similar whether the father or mother is considered. Educational attainment, when considered as the sole pathway (model 1), accounts for approximately 45 percent of the son-father IGE and 40 percent of the son-mother IGE. Occupation, when considered as the sole pathway (model 2), accounts for 31 percent of both the son-father and son-mother IGE estimates. When both pathways are considered jointly (model 3), the mediating effect of education drops to 29 percent

[170] Occupations have been reverse-ranked on the NZSCO classification so that higher scores represent higher-skilled occupations.
[171] Björklund et al. (2017, p. F80), who also use incremental indicators of educational attainment in their pathway decomposition, explain that "[b]y using such decumulative variables, rather than, e.g., exclusive dummies, we avoid ambiguity in the expected association between parental income and offspring education".
[172] For example, postgraduate degree compared to undergraduate degree, rather than postgraduate degree compared to no school qualifications; the effect of a postgraduate degree compared to having no school qualifications is the sum of the coefficient estimates on all of the education dummies. When split into these incremental dummies, some specific qualifications and occupations have *negative* returns to income relative to the qualification/occupation ranked immediately below it. As a result, these particular dummy-coded pathways explain a *negative* proportion of the IGE, thus detracting from an explanation of persistence.

**Figure 48**. *Summary of the sequential decomposition of the NZLC intergenerational income elasticity estimates.*




**Notes:** The figure reports a decomposition of intergenerational income persistence estimated with NZLC data into components mediated by educational attainment alone (model 1), occupational attainment alone (model 2), and educational and occupational attainment considered jointly (model 3).

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

among son-father pairs and 24 percent among son-mother pairs, and the mediating effect of occupation drops to 20 percent among son-father pairs and 23 percent among son-mother pairs.

In total, approximately half of intergenerational income persistence between sons and their parents can be explained by sons' educational attainment and occupation, leaving 51 percent and 53 percent of the son-father and son-mother IGE estimates unexplained by these factors, respectively. Comparing model 1 with model 3 (a comparison which assumes educational attainment occurs chronologically prior to engagement in an occupation), suggests that approximately 35 to 40 percent of the mediating role of sons' education is channelled through sons' occupation. In other words, a son's educational attainment affects his income partially via its influence on the occupational choices that it opens up for him (and which he takes up). The comparison also shows that adding occupation contributes little to the overall proportion of persistence explained. The total proportion of the IGE explained is 40 to 45 percent in model 1 which considers education alone, and this increases only modestly to 47 to 49 percent when sons' occupation is also considered.

Turning to the decomposition for daughters, the lower panel of Figure 48 shows that results are strikingly different between mothers and fathers. Among daughter-mother pairs, educational attainment (model 1) accounts for 13 percent of the IGE and occupation (model 2) accounts for 20 percent of the IGE. When considered together (model 3), the mediating roles of education and occupation decrease to five percent and 18 percent, respectively. In total, 23 percent of the daughter-mother IGE estimate is explained in model 3, leaving 77 percent unexplained by these pathways. Among daughter-father pairs,[173] education (model 1) accounts for 63 percent of the IGE and occupation (model 2) accounts for 70 percent of the IGE. When considered together (model 3), the mediating roles of education and occupation decrease to 25 percent and 60 percent, respectively. In total, the vast majority (85 percent) of the daughter-father IGE can be explained by the two pathways. Comparing model 1 with model 3 (for both daughter-mother and daughter-father pairs) implies that about 60 percent of the mediating effect of daughters' education is channelled through daughters' occupation, which is a larger proportion than among sons.

There is a notable gender difference in the relative contributions of the two pathways to explaining intergenerational persistence. Among sons, education explains a larger proportion of persistence than occupation (both when considered separately and jointly),[174] but among daughters, the reverse is the case: occupation explains a larger proportion of persistence than education (when considered separately but, especially, when considered jointly). Furthermore, occupation bolsters the explained component of the IGE to a greater extent among daughters than it does among sons.

Table 21 (sons) and Table 22 (daughters) present the more detailed results for the incremental dummies that lie behind Figure 48. The tables report, for each offspring-parent pair, the mediating role of each indicator for models 1, 2, and 3, expressed as both the product-of-coefficients term shown in equation (8) and as a percentage of the IGE. In addition, the tables also report, for model 3 only, the influence of parental income on the pathway ($\lambda$) and the return to offspring's income of the pathway ($\gamma$).

Among sons, the key education indicators contributing to persistence are the increment from having no qualifications to attaining a 5$^{th}$ form/Year 11 qualification (accounting for 12 percent and 10 percent of son-father and son-mother persistence in model 3, respectively) and the shift from a trade certificate to an undergraduate degree (10 percent and eight percent, respectively), both primarily due to relatively strong incremental income returns to these shifts. The key occupation indicators contributing to persistence are the shift from clerk to technician/associate

[173] The pathway decomposition of daughter-father persistence is performed despite the fact that the estimate of the daughter-father IGE, while positive, was not statistically significant at the five percent level.
[174] Although for son-mother pairs, model 3 indicates roughly equal contributions of education and occupation in explaining persistence.

**Table 21**. *Sequential decomposition of the NZLC intergenerational income elasticity estimates for samples of sons.*

| Son-father pairs (n=4,617) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Pathways mediating son-father $\beta$** | **Model 1** | | **Model 2** | | **Model 3** | | | |
| | **Component of $\beta$** | **Percent of $\beta$** | **Component of $\beta$** | **Percent of $\beta$** | **Influence of father's income on pathway ($\lambda$)** | **Son's income return to pathway ($\gamma$)** | **Component of $\beta$** | **Percent of $\beta$** |
| At least 5th form/Year 11/overseas or other school qual. | 0.037 | 15.4% | | | 0.109 (0.013)*** | 0.262 (0.052)*** | 0.029 | 11.9% |
| At least 6th form/Year 12 | 0.028 | 11.8% | | | 0.146 (0.017)*** | 0.122 (0.052)** | 0.018 | 7.5% |
| At least 7th form/Year 13 | -0.007 | -2.8% | | | 0.141 (0.019)*** | -0.078 (0.060) | -0.011 | -4.6% |
| At least trade certificate | 0.012 | 4.9% | | | 0.116 (0.020)*** | 0.106 (0.053)* | 0.012 | 5.1% |
| At least undergraduate degree | 0.038 | 15.9% | | | 0.132 (0.016)*** | 0.173 (0.047)*** | 0.023 | 9.5% |
| At least postgraduate degree | -0.002 | -0.7% | | | 0.046 (0.008)*** | -0.025 (0.069) | -0.001 | -0.5% |
| *Educational attainment total* | | | | | | | *0.069* | *28.9%* |
| At least plant or machine operator | | | 0.008 | 3.5% | 0.024 (0.009)** | 0.359 (0.068)*** | 0.009 | 3.6% |
| At least trades worker | | | 0.008 | 3.4% | 0.085 (0.015)*** | 0.013 (0.052) | 0.001 | 0.5% |
| At least agriculture/fishery worker | | | -0.035 | -14.8% | 0.131 (0.018)*** | -0.221 (0.059)*** | -0.029 | -12.1% |
| At least service/sales worker | | | 0.011 | 4.7% | 0.148 (0.020)*** | 0.024 (0.070) | 0.004 | 1.5% |
| At least clerk | | | -0.009 | -3.6% | 0.155 (0.020)*** | -0.054 (0.079) | -0.008 | -3.5% |
| At least technician/associate professional | | | 0.059 | 24.5% | 0.163 (0.019)*** | 0.307 (0.072)*** | 0.050 | 20.8% |
| At least professional | | | 0.033 | 13.6% | 0.138 (0.018)*** | 0.137 (0.055)* | 0.019 | 7.9% |
| At least legislator/administrator/manager | | | -0.002 | -0.7% | 0.053 (0.016)** | 0.066 (0.050) | 0.004 | 1.5% |
| *Occupation total* | | | | | | | *0.048* | *20.1%* |
| Explained component of son-father $\beta$ | 0.106 | 44.5% | 0.073 | 30.7% | | | 0.118 | 49.2% |
| Unexplained component of son-father $\beta$ | 0.133 | 55.5% | 0.166 | 69.3% | | | 0.121 | 50.9% |
| Total son-father $\beta$ | 0.239 | 100.0% | 0.239 | 100.0% | | | 0.239 | 100.1% |
| **Son-mother pairs (n=14,526)** | | | | | | | | |
| **Pathways mediating son-mother $\beta$** | **Model 1** | | **Model 2** | | **Model 3** | | | |
| | **Component of $\beta$** | **Percent of $\beta$** | **Component of $\beta$** | **Percent of $\beta$** | **Influence of mother's income on pathway ($\lambda$)** | **Son's income return to pathway ($\gamma$)** | **Component of $\beta$** | **Percent of $\beta$** |
| At least 5th form/Year 11/overseas or other school qual. | 0.008 | 14.1% | | | 0.021 (0.003)*** | 0.246 (0.026)*** | 0.005 | 9.5% |
| At least 6th form/Year 12 | 0.004 | 8.2% | | | 0.030 (0.004)*** | 0.089 (0.026)** | 0.003 | 4.9% |
| At least 7th form/Year 13 | 0.000 | -0.7% | | | 0.027 (0.005)*** | -0.050 (0.031) | -0.001 | -2.5% |
| At least trade certificate | 0.001 | 2.4% | | | 0.028 (0.005)*** | 0.079 (0.028)** | 0.002 | 4.0% |
| At least undergraduate degree | 0.009 | 16.6% | | | 0.030 (0.004)*** | 0.151 (0.023)*** | 0.005 | 8.4% |
| At least postgraduate degree | 0.000 | -0.4% | | | 0.010 (0.002)*** | -0.032 (0.032) | 0.000 | -0.6% |
| *Educational attainment total* | | | | | | | *0.013* | *23.7%* |
| At least plant or machine operator | | | 0.004 | 6.6% | 0.012 (0.003)*** | 0.296 (0.034)*** | 0.003 | 6.5% |
| At least trades worker | | | 0.002 | 4.2% | 0.027 (0.004)*** | -0.007 (0.026) | 0.000 | -0.4% |
| At least agriculture/fishery worker | | | -0.005 | -9.0% | 0.034 (0.005)*** | -0.090 (0.030)** | -0.003 | -5.6% |
| At least service/sales worker | | | 0.002 | 4.0% | 0.025 (0.005)*** | 0.038 (0.034) | 0.001 | 1.7% |
| At least clerk | | | -0.002 | -3.3% | 0.024 (0.005)*** | -0.074 (0.037)* | -0.002 | -3.3% |
| At least technician/associate professional | | | 0.009 | 16.7% | 0.032 (0.005)*** | 0.237 (0.035)*** | 0.008 | 14.2% |
| At least professional | | | 0.007 | 12.1% | 0.026 (0.005)*** | 0.170 (0.026)*** | 0.004 | 8.2% |
| At least legislator/administrator/manager | | | 0.000 | -0.2% | 0.016 (0.004)*** | 0.070 (0.024)** | 0.001 | 2.1% |
| *Occupation total* | | | | | | | *0.013* | *23.2%* |
| Explained component of son-mother $\beta$ | 0.022 | 40.3% | 0.017 | 31.1% | | | 0.025 | 47.0% |
| Unexplained component of son-mother $\beta$ | 0.032 | 59.7% | 0.037 | 68.9% | | | 0.029 | 53.0% |
| Total son-mother $\beta$ | 0.054 | 100.0% | 0.054 | 100.0% | | | 0.054 | 100.0% |

**Symbols**: ***$p$<0.001 **$p$<0.01 *$p$<0.05
**Notes**: The omitted reference category for educational attainment is 'No school qualifications' and for occupation is 'Elementary occupations'. Robust standard errors in parentheses.
**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

**Table 22**. Sequential decomposition of the NZLC intergenerational income elasticity estimates for samples of daughters.

| Daughter-mother pairs (n=9,312) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Pathways mediating daughter-mother $\beta$** | **Model 1** | | **Model 2** | | **Model 3** | | | |
| | **Component of $\beta$** | **Percent of $\beta$** | **Component of $\beta$** | **Percent of $\beta$** | **Influence of mother's income on pathway ($\lambda$)** | **Daughter's income return to pathway ($\gamma$)** | **Component of $\beta$** | **Percent of $\beta$** |
| At least 5th form/Year 11/overseas or other school qual. | 0.005 | 3.5% | | | 0.015 (0.004)*** | 0.035 (0.084) | 0.001 | 0.4% |
| At least 6th form/Year 12 | 0.002 | 1.6% | | | 0.022 (0.005)*** | 0.065 (0.066) | 0.001 | 1.0% |
| At least 7th form/Year 13 | -0.007 | -4.8% | | | 0.031 (0.006)*** | -0.304 (0.096)** | -0.010 | -6.6% |
| At least trade certificate | 0.011 | 7.5% | | | 0.035 (0.006)*** | 0.315 (0.095)** | 0.011 | 7.6% |
| At least undergraduate degree | 0.007 | 5.1% | | | 0.023 (0.005)*** | 0.144 (0.066)* | 0.003 | 2.3% |
| At least postgraduate degree | 0.001 | 0.4% | | | 0.004 (0.003) | 0.140 (0.091) | 0.001 | 0.4% |
| *Educational attainment total* | | | | | | | *0.007* | *5.1%* |
| At least plant or machine operator | | | 0.000 | 0.1% | 0.007 (0.002)** | 0.016 (0.197) | 0.000 | 0.1% |
| At least trades worker | | | 0.001 | 1.0% | 0.010 (0.003)** | 0.110 (0.267) | 0.001 | 0.8% |
| At least agriculture/fishery worker | | | 0.002 | 1.4% | 0.010 (0.003)** | 0.194 (0.238) | 0.002 | 1.4% |
| At least service/sales worker | | | -0.003 | -1.8% | 0.014 (0.004)*** | -0.185 (0.123) | -0.003 | -1.8% |
| At least clerk | | | 0.005 | 3.7% | 0.027 (0.005)*** | 0.190 (0.071)** | 0.005 | 3.6% |
| At least technician/associate professional | | | 0.005 | 3.4% | 0.031 (0.006)*** | 0.137 (0.067)* | 0.004 | 2.9% |
| At least professional | | | 0.015 | 10.3% | 0.040 (0.006)*** | 0.260 (0.070)*** | 0.010 | 7.2% |
| At least legislator/administrator/manager | | | 0.003 | 2.2% | 0.022 (0.004)*** | 0.244 (0.068)*** | 0.005 | 3.7% |
| *Occupation total* | | | | | | | *0.026* | *17.9%* |
| Explained component of daughter-mother $\beta$ | 0.019 | 13.4% | 0.029 | 20.2% | | | 0.033 | 22.9% |
| Unexplained component of daughter-mother $\beta$ | 0.126 | 86.6% | 0.116 | 79.8% | | | 0.112 | 77.1% |
| Total daughter-mother $\beta$ | 0.145 | 100.0% | 0.145 | 100.0% | | | 0.145 | 100.0% |
| **Daughter-father pairs (n=1,944)** | | | | | | | | |
| **Pathways mediating daughter-father $\beta$** | **Model 1** | | **Model 2** | | **Model 3** | | | |
| | **Component of $\beta$** | **Percent of $\beta$** | **Component of $\beta$** | **Percent of $\beta$** | **Influence of father's income on pathway ($\lambda$)** | **Daughter's income return to pathway ($\gamma$)** | **Component of $\beta$** | **Percent of $\beta$** |
| At least 5th form/Year 11/overseas or other school qual. | 0.044 | 32.5% | | | 0.085 (0.019)*** | 0.027 (0.213) | 0.002 | 1.7% |
| At least 6th form/Year 12 | -0.025 | -18.6% | | | 0.109 (0.025)*** | -0.080 (0.169) | -0.009 | -6.4% |
| At least 7th form/Year 13 | -0.011 | -7.9% | | | 0.098 (0.029)** | -0.258 (0.215) | -0.025 | -18.8% |
| At least trade certificate | 0.026 | 19.3% | | | 0.075 (0.029)* | 0.398 (0.211) | 0.030 | 22.0% |
| At least undergraduate degree | 0.056 | 41.2% | | | 0.114 (0.025)*** | 0.358 (0.172)* | 0.041 | 30.4% |
| At least postgraduate degree | -0.005 | -4.0% | | | 0.035 (0.014)* | -0.139 (0.253) | -0.005 | -3.6% |
| *Educational attainment total* | | | | | | | *0.034* | *25.3%* |
| At least plant or machine operator | | | 0.016 | 11.9% | 0.036 (0.014)* | 0.476 (0.421) | 0.017 | 12.8% |
| At least trades worker | | | -0.069 | -51.5% | 0.048 (0.016)** | -1.464 (0.598)* | -0.070 | -52.2% |
| At least agriculture/fishery worker | | | 0.082 | 60.6% | 0.046 (0.017)** | 1.746 (0.563)** | 0.080 | 59.6% |
| At least service/sales worker | | | -0.023 | -16.8% | 0.048 (0.022)* | -0.463 (0.297) | -0.022 | -16.4% |
| At least clerk | | | 0.015 | 10.9% | 0.072 (0.026)** | 0.226 (0.179) | 0.016 | 12.0% |
| At least technician/associate professional | | | 0.013 | 9.7% | 0.103 (0.029)*** | 0.086 (0.169) | 0.009 | 6.5% |
| At least professional | | | 0.031 | 23.3% | 0.123 (0.028)*** | 0.051 (0.183) | 0.006 | 4.7% |
| At least legislator/administrator/manager | | | 0.029 | 21.6% | 0.084 (0.023)*** | 0.525 (0.178)** | 0.044 | 32.6% |
| *Occupation total* | | | | | | | *0.080* | *59.7%* |
| Explained component of daughter-father $\beta$ | 0.084 | 62.5% | 0.094 | 69.7% | | | 0.114 | 84.7% |
| Unexplained component of daughter-father $\beta$ | 0.051 | 37.5% | 0.041 | 30.3% | | | 0.021 | 15.1% |
| Total daughter-father $\beta$ | 0.135 | 100.0% | 0.135 | 100.0% | | | 0.135 | 100.0% |

**Symbols**: ***$p$<0.001 **$p$<0.01 *$p$<0.05
**Notes**: The omitted reference category for educational attainment is 'No school qualifications' and for occupation is 'Elementary occupations'. Robust standard errors in parentheses.
**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

professional (21 percent and 14 percent for son-father and son-mother persistence in model 3, respectively) and the next increment upwards from technician to professional (eight percent for both son-parent pairs), primarily due to strong income returns to these shifts, but also, in the case of the former, a relatively strong influence of parents' income.[175]

Among daughters, results must be discussed separately for mothers versus fathers given the substantial differences. Among daughter-mother pairs, the key indicators are the increment from a 7th form/Year 13 qualification to a trades certificate (eight percent of persistence in model 3) and the shift from technician/associate professional to professional (seven percent of persistence), both due to a combination of relatively strong parental influence and strong incremental returns to income. Among daughter-father pairs, the standout education indicators are 'At least trade certificate' and 'At least undergraduate degree' (22 percent and 30 percent of persistence from model 3, respectively). The key occupation dummies are 'At least agriculture/fishery worker', 'At least clerk', and 'At least legislator/administrator/manager' (60 percent, 12 percent, and 33 percent of persistence, respectively), mostly driven by strong returns to income.

In summary, about half of intergenerational persistence among sons can be accounted for by his educational and occupational attainment (whether in relation to his father or mother). For daughters, about a quarter of persistence from mothers can be similarly accounted for, with occupation playing a greater role than education. Daughter-father persistence, while not a statistically significant relationship, can nonetheless be largely accounted for by daughter's education but especially her occupation. In general, education plays a greater role than occupation in intergenerational income persistence for sons, while for daughters, occupation plays a greater role than education. Across all offspring-parent pairs, a comparison of the estimates of the link between parental income and offspring's education and occupation and the return to offspring's income of the attainment indicators suggests that the former – statistically significant in all cases but one – is the main driver of intergenerational persistence, as many pathway indicators have non-significant and/or negative returns to income.

## 5.3 Decomposing the CHDS all offspring estimate of the intergenerational income elasticity

### *5.3.1 Pathway variables used in the decomposition of the CHDS intergenerational income elasticity estimates*

The analysis of CHDS data in chapter 4 estimated the elasticity of offspring's lifetime average earnings with respect to their family income in childhood to be 0.479. Because this 'all offspring' sample pools sons and daughters, offspring's sex is added to all equations in the decomposition

[175] Also noteworthy are the shifts from a 6th form to 7th form qualification and from trades worker to agriculture/fishery worker, both of which have negative incremental return to sons' income and thus account for negative proportions of the son-father and son-mother IGEs, thereby detracting from an explanation of persistence. Using 1981, 1986, 1991, and 1996 census microdata, Maani (1999, p. 12) similarly finds that "the returns to Bursary [7th form qualification] are unusually low, and at negative rates when compared to…Sixth Form Certificate" and that individuals whose highest qualification was a Bursary had a substantially lower labour force participation rate and higher unemployment rate than all other qualification levels except 'No qualifications'. These results raise questions "as to whether those with a Bursary for their highest qualification are at a disadvantage in comparison to those who commence work after the Sixth Form (Year 12), since the Seventh Form (Year 13) Bursary degree is designed as preparation for higher education" (Maani, 1999, p. 12), a higher education which was ultimately not pursued by the time the 7th-form-qualified sons and daughters in the NZLC samples reached their thirties or forties.

(equations (9) through (14)). The decomposition considers six pathways drawn from the extensive CHDS longitudinal dataset spanning birth to adulthood. These are two childhood non-cognitive behavioural traits (conduct problems and anxiety problems in middle childhood), two childhood cognitive ability measures (IQ score in middle childhood and reading ability in adolescence), plus educational and occupational attainment in adulthood.

When cohort members were aged 7, 8, and 9 years, symptoms of conduct disorder (aggressive, antisocial, oppositional, defiant behaviour) and anxiety/withdrawal (anxious, fearful, shy, sad or withdrawn behaviour) were measured with items drawn from behavioural questionnaires by Rutter, Tizard and Whitmore (1970) and Conners (1969, 1970). These questionnaires ask respondents to rate current symptoms of behavioural and emotional difficulties in children and were obtained from both mothers and teachers of the CHDS cohort. The CHDS researchers used responses to these behavioural questions to construct scales of conduct disorder and anxiety disorder, selecting items from the questionnaires that broadly conformed to diagnostic criteria for these disorders.[176] Higher scores on the conduct disorder scale correspond to the child exhibiting more symptoms of conduct problems. Higher scores on the anxiety disorder scale correspond to the child exhibiting more symptoms of anxiety and withdrawal problems.[177] Conduct and anxiety problems are hypothesised to be negatively related to parental income and have a negative return to offspring's income, thus contributing to intergenerational persistence of disadvantage.

At ages 8 and 9 years, cohort members' IQ was measured with the Revised Wechsler Intelligence Scale for Children (WISC-R; Wechsler, 1974). At each age, children were administered four verbal subtests (information, similarities, arithmetic, vocabulary) and four performance subtests (picture completion, block design, object assembly, coding). Scores on these subtests were combined to produce a total IQ score at ages 8 and 9, such that higher scores correspond to higher IQ. For analytical purposes, CHDS researchers frequently use an average of the IQ scores over the two administrations with missing scores imputed, and it is this average IQ score over ages 8 and 9 that is used in the current analysis.[178] IQ scores are hypothesised to be positively related to parental income and have a positive return to offspring's income.

At age 18, cohort members' reading ability was assessed with the New Zealand revision of the Burt Word Reading Test (Gilmore, Croft, & Reid, 1981), a word recognition test. The same test had also been administered annually between ages 8 to 13. Higher scores on this test correspond to higher reading ability. Of the 771 offspring in the 'all offspring' sample, 28 had missing Burt Reading Test scores at age 18. These offspring were therefore assigned their score at age 13, 12, 11, 10, 9, or 8 (in that order until a valid score was available). Given that reading ability

[176] For each set of items in each year, single factor models were fitted to the items using structural equation modelling. See Fergusson, Horwood, and Lloyd (1991) for a list of the conduct disorder items and the results of the single factor modelling, and Fergusson and Horwood (1993) for the anxiety items and modelling results.

[177] Conduct disorder symptoms are a measure of 'externalising' behavioural problems (aggressive behaviour directed towards other children) and anxiety symptoms are a measure of 'internalising' behavioural problems (emotionally reactive, anxious, and withdrawn behaviours directed mainly inwards at oneself). Blanden et al. (2007), who use data from the British Cohort Study to estimate and decompose intergenerational income persistence, also use scales of externalising and internalising behaviours as mediating variables, and these are derived from the same Rutter et al. (1970) questionnaire utilised in the CHDS.

[178] See Fergusson, Horwood and Ridder (2005) for evidence of high reliability of the total IQ scores at each age and for details on imputation of IQ scores for children living outside of the Canterbury region at ages 8 and 9 who were not administered the WISC-R.

increases over age, reading test scores at all ages were standardised (that is, scaled to have a mean of zero and standard deviation of one). Reading ability is expected to be positively related to parental income and have a positive return to offspring's income.

Educational attainment is measured by the highest educational qualification obtained by age 40 and occupation is measured at age 40. Offspring with missing education and occupation data at age 40 (including not being in the labour force at that age) are assigned their highest qualification and their occupation at ages 35 or 30 (in that order).[179] Of the 771 offspring in the sample, only five percent gained higher qualifications at age 35 or 40, meaning the educational attainment variable effectively measures attainment by age 30. As in the NZLC decomposition, the education and occupation variables are recoded into a hierarchy of dummy variables. For educational attainment, they are 'At least high school qualification', 'At least tertiary level 5 or 6 qualification, 'At least undergraduate degree', and 'At least postgraduate degree', with the omitted reference category being 'No formal educational qualifications'. Occupation is coded into eight dummy variables reflecting the level one (broadest) classification of the *Australian and New Zealand Standard Classification of Occupations* (ANZSCO) *2006*: 'At least machine operator or driver', 'At least sales worker', 'At least clerical and administrative worker', 'At least community and personal service worker', 'At least technician and trades worker', 'At least professional', and 'At least manager', with the omitted reference category being 'Labourers'. The coefficients on these dummy variables thus become estimates of the incremental return of that education/occupation level compared to the next lower level, rather than that compared to the omitted education/occupation level.

Of the 771 offspring in the sample, 763 had complete data on all the pathway variables. The IGE estimate with this marginally reduced sample is the same to three decimal places as the estimate with the full sample (perhaps not surprising given the small size of the sample loss). Table 23 presents descriptive statistics of the six pathway variables for the 763 offspring included in the decomposition.

### *5.3.2 Results of the decomposition of the CHDS 'all offspring' intergenerational income elasticity estimate*

The 'big picture' of the decomposition results is again presented first, in which the mediating effects of the education and occupation indicators have been summed to show the *total* mediating role of educational attainment and occupation. This is followed by more detailed results breaking down the mediating effects of education and occupation into each of the incremental dummies they comprise. As with the NZLC decomposition, the more detailed results also show the influence of parental income on, and the earnings return to, each of the incremental indicators (but only for a 'full' model in which all pathways are considered simultaneously).

[179] One daughter reported earnings at age 40 but did not report an occupation at that age, nor at ages 35 and 30 (she was classified as 'Not in paid employment' at all three ages). For the purposes of this decomposition, this daughter was coded to the lowest-ranking occupational category, 'Labourers'.

**Table 23**. *Descriptive statistics of the pathway variables used in the decomposition of the CHDS intergenerational income elasticity.*

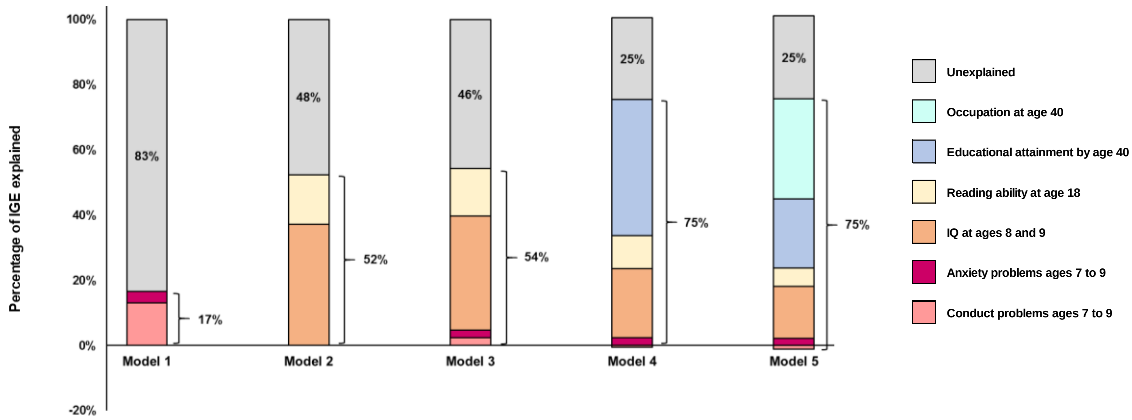
| CHDS all offspring sample with valid pathway data (n=763) | | | | | |
|---|---|---|---|---|---|
| **Continuous pathway variables** | **Minimum** | **Median** | **Maximum** | **Mean** | **SD** |
| *Non-cognitive traits:* | | | | | |
| Conduct problems (averaged over ages 7 to 9) | 41.00 | 47.00 | 85.00 | 49.45 | 7.00 |
| Anxiety problems (averaged over ages 7 to 9) | 20.00 | 25.00 | 62.00 | 26.15 | 4.45 |
| *Cognitive ability:* | | | | | |
| IQ score (averaged over ages 8 to 9) | 45.50 | 104.50 | 144.50 | 104.48 | 14.08 |
| Reading ability (standardised) (at age 18) | -5.51 | 0.28 | 1.18 | 0.00 | 1.01 |
| | | | | | |
| **Categorical pathway variables** | **n** | **%** | | | |
| *Educational attainment by age 40:* | | | | | |
| No formal qualification | 34 | 4.5 | | | |
| Secondary or tertiary level 4 qualification | 371 | 48.6 | | | |
| Tertiary level 5 or 6 qualification | 94 | 12.3 | | | |
| Bachelors degree | 206 | 27.0 | | | |
| Postgraduate degree | 58 | 7.6 | | | |
| *Occupational attainment at age 40:* | | | | | |
| Managers | 183 | 24.0 | | | |
| Professionals | 205 | 26.9 | | | |
| Technicians and trades workers | 105 | 13.8 | | | |
| Community and personal service workers | 68 | 8.9 | | | |
| Clerical and administrative workers | 80 | 10.5 | | | |
| Sales workers | 28 | 3.7 | | | |
| Machinery operators and drivers | 40 | 5.2 | | | |
| Labourers | 54 | 7.1 | | | |

**Notes**: Reading ability is measured with standardised scores from the Burt Word Reading Test administered at age 18. Of the 28 offspring with missing scores at age 18, 21 are assigned standardised scores on the same test administered at ages 13, 12, 11, 10, 9, or 8 years of age (in that order until a valid score was available).

**Source**: Author's calculations from Christchurch Health and Development Study dataset.

Figure 49 presents the sequential decomposition in diagrammatic form. In model 1, only the two measures of non-cognitive behavioural traits are included. In model 2, only the two measures of cognitive ability are included. In model 3, the non-cognitive and cognitive measures are considered jointly. In model 4, educational attainment by age 40 is added as a pathway, and finally in model 5, occupation at age 40 is added (shown in Figure 49 as the sum of their constituent dummies). This sequencing of pathways broadly reflects the chronological order in which the set of pathway variables was collected from the cohort and is designed to roughly approximate the causal chain of events often assumed in studies that sequentially decompose intergenerational persistence (i.e., that childhood non-cognitive traits and cognitive skills influence educational attainment which in turn influences occupational attainment, as in Blanden et al. (2007) and Blanden et al. (2014)). Figure 49 presents, for each model, the proportion of the IGE explained and the proportion unexplained by the set of mediating variables included in the model (in total across the set). More detailed results of the proportions explained by each of the individual variables within each set are reported in Table 24 below.

**Figure 49**. *Summary of the sequential decomposition of the CHDS 'all offspring' intergenerational income elasticity estimate.*




**Notes:** The figure reports a decomposition of intergenerational income persistence estimated with CHDS data into components mediated by the non-cognitive traits of conduct problems and anxiety problems (model 1), the cognitive abilities of IQ and reading ability (model 2), non-cognitive traits and cognitive abilities considered jointly (model 3), non-cognitive traits, cognitive abilities, and educational attainment (model 4), and non-cognitive traits, cognitive abilities, educational attainment, and occupational attainment (model 5).
**Source**: Author's calculations from Christchurch Health and Development Study dataset.

In model 1, only the two non-cognitive traits (conduct and anxiety problems at ages 7 to 9) are included as mediating pathways. At 13.2 percent of the IGE, conduct problems play a larger mediating role than anxiety problems (3.4 percent; see Table 24 below). Together they account for about 17 percent of intergenerational persistence, leaving 83 percent of persistence unexplained.

In model 2, only the two cognitive ability variables (IQ at ages 8 and 9 and reading ability at age 18) are included. IQ accounts for 37.2 percent and reading ability accounts for 15.2 percent of persistence, together accounting for about half (52 percent) of persistence. A comparison of the results from models 1 and 2 suggests cognitive abilities are a much more important conduit for intergenerational income persistence than non-cognitive traits.

In model 3, the non-cognitive and cognitive pathways are considered jointly. Compared to model 1, the mediating effect of conduct problems is notably reduced to 2.5 percent and that of anxiety problems falls to 2.1 percent of the IGE. On the other hand, the explanatory contributions of the cognitive variables (35.0 percent for IQ, 14.6 percent for reading ability) are only marginally changed from model 2. These results imply that the mediating effects of non-cognitive behavioural traits are largely channelled through their impact on cognitive skills. Altogether, the non-cognitive and cognitive variables account for 54 percent of persistence.

In model 4, educational attainment by age 40 is added as a pathway, and it makes a substantial explanatory contribution, accounting for 41.8 percent of persistence (when summed across its incremental dummies). As might be expected, the influence of the cognitive variables declines to 21.2 percent for IQ and 10.0 percent for reading ability, indicating that cognitive skills affect earnings partially via their influence on educational attainment. What is perhaps surprising is the size of the independent mediating effect that cognitive ability, especially IQ, continues to exert even after adjusting for educational attainment, a result discussed further below. The role of anxiety problems increases marginally to 2.4 percent, while that of conduct problems decreases to -0.4 percent, indicating that the remaining effect of conduct problems (that which had not already been channelled through the cognitive variables) is channelled through educational attainment (and there is a small residual negative effect on persistence, which arises because conduct problems have a small positive return to earnings after adjustment for the other covariates in model 4, thus driving mobility to a small extent).[180] Collectively, the non-cognitive, cognitive, and educational variables account for 75 percent of persistence. For comparative purposes, a separate model with educational attainment as the only mediator was also estimated (not shown in Figure 49 or reported in Table 24), analogous to many other studies. In this model, educational attainment accounts for 61 percent of persistence.

In model 5, occupational attainment at age 40 is added as a pathway, and it explains 30.7 percent of persistence (when summed across its dummies). The explanatory contribution of educational attainment decreases to 21.1 percent, implying that about half of the mediating effect of education is channelled through occupational attainment, which is to be expected given that ANZSCO ranks occupations based on skill level. The mediating effect of IQ declines to 16.0 percent, while that of reading ability is halved to 5.6 percent, again implying that the effects of these cognitive abilities are partially channelled through occupational attainment. The mediating effect of anxiety problems is virtually unchanged from model 4 (and indeed model 3), while that of conduct problems is pushed (slightly) further negative. In total, the pathways in model 5 account for 75 percent of persistence, which is unchanged from model 4. This suggests that occupational attainment exerts no independent effect in mediating intergenerational persistence over and above its role in channelling educational attainment, cognitive skills, and non-cognitive traits. In other words, occupation acts as a transmission mechanism linking parents' incomes with their offspring's earnings insofar as it manifests the transmission effects of offspring outcomes presumed to occur earlier in life.

Table 24 present the more detailed results that lie behind Figure 49. The table reports, for each model, the mediating role of each pathway including all the incremental dummies, expressed as both the product-of-coefficients term shown in equation (8) and as a percentage of the IGE. In addition, the table also reports, for model 5 only, the influence of parental income on the pathway ($\lambda$) and the return to offspring's earnings of the pathway ($\gamma$). In the decompositions incorporating educational attainment as a pathway (models 4 and 5), all the incremental educational dummies contribute to persistence, but attaining a level 5 or 6 tertiary qualification (many of which will be a trade or technical qualification) and an undergraduate degree make the largest contributions. In model 4 (prior to adding occupation), a pre-degree tertiary qualification accounts for 14.5 percent of persistence and an undergraduate degree accounts for 17.0 percent of persistence. In model 5, all the incremental occupational dummies are conduits of persistence except for 'at least sales worker' and 'at least community or personal service worker' which account for negative proportions of the IGE (-7.9 percent and -3.5 percent, respectively, both owing to negative incremental returns to offspring's earnings) and therefore are conduits of mobility.

[180] Using data from the British Cohort Study tracking a group of children born in Great Britain in 1970, Knapp et al. (2011) find that sons' conduct problems at age 10 are associated with higher earnings at age 30 *if employed* (but also a higher probability of unemployment). On the other hand, attention deficit and anxiety problems in childhood are both associated with lower earnings at age 30.

**Table 24**. *Sequential decomposition of the CHDS intergenerational income elasticity estimates for all offspring sample.*

| Offspring-parent pairs (n=763)[1] | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | | | |
| Pathways mediating offspring-parent $\beta$ | Component of $\beta$ | Percent of $\beta$ | Component of $\beta$ | Percent of $\beta$ | Component of $\beta$ | Percent of $\beta$ | Component of $\beta$ | Percent of $\beta$ | Influence of parents' income on pathway ($\lambda$) | Offspring's earnings return to pathway ($\gamma$) | Component of $\beta$ | Percent of $\beta$ |
| *Non-cognitive traits:* | | | | | | | | | | | | |
| Conduct problems (averaged over ages 7 to 9) | 0.063 | 13.2% | | | 0.012 | 2.5% | -0.002 | -0.4% | -4.968 (0.760)*** | 0.001 (0.004) | -0.005 | -1.0% |
| Anxiety problems (averaged over ages 7 to 9) | 0.016 | 3.4% | | | 0.010 | 2.1% | 0.012 | 2.4% | -1.232 (0.520)* | -0.009 (0.006) | 0.011 | 2.3% |
| *Cognitive ability:* | | | | | | | | | | | | |
| IQ score (averaged over ages 8 to 9) | | | 0.178 | 37.2% | 0.168 | 35.0% | 0.102 | 21.2% | 15.540 (1.399)*** | 0.005 (0.002)* | 0.077 | 16.0% |
| Reading ability (at age 18)[2] | | | 0.073 | 15.2% | 0.070 | 14.6% | 0.048 | 10.0% | 0.819 (0.105)*** | 0.033 (0.032) | 0.027 | 5.6% |
| *Educational attainment by age 40:*[3] | | | | | | | | | | | | |
| At least high school qualification | | | | | | | 0.014 | 3.0% | 0.128 (0.022)*** | 0.049 (0.130) | 0.006 | 1.3% |
| At least tertiary level 5 or 6 qualification | | | | | | | 0.069 | 14.5% | 0.513 (0.050)*** | 0.080 (0.081) | 0.041 | 8.6% |
| At least undergraduate degree | | | | | | | 0.082 | 17.0% | 0.517 (0.047)*** | 0.052 (0.088) | 0.027 | 5.6% |
| At least postgraduate degree | | | | | | | 0.035 | 7.4% | 0.178 (0.028)*** | 0.152 (0.102) | 0.027 | 5.6% |
| *Educational attainment total* | | | | | | | *0.201* | *41.8%* | | | *0.101* | *21.1%* |
| *Occupation at age 40:*[4] | | | | | | | | | | | | |
| At least machine operator or driver | | | | | | | | | 0.129 (0.027)*** | 0.243 (0.143) | 0.031 | 6.5% |
| At least sales worker | | | | | | | | | 0.187 (0.035)*** | -0.203 (0.171) | -0.038 | -7.9% |
| At least clerical or administrative worker | | | | | | | | | 0.191 (0.039)*** | 0.291 (0.149)* | 0.056 | 11.6% |
| At least community or personal service worker | | | | | | | | | 0.183 (0.047)*** | -0.092 (0.113) | -0.017 | -3.5% |
| At least technician or trade worker | | | | | | | | | 0.232 (0.050)*** | 0.124 (0.107) | 0.029 | 6.0% |
| At least professional | | | | | | | | | 0.280 (0.053)*** | 0.304 (0.091)** | 0.085 | 17.8% |
| At least manager | | | | | | | | | 0.039 (0.046) | 0.033 (0.073) | 0.001 | 0.3% |
| *Occupation total* | | | | | | | | | | | *0.147* | *30.7%* |
| Explained component of offspring-parent $\beta$ | 0.080 | 16.6% | 0.251 | 52.4% | 0.260 | 54.3% | 0.360 | 75.0% | | | 0.358 | 74.6% |
| Unexplained component of offspring-parent $\beta$ | 0.400 | 83.4% | 0.228 | 47.6% | 0.219 | 45.7% | 0.120 | 25.0% | | | 0.122 | 25.4% |
| Total offspring-parent $\beta$ | 0.479 | 100.0% | 0.479 | 100.0% | 0.479 | 100.0% | 0.479 | 100.0% | | | 0.479 | 100.0% |

**Symbols:** ***$p$<0.001 **$p$<0.01 *$p$<0.05

**Notes**:
[1] All regressions control for offspring's sex. Robust standard errors in parentheses.
[2] Reading ability is measured with standardised scores from the Burt Word Reading Test administered at age 18. Of the 28 offspring with missing scores at age 18, 21 are assigned standardised scores on the same test administered at ages 13, 12, 11, 10, 9, or 8 years of age. The remaining seven offspring have missing scores at all ages and are excluded from the decomposition.
[3] The omitted reference category for educational attainment is 'No formal qualifications'. Where educational attainment is missing at age 40, education at ages 35 or 30 are used instead.
[4] The omitted reference category for occupation is 'Labourers'. Where occupation is missing at age 40, occupation at ages 35 or 30 are used instead.

**Source**: Author's calculations from Christchurch Health and Development Study dataset.

Being a clerical or administrative worker and being a professional account for the largest proportions of persistence, at 12 percent and 18 percent, respectively, owing to relatively strong links with parental income and significant returns to offspring's earnings. Comparing the estimates of $\lambda$ and $\gamma$ suggests that the link between parental income and offspring's characteristics is the main driver of intergenerational persistence, as most characteristics have non-significant (and often small) returns to earnings.

In summary, three-quarters of intergenerational income persistence among CHDS offspring can be accounted for by offspring's non-cognitive traits, cognitive skills, and educational attainment. Adding offspring's occupation does not increase the total proportion of persistence explained and non-cognitive traits make a relatively small contribution to persistence. Thus, cognitive skills in childhood and educational attainment stand out as the key mechanisms linking parents' and offspring's incomes. Furthermore, the link appears to be primarily driven by the gradient between parental income and offspring's cognitive skills and educational attainment. Thus, the initial seeds of intergenerational income persistence appear to be sown early in the offspring's life and these are cultivated throughout their educational career with cumulative flow-on effects into adulthood.

### 5.4 Comparison of decomposition results with results from other studies

Several studies, including one for New Zealand, examine the role that various pathway factors play in mediating intergenerational income persistence. Gibbons (2010) uses two different methods to quantify the role of educational attainment as a mechanism through which parental income affects offspring's income in New Zealand. Using data from the Dunedin Multidisciplinary Health and Development Study with sons and daughters pooled, he estimates the IGE of offspring's log total income at age 32 with respect to their parents' log family income during the offspring's childhood, controlling for offspring's sex, to be 0.272. When Gibbons then adds offspring's highest educational qualification to the regression equation, the estimated IGE drops to 0.144, indicating that "on average about 47% of the effects of family background on income were mediated through educational qualifications, and about 53% occurred because of other influences [the unexplained component of the IGE estimate]" (Gibbons, 2010, p. 23). As an alternative, he also uses the decomposition technique utilised in this chapter to decompose the IGE into a component channelled through offspring's educational attainment – this time measured as years of education – and finds that "about half [48 percent] the intergenerational income effect may occur because children from better-off families tend to continue their education for longer than children from less well-off families, and that about half is attributable to other factors", which is an almost identical result to his first method (Gibbons, 2010, p. 23).

Gibbons' (2010) result can be compared to analogous results from the NZLC and CHDS decompositions reported here. However, caution is required when making such comparisons. The NZLC decomposition analyses sons and daughters separately and uses an individual parent's income as the regressor, whereas the estimate that Gibbons decomposes pools sons and daughters and uses family income pooled across mother and father. The most comparable result is the CHDS model with educational attainment only (reported in section 5.3.2), which pools sons with daughters and mothers with fathers as in Gibbons. In the NZLC decomposition, educational attainment accounts for 45 percent and 40 percent of the son-father and son-mother IGE estimates, respectively, which are broadly similar to Gibbons' result. However, analogous results for daughters (13 percent and 63 percent of the daughter-mother and daughter-father IGE estimates, respectively) differ markedly from Gibbons' result. In the CHDS decomposition, 61 percent of the IGE can be accounted for by educational attainment considered on its own, which is moderately higher than Gibbons' result.

Discussion now turns to the most comparable pieces of research from the international literature, each discussed in turn, before summarising the relevant patterns that emerge. Mood et al. (2012) use administrative data to estimate the role of cognitive ability, non-cognitive personality traits, physical characteristics (height, body mass index, fitness), and educational attainment in mediating income persistence among son-father pairs in Sweden.[181] When four cognitive, four non-cognitive, and three physical variables are considered jointly, they find that cognitive ability accounts for 20 percent, non-cognitive traits account for 13 percent, and physical characteristics account for four percent of the intergenerational correlation (IGC). In total, all three sets of variables collectively account for 37 percent of the IGC. When sons' years of education is then added as a mediator, it accounts for 33 percent of the IGC and the role of the cognitive variables is halved (down to 10 percent of the IGC), the role of non-cognitive traits is barely changed (12 percent), and the role of physical characteristics is reduced to about two percent. In total, all variables collectively account for 57 percent of the IGC. Mood et al. (2012, pp. 72, 74) note that, "[p]redictably, part of the cognitive ability component appears to be mediated by education, but the contribution of cognitive ability to the income correlation remains large even after controlling for education", while non-cognitive traits are only mediated "to a small extent through education" implying that such traits "mainly secure their income returns directly in the labour market". Mood et al. (2012, p. 72) summarise the results of their decomposition thus:

> Cognitive ability…contributes strongly to the intergenerational transmission of income, and part of this impact stems from its effect on education. However, a non-negligible proportion of the income correlation is accounted for by cognitive ability even after controlling for education, and personality traits account for a substantial share of the income correlation. The physiological characteristics analyzed here, on the other hand, do not contribute much to [intergenerational persistence].

These findings from Mood et al. (2012) are consistent with results from the CHDS decomposition, which also finds that offspring's cognitive ability contributes to intergenerational persistence partially via its influence on educational attainment but also independently of educational attainment (however, the size of the mediating effect is much larger in the CHDS results and Mood et al. analyse son-father pairs whereas the CHDS pools sons with daughters and mothers

[181] Persistence is measured with the intergenerational correlation in incomes. Data on cognitive ability, non-cognitive traits, and physical characteristics are drawn from Swedish military enlistment records. Cognitive ability is measured by tests of logic-inductive reasoning, verbal comprehension, spatial ability, and technical understanding, the test scores from which are summed to provide an overall measure of cognitive ability. Data on non-cognitive traits, drawn from an interview with a psychologist designed to assess suitability for military service, are measured by psychologists' ratings of social maturity (having friends, taking responsibility, independence), intensity (self-motivation), psychological energy (perseverance, ability to fulfil plans, ability to remain focused), and emotional stability (ability to control and channel nervousness, tolerance of stress, disposition to anxiety). Physical characteristics are measured by height, body mass index (scored as the absolute deviation from the mean body mass index), and a test of physical fitness. Sons' educational attainment, drawn from administrative records, is measured two ways: firstly, as years of schooling, and secondly, as a set of 43 dummy variables combining education level and field of study.

with fathers). The non-cognitive traits analysed in Mood et al. make a larger contribution to persistence than the non-cognitive traits analysed in the CHDS decomposition, which is likely due in part because the latter measure behavioural problems only (a subset of non-cognitive traits capturing maladaptive characteristics) while the former measure a wider set of non-cognitive traits that are primarily adaptive characteristics (see footnote 181 for details of the non-cognitive variables used in Mood et al. (2012)).

Hirvonen (2010) uses administrative data to estimate the role of IQ, non-cognitive traits, body mass index (all measured around age 18) and years of schooling attained by adulthood in mediating intergenerational income persistence among son-father pairs in Sweden.[182] When each pathway is considered alone, she finds that IQ accounts for 29 percent, non-cognitive traits for 21 percent, body mass index for one percent, and years of schooling for 45 percent of the IGE.[183] The non-cognitive effect is similar in magnitude to that from the CHDS decomposition (17 percent) but the IQ effect is smaller than the CHDS estimate of 37 percent. The years of schooling effect is similar to the analogous result for sons in the NZLC (40 to 45 percent) but is smaller than the CHDS result of 61 percent.

Hirvonen (2010) then performs a separate decomposition for daughters, estimating the role of years of schooling and field of education in mediating persistence among daughter-father pairs in Sweden (military enlistment records containing cognitive, non-cognitive, and health-related data being unavailable for Swedish women). She finds that years of schooling and field of education collectively account for 55 percent of the daughter-father IGE, which is slightly lower than the 63 percent found among daughter-father pairs in the NZLC decomposition.

Blanden et al. (2007) use survey data from the British Cohort Study (BCS, born 1970) to estimate the role of non-cognitive skills, cognitive skills, educational attainment, and labour market attachment in mediating intergenerational income persistence in Great Britain.[184] Their decomposition reveals that their set of non-cognitive variables, when considered alone, account for 19 percent of the IGE. This is similar to the 17 percent found in the CHDS decomposition. Their set of cognitive variables, when considered alone, account for 29 percent of the IGE, considerably lower than the 52 percent found in the CHDS. When the non-cognitive and cognitive variables are considered together, they collectively account for 32 percent of the IGE, again lower than the 54 percent in the CHDS. When education and labour market attachment are added, so that all sets of variables are included in the decomposition, the non-cognitive variables account

[182] Data on IQ, non-cognitive traits, and body mass index are obtained from Swedish military enlistment records covering all males compulsorily enlisted at around age 18. IQ is stanine-scored (i.e., scaled to range from one to nine with a mean of five and standard deviation of two) and converted into dummy variables for each stanine. Non-cognitive traits are stanine-scored from one to nine by a certified psychologist designed to measure conscripts' ability to cope with the psychological requirements of military service covering traits such as willingness to assume responsibility, independence, persistence, emotional stability, and initiative, and are converted to dummy variables for each stanine. Educational attainment by 1999 (ages 32 to 38) is obtained from administrative records and converted into years of schooling.

[183] When all pathways are considered simultaneously, they account for 53 percent of persistence (IQ accounts for 11 percent, non-cognitive traits for 12 percent, body mass index for zero percent, and years of schooling for 30 percent of the IGE).

[184] Non-cognitive skills consist of 10 variables measuring various behavioural problems and personality traits at ages 5, 10, and 16. Cognitive skills consist of five variables measuring language comprehension, visual motor co-ordination, word recognition, mathematical competence, and an ability test similar to an IQ test at ages 5 and 10. Educational attainment consists of five variables measuring number of O-level exams passed at age 16, staying on at school after age 16, number of A-level exams passed after age 16, staying in education after age 18, and degree completion. Finally, labour market attachment consists of two variables measuring the proportion of months spent unemployed between ages 16 and 30 and the proportion of months spent not in the labour force or education between ages 16 and 30.

for six percent, the cognitive variables account for seven percent, educational attainment accounts for 38 percent, and the labour market attachment variables account for nine percent of the IGE. In all, the full set of variables can account for about half (52 percent) of intergenerational income persistence, considerably less than found here for the CHDS data.[185]

Blanden et al. (2014) use survey data from the BCS for Great Britain and the PSID for the US to estimate the role of educational attainment, labour market attachment, marital status, health, and occupation in mediating intergenerational income persistence in Great Britain and the US.[186] When educational attainment is considered alone, it accounts for 29 percent of the IGE in Great Britain and 48 percent in the US, the latter result similar to the 40 to 45 percent among sons in the NZLC but both results lower than the 61 percent found among the CHDS sample.[187]

Gregg, Jonsson, et al. (2017) use administrative data for Sweden and survey data from the BCS for Great Britain and National Longitudinal Survey of Youth for the US to quantify the role of cognitive ability and educational attainment in mediating intergenerational income persistence among sons in the three countries.[188] They find that the combination of cognitive ability and educational attainment explains 57 percent of the IGE in Sweden (four percent accounted for by cognitive ability and 53 percent by educational attainment), 43 percent of the IGE in the US (eight percent accounted for by cognitive ability and 35 percent by educational attainment), and 35 percent of the IGE in Britain (two percent accounted for by cognitive ability and 33 percent by educational attainment). These results differ from the general picture painted by the CHDS results. Whereas the results from Gregg, Jonsson, et al. (2017) suggest that cognitive ability has little influence on offspring's earnings and intergenerational persistence beyond its impact via

[185] Blanden et al. (2007) also use data on sons drawn from the National Child Development Study (born 1958) to estimate the IGE of son's log earnings at age 33 with respect to their parents' log family income when the son was aged 16, which they find to be 0.205. They then decompose this estimate into the same four sets of variables as in the BCS analysis (non-cognitive skills, cognitive skills, educational attainment, and labour market attachment, although measured with different items than in the BCS). When all transmission mechanisms are considered together, they find that non-cognitive skills account for close to zero percent, cognitive skills account for 11 percent, educational attainment accounts for 38 percent, and labour market attachment accounts for 10 percent of the IGE, meaning in total 53 percent of the IGE can be explained by these factors. Blanden et al. (2007) then go on to show that over 80 percent of the increase in intergenerational persistence between the 1958 cohort and the 1970 cohort can be explained by changes in these transmission mechanisms (mainly via a strengthening of the link between family income and educational attainment and labour market attachment).

[186] Educational attainment at age 30 is measured with the following incremental indicators: 'no school qualifications', 'at least high school qualifications', 'at least some college', and 'completed college'. Labour market attachment is measured as the proportion of years between ages 22-25 and 26-29 that the son was not in the labour market or in education, and the proportion of years between the same ages that the son was in full-time or near full-time work or education. Marital status is measured with indicators for being younger than 23 at first marriage and being married at age 30. Health status at age 30 is measured with indicators for excellent self-reported health and fair, poor, or very poor self-reported health. Finally, occupation at ages 30 and 34 is measured with incremental indicators over a seven-point hierarchy ranging from 'at least semi-routine' up to 'at least higher managerial and professional'.

[187] Adding variables for early marriage, labour market attachment, marital status, health status, occupation at age 30, and occupation at age 34 increases the explained percentage to 52 percent in Great Britain and 55 percent in the US, meaning the set of factors as a whole can explain about half of intergenerational income persistence among British and American sons. In this full decomposition, educational attainment accounts for 12 percent of the IGE in Great Britain but more than twice as much – 26 percent – in the US, while occupation at age 34 accounts for 19 percent of the IGE in Great Britain but about half as much – 10 percent – in the US. Marriage and health variables "have very little explanatory power for either country" (Blanden et al., 2014, p. 436).

[188] Educational attainment is coded to the following internationally-comparable four-category classification: (1) no formal qualifications (high school dropout in the US, less than O-level in Britain, and comprehensive in Sweden); (2) lower secondary qualification (high school graduate in the US, O-level in Britain, short upper secondary in Sweden); (3) upper secondary or pre-degree tertiary qualification (associate's in the US, A-levels in Britain, long upper secondary or short post-secondary in Sweden), and; (4) university qualification. Cognitive ability is measured by the Armed Forces Qualification test taken at age 17 among the US sample, a derived score from various tests of cognitive ability at age 16 (or if missing, at age 10) among the British sample, and a test of cognitive ability administered at age 18 as part of compulsory military enlistment in the Swedish sample.

offspring's education, the CHDS results suggest cognitive ability has a much more direct effect on earnings and persistence.

Björklund et al. (2017) use administrative data for Sweden and BCS data for Great Britain to estimate the role of birthweight, height in adolescence, average school grade at age 16, and educational attainment in adulthood in mediating intergenerational income persistence in the two countries, separately for sons and daughters.[189] Among sons, when each trait is considered alone, they find that birthweight and height account for almost none of the IGC in both countries (for birthweight, zero percent in Sweden and at most two percent in Great Britain; for height, two percent in Sweden and six percent in Great Britain), average school grade accounts for 35 percent of persistence in Sweden and 33 percent in Great Britain, and educational attainment accounts for 28 percent of persistence in Sweden and 34 percent in Great Britain. The latter results on educational attainment are both lower than the 40 to 45 percent found for sons in the NZLC and the 61 percent found in the CHDS. Among daughters, Björklund et al. find that birthweight accounts for almost none of the IGC in both countries, average school grade accounts for 29 percent of persistence in Sweden and 33 percent in Great Britain, and educational attainment accounts for 23 percent of persistence in Sweden and 45 percent in Great Britain.[190] The latter result differs markedly from the NZLC results for daughters (13 percent among daughter-mother pairs and 63 percent among daughter-father pairs), although, again, if mothers and fathers were pooled in the NZLC as in Björklund et al. (2017) then the results may have been closer.

Taking these study comparisons as a whole, the results are mixed, but some patterns are evident. The estimated mediating effect of sons' educational attainment from the NZLC decomposition is similar to analogous estimates discussed above for sons in Sweden from Hirvonen (2010) and sons in the US from Blanden et al. (2014), as well as being similar to a New Zealand estimate pooling sons and daughters (hence less analogous) from Gibbons (2010). For daughters, only Hirvonen (2010) estimates daughter-specific mediating effects for education as in the NZLC decomposition and her estimate for daughter-father pairs is broadly similar to (albeit slightly lower than) the equivalent daughter-father estimate from the NZLC. Note that, to the author's knowledge, there are no studies in the literature that decompose daughter-mother income persistence into mediating pathways, so the NZLC daughter-mother decomposition results are novel in this respect. In contrast to these NZLC results, the mediating role of educational attainment as estimated in the CHDS decomposition is generally higher than equivalent estimates from other studies, being higher than estimates from Gibbons (2010) for New Zealand and from Björklund et al. (2017) for Sweden (both of which also pool sons and daughters) as well as being

[189] Persistence is measured by the IGC. Data on birthweight are obtained from hospital records in both countries and measured in three ways; in kilograms, in logarithms, and a dummy indicator for low birthweight. Data on height (in metres) are obtained from assessments of BCS cohort members at age 16 and from Swedish military enlistment records at age 18 (for men only) which are then age-standardised. Because height data are not available for Swedish women, the role of height in mediating intergenerational persistence is not examined in their analysis of daughters. To derive average school grade, data on offspring's grades in selected subjects is transformed into a percentile rank and then the average percentile rank across subjects is used as the offspring's grade (grades in O-level or CSE exams in 12 subjects in Britain and grades in exams in 10 compulsory subjects taken at the end of compulsory schooling in Sweden). For educational attainment, Björklund et al. use the ISCED classification to code offspring's educational attainment at age 34 in both countries into four categories: less than high school, at least high school, at least some college or post-secondary, and completed college. Only sons and daughters with valid data on all pathway variables are included in the decomposition.

[190] When all sons' (daughters') traits are considered simultaneously, Gregg, Jonsson, et al. (2017) find they collectively account for 38 percent (34 percent) of persistence in Sweden and 46 percent (56 percent) in Great Britain.

higher than (the less-analogous) son-only estimates from Blanden et al. (2014) for Great Britain and the US. Thus, when estimated with NZLC data, the importance of education to intergenerational income persistence in New Zealand is similar to other countries (Sweden and the US) and to a previous estimate for New Zealand, but when estimated with CHDS data, is higher than in other countries (Sweden, the US, and Great Britain) and a previous estimate for New Zealand.

There is no such ambiguity, however, when comparing the estimated mediating effect of cognitive ability from the CHDS decomposition with analogous results from other studies, which is found to be consistently higher. Of all the results reported in this chapter, it is this result that stands out as the most distinctive and is therefore discussed in some detail here including its potential policy implications. Recall that cognitive ability accounts for half of persistence when considered alone, which is higher than analogous results discussed above from Mood et al. (2012), Blanden et al. (2007), and Gregg, Jonsson, et al. (2017). Furthermore, IQ accounts for 37 percent of persistence when only cognitive skills are included as pathways, and still accounts for a non-trivial proportion of persistence (16 percent) after education and occupation are added, both of which are higher than the equivalent figures from Hirvonen (2010).

This lingering influence of childhood IQ on economic outcomes in adulthood is consistent with analysis of CHDS data by Fergusson et al. (2005) who use multiple linear regression to test for an association between childhood IQ and annual personal income at age 25 (and other economic, educational, and behavioural outcomes across ages 15 to 25), in both a bivariate 'unadjusted' model and a multivariable model that adjusts for a wide range of potential confounders drawn from the CHDS dataset (including childhood measures of family socioeconomic disadvantage, family instability, child conduct problems, parental adjustment problems, and child abuse).[191] They find a statistically significant positive bivariate association between IQ and income at age 25 ($p$<0.0001) with a coefficient of $2,966 (that is, a one-level increase in IQ is associated with a $2,966 increase in annual income at age 25 – about nine percent of the average income in the sample – unadjusted for potential confounders). In the multivariable model, the association remained significant ($p$<0.05) but the coefficient had reduced to $1,595 after adjustment for all covariates, indicating that the covariates explained about half of the association and that children from the highest IQ group earned on average approximately $8,000 more per annum at age 25 than children from the lowest IQ group, controlling for confounders. Childhood IQ was also significantly associated with educational attainment by age 25 and duration of unemployment between ages 18 and 25. Fergusson et al. (2005, p. 856) conclude that "increasing IQ was associated with increasing educational success at school, higher rates of post-school educational/vocational attainment, degree success, lower rates of unemployment and higher income at age 25" and "[s]tatistical control for a wide range of factors including early conduct problems and family, social and childhood circumstances failed to explain these associations,

[191] The continuous IQ scores were classified into five levels (the lowest level equal to an IQ score of 84 or less, the highest level equal to an IQ score of 115 or more), however Fergusson et al. (2005, p. 852) note that "[f]urther analysis…in which IQ was scored as a continuous variable produced an identical set of conclusions". Each potential confounder was measured with a composite index. For example, family socioeconomic disadvantage was scored on the basis of low socioeconomic status at birth, both parents having no formal educational qualifications at birth, and a below-average standard of living as rated by interviewers at three or more annual interviews between ages 1 to 12.

supporting the view that intelligence had a direct relationship to later educational, occupational and related outcomes independently of other childhood characteristics and family environment".

Also using CHDS data, Maani and Kalb (2007) estimate the effects of IQ at age 8 and family income decile over childhood (alongside a range of other individual and family characteristics) on academic achievement at age 15 (measured by average grade in School Certificate national examinations) and the subsequent decision to leave or stay in school at age 16. In a model that jointly estimates academic achievement and school leaving simultaneously, they find that IQ and family income over childhood are positively and significantly related to academic achievement, and academic achievement in turn strongly affects the decision to leave school. Thus, "[i]n the school-leaving choice, the student's academic performance is an important channel through which personal ability [IQ] and economic factors [family income over childhood] exert their influence" and "[t]o the extent that the academic performance of students can be influenced throughout their education years, these results point to the role that childhood and teenage academic performance can play in breaking cycles of disadvantage" (Maani & Kalb, 2007, p. 372). These results suggest a sequential link between family income, child IQ, and academic achievement at secondary school that is akin to the sequential links implied in the CHDS decomposition reported here.

Other evidence points to the importance of inequalities in cognitive abilities to intergenerational income persistence. Using country-level data on son-father pairs in seven high-income countries, Corak (2006, p. 159) plots a measure of each country's gradient between sons' cognitive abilities and their fathers' education[192] against an estimate of the country's son-father earnings elasticity and finds "a very strong positive relationship…with the raw correlation between the two statistics being 0.856". This result, while correlational, supports the notion that parental influence on offspring's cognitive abilities is a key mechanism driving intergenerational income persistence. In New Zealand, the gradient between parental education and offspring's cognitive skills appears to be relatively strong compared to other countries. Maré and Stillman (2010) use 2003 Programme for International Student Assessment (PISA) data on 15-year-old secondary school students from 29 OECD countries including New Zealand[193] to examine the relationship between parental education (measured on a seven-point scale, ranging in New Zealand from 'Did not finish intermediate school' up to 'Has a university degree') and the cognitive skills of their children (measured by PISA standardised test scores in mathematics, science, reading, and problem-solving). They find a strong positive relationship between parental education and children's test scores; children with a university-educated parent scored 0.75-0.90 standard deviations higher on PISA cognitive tests than children whose parents finished only intermediate schooling. Focusing on maths scores only, they then compare parental education-test score gradients across the 29 countries and find that the gradient is stronger in New Zealand than in all but nine other OECD countries (among the nine countries with a stronger gradient are the US, Japan,

[192] This gradient is measured with data from the International Adult Literacy Survey (IALS) as the relationship between the literacy test scores of male IALS participants aged in their thirties and their fathers' education, as reported in Esping-Andersen (2004).

[193] The New Zealand sample size was n=2,694. The sample for each country was restricted to students who were born in that country and whose parents were also born in that country.

Germany, and Slovakia which had the strongest gradient).[194] Park (2008) uses 2001 Progress in International Reading Study (PIRLS) data on 10-year-old students from 25 countries including New Zealand[195] to predict student's reading performance (as measured by the PIRLS reading test) based on the home literacy environment provided by parents (as measured by parental engagement in literacy activities with their child during preschool years, parental attitudes to reading, and number of books in the home), employing country-specific models separately for each parental input with controls for parental education, child's gender, number of children and adults in the home, rurality of child's school, and language minority status. Park finds that reading performance is significantly associated with parental attitudes towards reading and number of books in the home in all 25 countries, and significantly associated with parental engagement in child literacy activities in all but two countries. The effects of both parental engagement in child literacy activities and parental reading attitudes on reading score are larger in New Zealand than all other countries and the effect of number of books in the home is third-largest in New Zealand.[196] It seems, then, that the role played by the family in generating intergenerational persistence is greater in New Zealand than in most other countries.

By comparison, earnings returns to skills (numeracy, literacy, and problem-solving skills) in New Zealand are very close to the average for the 32 countries who participated in PIAAC 2016 (Hanushek, Schwerdt, Wiederhold, & Woessmann, 2017) and returns to tertiary education in New Zealand are consistently below the OECD average (Zuccollo, Maani, Kaye-Blake, & Zeng, 2013). Using country-level data from 14 high-income countries, Corak (2013, pp. 87-88) plots each country's "college earnings premium" – measured as the average earnings in 2009 of men aged 25 to 34 years with a university degree relative to the average earnings of their counterparts with a high school qualification – against an estimate of the country's son-father earnings elasticity and finds a noticeable positive relationship, indicating that "in countries where the return to college education is higher, intergenerational mobility tends to be lower". New Zealand appears in the lower-left quadrant of Corak's (2013) plot, with relatively low returns to higher education and a middling earnings elasticity. A comparison of these results to those of parental influence discussed above suggest that, in New Zealand, the family, rather than the labour market, is likely to be the stronger institutional force driving intergenerational persistence.

[194] Australia had a gradient of 0.58 which was similar to the Nordic countries. Maré and Stillman (2010) go on to show that the New Zealand gradients in reading and problem-solving disappear, and the gradients in maths and science are considerably flattened, once differences in children's student-school interactions, household characteristics, household educational resources, and to a lesser extent school characteristics, are statistically controlled for. See also Hernandez's (2019) study of the contribution made by schools to the educational achievement of secondary school students in New Zealand, independent of the socioeconomic background of their student bodies (the 'value-add' of schools), using administrative data on nearly 400,000 students from 480 secondary schools who attempted the National Certificate of Educational Achievement (NCEA) from 2008 to 2017. While estimating the influence of families on educational achievement is not his primary aim, Hernandez nevertheless finds that, among a range of family background variables, parents' educational attainment is the strongest predictor of students' NCEA achievement, while parents' log incomes averaged over an 18-year period are associated with *declines* in educational achievement, once adjusted for parents' educational attainment and other characteristics.

[195] The New Zealand sample size was n=2,086.

[196] When Park (2008) standardises the home literacy variables and includes them all simultaneously in country-specific models predicting reading performance, he finds that, of all 25 countries, the standardised effect of engagement in child literacy activities is third-largest in New Zealand, the standardised effect of parental attitudes to reading is largest in New Zealand, and the standardised effect of number of books is fifth-largest in New Zealand. He also finds that all three variables (but overwhelmingly the latter two) play a role in mediating the relationship between parental education and reading performance, collectively accounting for 39 percent of the reading score gap between children with a tertiary-qualified parent and children with lower-secondary-qualified parents.

It is important to note that intelligence is not an immutable capacity fixed at birth and impervious to intervention (Cianciolo & Sternberg, 2004). Intelligence does, of course, have a genetic component and IQ displays heritability of approximately 0.5, meaning that about half of the variation in IQ scores – averaged across various populations of various ages at various times and places – is attributable to genetic effects (Mandelman & Grigorenko, 2011; Plomin, 2019).[197] Yet intelligence is also highly susceptible to environmental influences including family effects (Flynn & Sternberg, 2020)[198] and appears to be subject to critical periods, being "malleable when children are very young but…difficult to change after about the age of 10" (Cunha et al., 2006, p. 757). As intelligence experts Cianciolo and Sternberg (2004, pp. 96-97, 108) note:

> …the presence of a genetic influence on intelligence does not indicate that people's levels of intelligence are predetermined and unmodifiable. In fact, the heritability and the modifiability of intelligence are unrelated. Consider, for example, height. Height is a highly heritable characteristic (i.e., differences among people in height frequently appear to be largely due to genetic differences), but height can be reduced through environmental stressors, such as prenatal exposure to toxins and childhood malnutrition, or enhanced through medical interventions, such as the injection of growth hormones. In addition, average height has increased dramatically and rapidly throughout the past several generations, as has IQ….there is substantial interplay between genes and environment in determining one's intellectual competency.

As noted in the quotation above, substantial increases in IQ scores over successive generations – known as the 'Flynn effect' – have been documented by intelligence researcher James Flynn (1984, 1987). The Flynn effect constitutes evidence for the malleability of intelligence, as Sternberg and Grigorenko (1999, pp. 541, 542) explain:

> Environment has a powerful effect on many attributes, including cognitive abilities. These effects may interact with genetic structures, but they nevertheless can result in massive modifications of demonstrated cognitive or other performances. Perhaps the simplest and most potent demonstration of this fact is called the "Flynn effect"….The basic phenomenon of the Flynn effect is an increase in IQ throughout successive generations around the world during the past 30 years…This effect is most likely environmental because a successive stream of genetic changes of such dramatic magnitude could not have occurred and exerted strong influence in such a short period of time….the environment can have and has had a massive effect on intellectual ability, regardless of what the heritability of intelligence may be.

The results in the current analysis say nothing about whether the intergenerational income persistence observed in the NZLC and CHDS samples is primarily driven by genetic or environmental factors. However, studies by Björklund, Lindahl and Plug (2006), Björklund, Jäntti,

[197] Even the notion of a gene versus environment dichotomy – or at least some conceptions of it – has been subject to criticism. Sternberg and Grigorenko (1999, p. 540) assert that "there are no pure genetic effects on behaviour", but rather "[g]enes express themselves through covariation and interaction with the environment". Grigorenko (2000, pp. 76, 77) questions the assumption of the separability of the main effects of genes and environments given that "development is messy and cannot be divided neatly into causal forces of genes and environments" and notes that "relatively little is known about detailed aspects of coactions of genes and environments in regard to human intellectual functioning". In a similar vein, Cunha and Heckman (2007, p. 32) assert that "the nature versus nurture distinction is obsolete" and that "[a]dditive "nature" and "nurture" models, while traditional and still used in many studies of heritability and family influence, mischaracterize how ability is manifested" and "[g]enes and environment cannot be parsed meaningfully by traditional linear models that assign variance to each component".

[198] As with income, IQ is itself correlated across generations (e.g., Björklund, Eriksson, & Jäntti, 2010; Black, Devereux, & Salvanes, 2009) as are cognitive abilities measured more generally (Grönqvist, Öckert, & Vlachos, 2017), for both genetic and environmental reasons. Black et al. (2009) estimate the correlation (elasticity) between the IQ of 24,754 Norwegian fathers born 1932 and 1933 and that of their sons to be 0.38 (0.32). Björklund et al. (2010) estimate the correlation (elasticity) between the IQ of 22,626 Swedish fathers born 1961 to 1956 and that of their sons born 1966 to 1980 to be 0.347 (0.327) and the brother correlation in IQ (for 379,456 brothers born in the period 1951 to 1968) to be 0.473. Grönqvist et al. (2017) estimate the correlation between the cognitive ability of 50,171 Swedish fathers (average year of birth 1954) and that of their sons (average year of birth 1981) to be 0.350, but when they correct for measurement error by using the father's brother's (the son's uncle's) cognitive ability as an instrument for the father's ability, the correlation rises to 0.478. Grönqvist et al. (2017, p. 889) conclude that "the transmission of productive abilities can explain a substantial part of the intergenerational correlation in economic outcomes".

and Solon (2007), and Black, Devereux, Lundborg, and Majlesi (2020) all use Swedish administrative data on adopted children matched to their biological and adoptive parents to disentangle the role of nature versus nurture in the intergenerational transmission of earnings, income, years of schooling, and other economic outcomes.[199] Given their samples are of children adopted out at very young ages, they consider biological parents' characteristics to be an indicator of pre-birth factors including genes and the prenatal environment, while adoptive parents' characteristics are an indicator of post-birth factors such as childhood environment. If adoptees' economic outcomes are more closely linked to that of their biological parents than their adoptive parents, then that implies that 'nature' or pre-birth factors are more important than 'nurture' or post-birth factors in intergenerational persistence. All three studies find that both biological parents and adoptive parents influence adoptee's economic outcomes, and Björklund et al. (2006) and Black et al. (2020) find that adoptee's earnings are more strongly associated with their adoptive father's earnings than with their biological father's earnings, and similarly for income, suggesting that post-birth environmental channels play a larger role in the intergenerational transmission of earnings and income than pre-birth factors.[200]

Furthermore, in their review of the evidence on gene-environment interaction effects and epigenetic effects[201] on socioeconomic outcomes, Lundborg and Stenberg (2010, p. 320) conclude that "environments may affect heritable traits across generations" so "policies which neutralize adverse environments may also increase intergenerational mobility, given that genetic and/or environmental risk factors are more common in socially disadvantaged groups". Such policies "may compensate groups of individuals carrying genetic risks, without the need to identify anyone's genetic endowments" since they target environmental risks yet disproportionately benefit those at genetic risk due to gene-environment interactions (Lundborg & Stenberg, 2010, p. 320).

These findings on the malleability of intelligence and the importance of environmental factors to intergenerational transmission indicate that the persistence of incomes across generations found in the current analysis, and its mediation by childhood IQ, reading ability, and educational attainment, are all amenable to policy intervention.[202] In particular, the results suggest interventions in childhood aimed at promoting cognitive development and reading ability among children from poorer family backgrounds could help to lift their educational attainment and ultimately weaken the intergenerational link in incomes, were that to be a goal of public policy. As Corak (2004, p. 33) argues, "[t[he inheritance of education, occupation, and income is influenced in the first instance by the impact parents have on a child's cognitive performance, and societies that have leveled the playing field with respect to these circumstances have had the most success

[199] Björklund et al. (2006) use the log of earnings and income and also examine educational attainment, while Black et al. (2020) use (within-cohort) ranks of earnings and income and also examine educational attainment and wealth (the latter being their primary focus). In both studies, only fathers' earnings and income are analysed, as adoption guidelines in Sweden at the time the offspring were adopted indicated an expectation that adoptive mothers would not be employed in paid work.

[200] In contrast to their findings on earnings and income, Black et al. (2020) find nature to be more important than nurture in the intergenerational transmission of educational attainment, while Björklund et al. (2006) find the two components have equal effects via transmission from fathers but nature effects dominate the transmission from mothers.

[201] These refer to traits that are affected by heritable changes in gene functions independent of the DNA sequence (Lundborg & Stenberg, 2010).

[202] The finding that there are large differences in intergenerational mobility across countries also suggests that public policy can create more mobile societies (OECD, 2018).

in promoting generational mobility". Similarly, in their review of the evidence on skill formation over the lifecycle, Cunha et al. (2006, pp. 703, 793) find that "[t]he quantitatively important constraints facing disadvantaged children are the ones determining their early environments" and "[h]aving access to more and higher-quality resources that contribute to improving cognitive ability early in life affects skill acquisition later in life" hence "[g]overnment policies to promote early accumulation of human capital should be targeted to the children of poor families". OECD (2018) catalogue a range of policy options that governments can implement that could potentially enhance mobility,[203] including high-quality early childhood education, supplementing the incomes of low-income parents, limiting early tracking in schools, and reducing socioeconomic segregation in neighbourhoods and schools.[204] Policies to support maternal mental health and reduce maternal stress during pregnancy have also been identified as critical to breaking cycles of intergenerational disadvantage because of their negative effects on child development (Low et al., 2021). OECD (2018, p. 56) conclude that "[a]ll in all, governments have various policy tools at hand that, depending on the country-specific conditions, can help them address one of the defining challenges of our time: promoting social mobility, within and across generations, and to give everyone a chance to fully express their talent and potential".

### 5.5 Bias and limitations

As prefaced earlier in section 5.1, to correctly identify the mediating role played by offspring's traits in transmitting income from parent to offspring, it is necessary to make the strong identifying assumption that there are no unobserved factors influencing both parental income and offspring's outcomes. As Blanden et al. (2014, p. 434) concede, this is a "stringent assumption which is unlikely to be met in reality". If unobserved parental characteristics such as parents' education influence both parents' income and the mediating offspring trait, then the importance of that trait in intergenerational transmission will be overstated.[205] Moreover, the estimated effect of any given

[203] Many such policies involving large-scale income redistribution are likely to be subject to trade-offs, notably compression of the after-tax wage distribution and resulting reductions in the returns to education, which may reduce financial incentives to undertake further education (Landersø & Heckman, 2017). Based on their empirical comparison of intergenerational mobility in Denmark and the US, Heckman and Landersø (2021) conclude that "[d]espite generous Danish social policies, family influence on important child outcomes in Denmark is about as strong as it is in the United States" and "Denmark achieves lower income inequality and greater intergenerational income mobility [than the US] primarily through its tax and transfer programs and not by building the skills of children across generations and promoting their human potential more effectively".

[204] Such policies may require acceptance by high-income parents of a reduction in their ability to pass on advantage to their offspring and indeed of downward mobility for their offspring. For example, in arguing for a reduction in socioeconomic segregation among New Zealand schools, Lauder and Hughes (1990a, p. 57) acknowledge that "[w]hile we would not want to deny that there would probably be some reduction in achievement for pupils presently attending privileged [private] schools, we would argue that this should not be used as a justification for the status quo" since "[i]n our view the existence of privileged private schools is socially divisive and in an area such as education we do not think the individual 'freedom' to buy privilege is justified" and that "[i]t should also be remembered that the reduction in performance by these privileged pupils will be matched by an increase in performance by low [socioeconomic status] students".

[205] Indeed, Mendolia and Siminski (2017), who use Australian data from the HILDA survey to quantify the role of education as a mechanism through which family background affects offspring's hourly earnings, find that the role played by offspring's educational attainment in mediating intergenerational transmission becomes overstated by a factor of roughly 1.6 to 1.8 when the definition of family background is changed from a comprehensive, multidimensional construct encompassing the sorts of parental characteristics subsumed within the IGE that are correlated with their own income and influence their offspring's earnings (Mendolia and Siminski use respondents' retrospective reports of their mothers' and fathers' education, mothers' and fathers' occupations, mothers' and fathers' countries of birth, and other demographic and socioeconomic parental characteristics) to simply parents' earnings (imputed from parents' occupation, as retrospective reports of parental income were not collected in HILDA). In other words, the role of education in intergenerational income persistence is exaggerated because it funnels the effects of other parental characteristics encapsulated by the IGE.

pathway variable is sensitive to the inclusion of other pathway variables and the order in which they are included in the equation (Hirvonen, 2010).

Yet the addition of more pathway variables to the decomposition may reduce bias in the estimation of mediating effects. As Hirvonen (2010, p. 22) notes, “the problem of omitted variable bias is likely to impair the results [from this decomposition technique], yet including more covariates in the equation should also decrease the degree of the total inconsistency in the estimates” and “[n]ormally, the more variables are involved in the equation, the higher the propensity for them to net out the effect of bias” and results can “still present a useful upper bound to the true impact of these factors on the persistence of earnings”. Mood et al. (2012, pp. 64, 70) likewise note that:

> …the [decomposition] models we test are descriptive in the sense that causality is a conjecture rather than an established fact….the included [mediating] variables can obviously be related to unmeasured variables that also transmit the intergenerational correlation, so the relative weights of different mediating variables could change if more variables were included. These caveats notwithstanding, in our case the problem of unobserved variables is smaller than usual because most of the prime suspects (cognitive and non-cognitive characteristics) [believed to be determinants of income but often unobserved in practice] are observed. Even if we encourage a healthy skepticism to interpreting the details of the decomposition….we still believe that the big picture is informative.

Despite these limitations of the decomposition approach, in practice there may be few feasible alternatives that do not also rely on strong assumptions. As Blanden et al. (2014, p. 428) note, “[t]here exist no data sources or complete structural model by which we can compare all aspects of parental inputs to child outputs from cradle to adulthood” hence “the best we can do is to isolate various channels by which status is transmitted from parent to child”.

## References


Aaronson, D., & Mazumder, B. (2008). Intergenerational Economic Mobility in the United States, 1940 to 2000. *The Journal of Human Resources, 43*(1), 139-172. https://doi.org/10.2307/40057341

Alimi, O. B., Maré, D., & Poot, J. (2018a). *International Migration and the Distribution of Income in New Zealand Metropolitan and Non-Metropolitan Areas*. Bonn, Germany: Institute for the Study of Labor. Retrieved from https://www.iza.org/publications/dp/11959/international-migration-and-the-distribution-of-income-in-new-zealand-metropolitan-and-non-metropolitan-areas

Alimi, O. B., Maré, D., & Poot, J. (2018b). *Who partners up? Educational assortative matching and the distribution of income in New Zealand*. Wellington: Motu Economic and Public Policy Research. Retrieved from https://motu.nz/our-work/population-and-labour/individual-and-group-outcomes/who-partners-up-educational-assortative-matching-and-the-distribution-of-income-in-new-zealand/

Altonji, J. G., & Dunn, T. A. (2000). An Intergenerational Model of Wages, Hours, and Earnings. *The Journal of Human Resources, 35*(2), 221-258. https://doi.org/10.2307/146324

Andrews, D., & Leigh, A. (2008). *More Inequality, Less Social Mobility*. Canberra. Retrieved from https://www.academia.edu/3599803/More_inequality_less_social_mobility

Andrews, D., & Leigh, A. (2009). More inequality, less social mobility. *Applied Economics Letters, 16*(15), 1489-1492. https://doi.org/10.1080/13504850701720197

Atkinson, A. B., & Leigh, A. (2008). Top Incomes in New Zealand 1921–2005: Understanding the Effects of Marginal Tax Rates, Migration Threat, and the Macroeconomy. *Review of Income and Wealth, 54*(2), 149-165. https://doi.org/10.1111/j.1475-4991.2008.00268.x

Atkinson, J., Shaw, C., Blakely, T., Stanley, J., & Sloane, K. (2010). *Linkage of Census and Cancer Registrations, 1981-2004: CancerTrends Technical Report: Number 1*. Wellington: University of Otago. Retrieved from https://www.otago.ac.nz/wellington/otago020222.pdf

Baldock, C. V. (1971). *Vocational Choice and Opportunity*. Christchurch: University of Canterbury.

Becker, G. S., Kominers, S. D., Murphy, K. M., & Spenkuch, J. L. (2018). A Theory of Intergenerational Mobility. *Journal of Political Economy, 126*(S1), S7-S25. https://doi.org/10.1086/698759

Becker, G. S., & Tomes, N. (1979). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy, 87*(6), 1153-1189. https://doi.org/10.2307/1833328

Becker, G. S., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics, 4*(3), S1-S39. https://doi.org/10.2307/2534952

Belich, J. (2001). *Paradise Reforged: A History of the New Zealanders from the 1880s to the Year 2000*. Auckland: Penguin Press.

Belsky, D. W., Domingue, B. W., Wedow, R., Arseneault, L., Boardman, J. D., Caspi, A., . . . Harris, K. M. (2018). Genetic analysis of social-class mobility in five longitudinal studies. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1801238115

Bertram, G. (2009). The New Zealand Economy, 1900-2000. In G. Byrnes (Ed.), *The New Oxford History of New Zealand* (pp. 537-572). Auckland: Oxford University Press.

Björklund, A., & Chadwick, L. (2003). Intergenerational income mobility in permanent and separated families. *Economics Letters, 80*(2), 239-246. https://doi.org/10.1016/S0165-1765(03)00084-3

Björklund, A., Eriksson, K. H., & Jäntti, M. (2010). IQ and Family Background: Are Associations Strong or Weak? *The B.E. Journal of Economic Analysis & Policy, 10*(1), 1-12.

Björklund, A., & Jäntti, M. (2009). Intergenerational Income Mobility and the Role of Family Background. In W. Salverda, B. Nolan, & T. M. Smeeding (Eds.), *The Oxford Handbook of Economic Inequality* (pp. 491-521). Auckland: Oxford University Press.

Björklund, A., & Jäntti, M. (2020). Intergenerational mobility, intergenerational effects, sibling correlations, and equality of opportunity: A comparison of four approaches. *Research in Social Stratification and Mobility, 70*, 100455. https://doi.org/10.1016/j.rssm.2019.100455

Björklund, A., Jäntti, M., & Nybom, M. (2017). The Contribution of Early-life Versus Labour Market Factors to Intergenerational Income Persistence: A Comparison of the UK and Sweden. *The Economic Journal, 127*(605), F71-F94. https://doi.org/10.1111/ecoj.12328

Björklund, A., Jäntti, M., & Solon, G. (2007). Nature and Nurture in the Intergenerational Transmission of Socioeconomic Status: Evidence from Swedish Children and Their

Biological and Rearing Parents. *The B.E. Journal of Economic Analysis & Policy, 7*(2), 1-21.
Björklund, A., Lindahl, M., & Plug, E. (2006). The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data. *The Quarterly Journal of Economics, 121*(3), 999-1028.
Björklund, A., Roine, J., & Waldenström, D. (2012). Intergenerational top income mobility in Sweden: Capitalist dynasties in the land of equal opportunity? *Journal of Public Economics, 96*(5), 474-484. https://doi.org/10.1016/j.jpubeco.2012.02.003
Black, S. E., & Devereux, P. J. (2011). Recent Developments in Intergenerational Mobility. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 4B, pp. 1487-1541): Elsevier. https://doi.org/10.1016/S0169-7218(11)02414-2
Black, S. E., Devereux, P. J., Lundborg, P., & Majlesi, K. (2020). Poor Little Rich Kids? The Role of Nature versus Nurture in Wealth and Other Economic Outcomes and Behaviours. *The Review of Economic Studies, 87*(4), 1683-1725. https://doi.org/10.1093/restud/rdz038
Black, S. E., Devereux, P. J., & Salvanes, K. G. (2009). Like father, like son? A note on the intergenerational transmission of IQ scores. *Economics Letters, 105*(1), 138-140. https://doi.org/10.1016/j.econlet.2009.06.022
Blanden, J. (2013). Cross-Country Rankings in Intergenerational Mobility: A Comparison of Approaches from Economics and Sociology. *Journal of Economic Surveys, 27*(1), 38-73. https://doi.org/10.1111/j.1467-6419.2011.00690.x
Blanden, J., Goodman, A., Gregg, P., & Machin, S. (2004). Changes in intergenerational mobility in Britain. In M. Corak (Ed.), *Generational Income Mobility in North America and Europe* (pp. 122-146). Cambridge, UK: Cambridge University Press.
Blanden, J., Gregg, P., & Macmillan, L. (2007). Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education. *The Economic Journal, 117*(519), C43-C60. https://doi.org/10.2307/4625498
Blanden, J., Gregg, P., & Macmillan, L. (2011). *Intergenerational Persistence in Income and Social Class: The Impact of Within-Group Inequality*. Bonn, Germany: Institute for the Study of Labor. Retrieved from http://ftp.iza.org/dp6202.pdf
Blanden, J., Haveman, R., Smeeding, T., & Wilson, K. (2014). Intergenerational Mobility in the United States and Great Britain: A Comparative Study of Parent–Child Pathways. *Review of Income and Wealth, 60*(3), 425-449. https://doi.org/10.1111/roiw.12032
Bloise, F., & Raitano, M. (2021). Intergenerational Earnings Persistence in Italy between Actual Father–Son Pairs Accounting for Lifecycle and Attenuation Bias. *Oxford Bulletin of Economics and Statistics, 83*(1), 88-114. https://doi.org/10.1111/obes.12375
Boden, J. M., Fergusson, D. M., & Horwood, L. J. (2013). Pathways to economic outcomes at age 30: Income and living standards in a New Zealand birth cohort. *New Zealand Sociology, 28*(3), 102-135.
Boden, J. M., Newton-Howes, G., Foulds, J., Spittlehouse, J., & Cook, S. (2019). Trajectories of alcohol use problems based on early adolescent alcohol use: Findings from a 35 year population cohort. *International Journal of Drug Policy, 74*, 18-25. https://doi.org/https://doi.org/10.1016/j.drugpo.2019.06.011
Böhlmark, A., & Lindquist, M. J. (2005). *Life-Cycle Variations in the Association Between Current and Lifetime Income: Country, Cohort and Gender Comparisons*. Stockholm: Swedish Institute for Social Research (SOFI), Stockholm University. Retrieved from https://www.sofi.su.se/english/research/dissertations-and-publications/sofi-working-papers
Böhlmark, A., & Lindquist, M. J. (2006). Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden. *Journal of Labor Economics, 24*(4), 879-896. https://doi.org/10.1086/506489
Bratberg, E., Davis, J., Mazumder, B., Nybom, M., Schnitzlein, D. D., & Vaage, K. (2017). A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the US. *The Scandinavian Journal of Economics, 119*(1), 72-101. https://doi.org/10.1111/sjoe.12197
Bratberg, E., Nilsen, Ø. A., & Vaage, K. (2005). Intergenerational Earnings Mobility in Norway: Levels and Trends. *Scandinavian Journal of Economics, 107*(3), 419-435. https://doi.org/10.1111/j.1467-9442.2005.00416.x
Bratsberg, B., Røed, K., Raaum, O., Naylor, R., Jäntti, M., Eriksson, T., & Österbacka, E. (2007). Nonlinearities in Intergenerational Earnings Mobility: Consequences for Cross-Country Comparisons. *The Economic Journal, 117*(519), C72-C92.

Brenner, J. (2010). Life-cycle variations in the association between current and lifetime earnings: Evidence for German natives and guest workers. *Labour Economics, 17*(2), 392-406. https://doi.org/10.1016/j.labeco.2009.03.006

Callister, P. (2006). Are New Zealanders heading for older age richer, better educated and more likely to be employed? In J. Boston & J. A. Davey (Eds.), *Implications of Population Ageing: Opportunities and Risks*. Wellington: Institute of Policy Studies, Victoria University of Wellington.

Callister, P., Von Randow, M., & Cotterell, G. (2011). *Women, men and the new economics of partnering in New Zealand: A research note*. Auckland: COMPASS Research Centre, University of Auckland. Retrieved from https://www.auckland.ac.nz/en/arts/our-research/research-institutes-centres-groups/compass/publications/working-paper-series.html

Carneiro, P., García, I. L., Salvanes, K. G., & Tominey, E. (2021). Intergenerational Mobility and the Timing of Parental Income. *Journal of Political Economy, 129*(3), 757-788. https://doi.org/10.1086/712443

Carter, K. N., Imlach-Gunasekara, F., McKenzie, S. K., & Blakely, T. (2012). Differential loss of participants does not necessarily cause selection bias. *Australian and New Zealand Journal of Public Health, 36*(3), 218-222. https://doi.org/10.1111/j.1753-6405.2012.00867.x

Caucutt, E. M., & Lochner, L. (2020). Early and Late Human Capital Investments, Borrowing Constraints, and the Family. *Journal of Political Economy, 128*(3), 1065-1147. https://doi.org/10.1086/704759

Caucutt, E. M., Lochner, L., & Park, Y. (2017). Correlation, Consumption, Confusion, or Constraints: Why Do Poor Children Perform so Poorly? *The Scandinavian Journal of Economics, 119*(1), 102-147. https://doi.org/10.1111/sjoe.12195

Center for Human Resource Research. (2018). *National Longitudinal Surveys: Retention & Reasons for Noninterview*. Retrieved from https://www.nlsinfo.org/content/cohorts/nlsy79/intro-to-the-sample/retention-reasons-noninterview/page/0/0/#table1

Centre for Longitudinal Studies. (2021). *Our studies*. Retrieved 25 May, 2021, from https://cls.ucl.ac.uk/cls-studies/

Chadwick, L., & Solon, G. (2002). Intergenerational Income Mobility among Daughters. *The American Economic Review, 92*(1), 335-344. https://doi.org/10.2307/3083337

Chen, W.-H., Ostrovsky, Y., & Piraino, P. (2017). Lifecycle variation, errors-in-variables bias and nonlinearities in intergenerational income transmission: new evidence from Canada. *Labour Economics, 44*, 1-12. https://doi.org/10.1016/j.labeco.2016.09.008

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics, 129*(4), 1553-1623. https://doi.org/10.1093/qje/qju022

Choi, S., Chung, I., & Breen, R. (2020). How Marriage Matters for the Intergenerational Mobility of Family Income: Heterogeneity by Gender, Life Course, and Birth Cohort. *American Sociological Review, 85*(3), 353-380. https://doi.org/10.1177/0003122420917591

Cianciolo, A. T., & Sternberg, R. J. (2004). *Intelligence: A Brief History*. Victoria, Australia: Blackwell Publishing. https://doi.org/10.1002/9780470693988

Clarkson, V. P. P. (2014). *Assessing Linkage Bias in the 1981–2006 Longitudinal Census Cohort*. Auckland: University of Auckland. Retrieved from https://www.auckland.ac.nz/en/arts/our-research/research-institutes-centres-groups/compass/social-research-projects/student-projects-compass.html

Coleman, A. (2006). *The Life-Cycle Model, Savings and Growth*. Wellington: Reserve Bank of New Zealand. Retrieved from https://www.rbnz.govt.nz/research-and-publications/seminars-and-workshops/housing-savings-and-the-household-balance-sheet

Coleman, A., & McDonald, H. (2010). *"No country for old men": a note on the trans-Tasman income divide*. Wellington: Motu Economic and Public Policy Research. Retrieved from https://www.motu.nz/our-expertise/urban-and-regional/new-zealand-and-the-world/no-country-for-old-men-a-note-on-the-trans-tasman-income-divide/

Conners, C. K. (1969). A Teacher Rating Scale for Use in Drug Studies with Children. *American Journal of Psychiatry, 126*(6), 884-888. https://doi.org/10.1176/ajp.126.6.884

Conners, C. K. (1970). Symptom Patterns in Hyperkinetic, Neurotic, and Normal Children. *Child Development, 41*(3), 667-682. https://doi.org/10.2307/1127215

Corak, M. (2004). Generational income mobility in North America and Europe: An introduction. In M. Corak (Ed.), *Generational Income Mobility in North America and Europe* (pp. 1-37). Cambridge, UK: Cambridge University Press.

Corak, M. (2006). Do Poor Children Become Poor Adults? Lessons from a Cross-Country Comparison of Generational Earnings Mobility. In J. Creedy & G. Kalb (Eds.), J. Bishop & Y. Amiel, *Dynamics of Inequality and Poverty* (pp. 143-188). Sydney: Elsevier JAI. https://doi.org/10.1016/S1049-2585(06)13006-9

Corak, M. (2013). Inequality from Generation to Generation: The United States in Comparison. In R. S. Rycroft (Ed.), *The Economics of Inequality, Poverty, and Discrimination in the 21st Century* (Vol. 1, pp. 107-123). California: ABC-CLIO, LLC.

Corak, M. (2020a). The Canadian Geography of Intergenerational Income Mobility. *The Economic Journal, 130*(631), 2134-2174. https://doi.org/10.1093/ej/uez019

Corak, M. (2020b). Intergenerational Mobility: What Do We Care About? What Should We Care About? *Australian Economic Review, 53*(2), 230-240. https://doi.org/10.1111/1467-8462.12372

Corak, M., Lindquist, M. J., & Mazumder, B. (2014). A comparison of upward and downward intergenerational mobility in Canada, Sweden and the United States. *Labour Economics, 30*, 185-200. https://doi.org/10.1016/j.labeco.2014.03.013

Cotterell, G. (2017). Social Mobility in Aotearoa New Zealand in the Neoliberal Era. In A. Bell, V. Elizabeth, T. McIntosh, & M. Wynyard (Eds.), *A Land of Milk & Honey? Making Sense of Aotearoa New Zealand* (pp. 158-169). Auckland: Auckland University Press.

Creedy, J. (1996). Earnings dynamics over the life cycle: New evidence for New Zealand. *New Zealand Economic Papers, 30*(2), 131-153. https://doi.org/10.1080/00779959609544254

Creedy, J., Gemmell, N., & Nguyen, L. (2018). Income Inequality in New Zealand, 1935–2014. *Australian Economic Review, 51*(1), 21-40. https://doi.org/10.1111/1467-8462.12240

Cunha, F., & Heckman, J. J. (2007). The Technology of Skill Formation. *The American Economic Review, 97*(2), 31-47. https://doi.org/10.1257/aer.97.2.31

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the Evidence on Life Cycle Skill Formation. In E. A. Hanushek & F. Welch (Eds.), K. J. Arrow & M. D. Intriligator, *Handbook of the Economics of Education* (Vol. 1, pp. 697-812). Sydney: Elsevier.

Davis, P. (1979). Social Mobility in New Zealand: Preliminary Results from a National Survey. *Journal of Sociology, 15*(1), 50-56. https://doi.org/10.1177/144078337901500107

De Broucker, P., & Underwood, K. (1998). Intergenerational education mobility: an international comparison with a focus on postsecondary education. *Education Quarterly Review, 5*(2), 30-51.

Dearden, L., Machin, S., & Reed, H. (1997). Intergenerational Mobility in Britain. *The Economic Journal, 107*(440), 47-66. https://doi.org/10.2307/2235270

Department of Statistics. (1977). *1976 Census of Population and Dwellings Canterbury Statistical Area Bulletin No. 10*. Wellington: Department of Statistics.

Department of Statistics. (1979). *1976 Census of Population and Dwellings Volume 2: Ages and Marital Status*. Wellington: Department of Statistics.

Department of Statistics. (1980a). *1976 Census of Population and Dwellings Volume 4: Labour Force*. Wellington: Department of Statistics.

Department of Statistics. (1980b). *1976 Census of Population and Dwellings Volume 5: Incomes and Social Security Benefits*. Wellington: Department of Statistics.

Department of Statistics. (1980c). *1976 Census of Population and Dwellings Volume 6: Education and Training*. Wellington: Department of Statistics.

Department of Statistics. (1980d). *1976 Census of Population and Dwellings Volume 7: Birthplaces and Ethnic Origin*. Wellington: Department of Statistics.

Department of Statistics. (1980e). *1976 New Zealand Census of Population and Dwellings Volume 5: Incomes and Social Security Benefits*. Wellington: Department of Statistics.

Department of Statistics. (1981). *1976 Census of Population and Dwellings Volume 10: Households, Families and Fertility*. Wellington: Department of Statistics.

Department of Statistics. (1983). *New Zealand Census of Population and Dwellings 1981 Volume 5: Income and Social Security Benefits*. Wellington: Department of Statistics.

Department of Statistics. (1985). *Population Perspectives 81: General Report* (Vol. 12). Wellington: Department of Statistics.

Department of Statistics. (1987). *1986 New Zealand Census of Population and Dwellings National Summary: Series C Report 2*. Wellington: Department of Statistics.

Deutscher, N., & Mazumder, B. (2020). Intergenerational mobility across Australia and the stability of regional estimates. *Labour Economics, 66*, 101861. https://doi.org/10.1016/j.labeco.2020.101861

Deutscher, N., & Mazumder, B. (2021). *Measuring Intergenerational Income Mobility: A Synthesis of Approaches*. Chicago, Illinois: Federal Reserve Bank of Chicago. https://doi.org/10.21033/wp-2021-09

Didham, R., & Nissen, K. (2015). Occupational Transitions: A Cohort Analysis of Transitions between Occupation Classes 1981 to 2006. *Labour, Employment and Work in New Zealand*. https://doi.org/10.26686/lew.v0i0.2201

Didham, R., Nissen, K., & Dobson, W. (2014). *Linking censuses: New Zealand longitudinal census 1981–2006*. Wellington: Statistics New Zealand. Retrieved from https://www.stats.govt.nz/methods/linking-censuses-new-zealand-longitudinal-census-19812006

Dworkin, R. (1981). What is Equality? Part 2: Equality of Resources. *Philosophy & Public Affairs, 10*(4), 283-345. https://doi.org/10.2307/2265047

Easton, B. (1996). Income distribution. In B. Silverstone, A. Bollard, & R. Lattimore (Eds.), J. Tinbergen, D. W. Jorgenson, J. J. Laffont, T. Persson, & H. K. V. Dijk, *A Study of Economic Reform: The Case of New Zealand* (pp. 101-138). Oxford: Elsevier.

Easton, B. (2013a). Economic Inequality In New Zealand: A User's Guide. *New Zealand Sociology, 28*(3), 19-66.

Easton, B. (2013b). Top Market Incomes 1981-2011. *New Zealand Sociology, 28*(2), 135-144.

Easton, B. (2014). Economic Inequality in New Zealand: Update to a User's Guide. *New Zealand Sociology, 29*(3), 7-23.

Easton, B. (2015a). Distribution of Pre-Tax Top Personal Incomes. *Policy Quarterly, 11*(1), 47-51.

Easton, B. (2015b). Top After-Tax Incomes. *Policy Quarterly, 11*(1), 52-55.

Easton, B. (2020). *Not in Narrow Seas: The Economic History of New Zealand*. Wellington: Victoria University of Wellington Press.

Elley, W. B., & Irving, J. C. (1976). Revised socio-economic index for New Zealand. *New Zealand Journal of Educational Studies, 11*(1), 25-36.

Errington, C., Cotterell, G., von Randow, M., & Milligan, S. (2008). *A guide to using data from the New Zealand Census: 1981–2006*. Wellington: Statistics New Zealand.

Esping-Andersen, G. (2004). Unequal opportunities and the mechanisms of social inheritance. In M. Corak (Ed.), *Generational Income Mobility in North America and Europe* (pp. 289-314). Cambridge, UK: Cambridge University Press.

Fairburn, M. (1979). Social Mobility and Opportunity in Nineteenth-Century New Zealand. *New Zealand Journal of History, 13*(1), 43-64.

Fawcett, J. (2005). *Socioeconomic Variation in Mortality in New Zealand 1981-1999: Cross-National And Times Series Comparisons.* University of Otago, Wellington. Retrieved from https://www.otago.ac.nz/wellington/otago025443.pdf

Fawcett, J., Atkinson, J., Herd, P., & Blakely, T. (2008). *Record Linkage of Census and Mortality 2001-04 Records: New Zealand Census-Mortality Study Technical Report No.6.* Wellington: University of Otago. Retrieved from https://www.otago.ac.nz/wellington/otago024510.pdf

Fergusson, D. M., & Horwood, L. J. (1993). The Structure, Stability and Correlations of the Trait Components of Conduct Disorder, Attention Deficit and Anxiety/Withdrawal Reports. *Journal of Child Psychology and Psychiatry, 34*(5), 749-766. https://doi.org/10.1111/j.1469-7610.1993.tb01069.x

Fergusson, D. M., & Horwood, L. J. (2001). The Christchurch Health and Development Study: review of findings on child and adolescent mental health. *Australian and New Zealand Journal of Psychiatry, 35*(3), 287-296. https://doi.org/10.1046/j.1440-1614.2001.00902.x

Fergusson, D. M., & Horwood, L. J. (2013). Christchurch Health and Development Study. In P. Joyce, G. Nicholls, K. Thomas, & T. Wilkinson (Eds.), *The Christchurch Experience: 40 Years of Research and Teaching* (pp. 79-87). Christchurch: University of Otago.

Fergusson, D. M., Horwood, L. J., & Boden, J. M. (2008). The transmission of social inequality: Examination of the linkages between family socioeconomic status in childhood and educational achievement in young adulthood. *Research in Social Stratification and Mobility, 26*(3), 277-295. https://doi.org/10.1016/j.rssm.2008.05.001

Fergusson, D. M., Horwood, L. J., & Lloyd, M. (1991). Confirmatory Factor Models of Attention Deficit and Conduct Disorder. *Journal of Child Psychology and Psychiatry, 32*(2), 257-274. https://doi.org/10.1111/j.1469-7610.1991.tb00305.x

Fergusson, D. M., Horwood, L. J., & Ridder, E. M. (2005). Show me the child at seven II: childhood intelligence and later outcomes in adolescence and young adulthood. *Journal of Child Psychology and Psychiatry, 46*(8), 850-858. https://doi.org/10.1111/j.1469-7610.2005.01472.x

Fergusson, D. M., Horwood, L. J., Shannon, F. T., & Lawton, J. M. (1989). The Christchurch Child Development Study: a review of epidemiological findings. *Paediatric and Perinatal Epidemiology, 3*(3), 302-325. https://doi.org/10.1111/j.1365-3016.1989.tb00382.x
Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*(1), 29-51.
Flynn, J. R. (1987). Massive IQ Gains in 14 Nations: What IQ Tests Really Measure. *Psychological Bulletin, 101*(2), 171-191.
Flynn, J. R., & Sternberg, R. J. (2020). Environment and Intelligence. In R. J. Sternberg (Ed.), *Human Intelligence: An Introduction* (pp. 253-278). Melbourne: Cambridge University Press. https://doi.org/10.1017/9781108610636
Gibb, S. J., Fergusson, D. M., & Horwood, L. J. (2012). Childhood family income and life outcomes in adulthood: Findings from a 30-year longitudinal study in New Zealand. *Social Science & Medicine, 74*(12), 1979-1986. https://doi.org/10.1016/j.socscimed.2012.02.028
Gibbons, M. (2010). *Income and Occupational Intergenerational Mobility in New Zealand*. Wellington: New Zealand Treasury. Retrieved from http://www.treasury.govt.nz/publications/research-policy/wp/2010/10-06
Gibbons, M. (2011). Intergenerational Economic Mobility in New Zealand. *Policy Quarterly, 7*(2), 53-60.
Gilmore, A., Croft, C., & Reid, N. A. (1981). *Burt Word Reading Test, New Zealand Revision: Teachers Manual*. Wellington: New Zealand Council for Educational Research.
Gouskova, E., Chiteji, N., & Stafford, F. (2010). Estimating the intergenerational persistence of lifetime earnings with life course matching: Evidence from the PSID. *Labour Economics, 17*(3), 592-597. https://doi.org/10.1016/j.labeco.2009.04.009
Grawe, N. D. (2006). Lifecycle bias in estimates of intergenerational earnings persistence. *Labour Economics, 13*(5), 551-570. https://doi.org/10.1016/j.labeco.2005.04.002
Gregg, P., Jonsson, J. O., Macmillan, L., & Mood, C. (2017). The Role of Education for Intergenerational Income Mobility: A comparison of the United States, Great Britain, and Sweden. *Social Forces, 96*(1), 121-152. https://doi.org/10.1093/sf/sox051
Gregg, P., Macmillan, L., & Vittori, C. (2017). Moving Towards Estimating Sons' Lifetime Intergenerational Economic Mobility in the UK. *Oxford Bulletin of Economics and Statistics, 79*(1), 79-100. https://doi.org/10.1111/obes.12146
Grigorenko, E. L. (2000). Heritability and Intelligence. In R. J. Sternberg (Ed.), *Handbook of Intelligence* (pp. 53-91). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511807947.005
Grönqvist, E., Öckert, B., & Vlachos, J. (2017). The Intergenerational Transmission of Cognitive and Noncognitive Abilities. *Journal of Human Resources, 52*(4), 887-918. https://doi.org/10.3368/jhr.52.4.0115-6882R1
Haider, S., & Solon, G. (2006). Life-Cycle Variation in the Association between Current and Lifetime Earnings. *The American Economic Review, 96*(4), 1308-1320. https://doi.org/10.1257/aer.96.4.1308
Hall, B., Thorns, D., & Willmott, B. (1983). *Community Formation and Change: A Study of Rural and Urban Localities in New Zealand*. Christchurch: University of Canterbury.
Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2017). Coping with change: International differences in the returns to skills. *Economics Letters, 153*, 15-19. https://doi.org/10.1016/j.econlet.2017.01.007
Hausman, J. (2001). Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left. *Journal of Economic Perspectives, 15*(4), 57-67. https://doi.org/10.1257/jep.15.4.57
Heckman, J. J., & Landersø, R. (2021). Lessons for Americans from Denmark about inequality and social mobility. *Labour Economics*, 101999. https://doi.org/10.1016/j.labeco.2021.101999
Heckman, J. J., & Mosso, S. (2014). The Economics of Human Development and Social Mobility. *Annual Review of Economics, 6*(1), 689-733. https://doi.org/10.1146/annurev-economics-080213-040753
Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics, 24*(3), 411-482. https://doi.org/10.1086/504455
Helsø, A.-L. (2021). Intergenerational Income Mobility in Denmark and the United States. *The Scandinavian Journal of Economics, 123*(2), 508-531. https://doi.org/10.1111/sjoe.12420

Hernandez, J. (2019). *In Fairness to Our Schools: Better measures for better outcomes*. Wellington: The New Zealand Initiative. Retrieved from https://nzinitiative.org.nz/reports-and-media/reports/in-fairness-to-our-schools/

Hertz, T. (2005). Rags, Riches, and Race: The Intergenerational Economic Mobility of Black and White Families in the United States. In S. Bowles, H. Gintis, & M. O. Groves (Eds.), *Unequal Chances: Family Background and Economic Success* (pp. 165-191). New York: Russell Sage Foundation.

Hertz, T. (2008). A group-specific measure of intergenerational persistence. *Economics Letters, 100*(3), 415-417. https://doi.org/10.1016/j.econlet.2008.03.008

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., & Verashchagina, A. (2007). The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends. *The B.E. Journal of Economic Analysis & Policy, 7*(2). https://doi.org/10.2202/1935-1682.1775

Hirvonen, L. H. (2008). Intergenerational Earnings Mobility Among Daughters and Sons: Evidence from Sweden and a Comparison with the United States. *American Journal of Economics and Sociology, 67*(5), 777-826. https://doi.org/10.1111/j.1536-7150.2008.00598.x

Hirvonen, L. H. (2010). *Accounting for Intergenerational Earnings Persistence: Can we Distinguish Between Education, Skills, and Health?* Stockholm: Swedish Institute for Social Research (SOFI), Stockholm University. Retrieved from https://www.sofi.su.se/english/research/dissertations-and-publications/sofi-working-papers

Hughes, D., & Lauder, H. (1990). Public Examinations and the Structuring of Inequality: A Case Study. In H. Lauder & C. Wylie (Eds.), *Towards Successful Schooling* (pp. 151-173). Hampshire: The Falmer Press.

Hughes, D., & Lauder, H. (1991). Human Capital Theory and the Wastage of Talent in New Zealand. *New Zealand Journal of Educational Studies, 26*(1), 5-20.

Hyslop, D. (2000). *A Preliminary Analysis of the Dynamics of Individual Market and Disposable Incomes*. Wellington: New Zealand Treasury. Retrieved from https://treasury.govt.nz/publications/wp/preliminary-analysis-dynamics-individual-market-and-disposable-incomes-wp-00-15

Hyslop, D., & Maré, D. (2005). Understanding New Zealand's Changing Income Distribution, 1983-1998: A Semi-Parametric Analysis. *Economica, 72*(287), 469-495.

Hyslop, D., Maré, D., & Timmins, J. (2003). *Qualifications, Employment and the Value of Human Capital, 1986-2001*. Wellington: New Zealand Treasury. Retrieved from https://www.treasury.govt.nz/publications/wp/qualifications-employment-and-value-human-capital-1986-2001-wp-03-35-html

Hyslop, D., Rice, A., & Skilling, H. (2019). *Understanding labour market developments in New Zealand, 1986–2017*. Wellington: Reserve Bank of New Zealand. Retrieved from https://www.rbnz.govt.nz/research-and-publications/discussion-papers/2019/dp2019-02

Hyslop, D., & Yahanpath, S. (2006). Income Growth and Earnings Variations in New Zealand, 1998–2004. *Australian Economic Review, 39*(3), 293-311. https://doi.org/10.1111/j.1467-8462.2006.00419.x

James, C. (2020). Social Laboratory: Reality or Myth? In B. Evan & K. Girol (Eds.), *Public Policy and Governance Frontiers in New Zealand* (Vol. 32, pp. 27-52): Emerald Publishing Limited. https://doi.org/10.1108/S2053-769720200000032022

Jäntti, M., Bratsberg, B., Røed, K., Raaum, O., Naylor, R., Österbacka, E., . . . Eriksson, T. (2006). *American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States*. Bonn, Germany: Institute for the Study of Labor. Retrieved from https://ftp.iza.org/dp1938.pdf

Jäntti, M., & Jenkins, S. P. (2015). Income Mobility. In A. B. Atkinson & F. Bourguignon (Eds.), *Handbook of Income Distribution* (Vol. 2A, pp. 807-935): Elsevier. https://doi.org/10.1016/B978-0-444-59428-0.00011-4

Jäntti, M., & Lindahl, L. (2012). On the variability of income within and across generations. *Economics Letters, 117*(1), 165-167. https://doi.org/10.1016/j.econlet.2012.05.007

Jenkins, S. P. (1987). Snapshots versus movies. *European Economic Review, 31*(5), 1149-1158. https://doi.org/10.1016/S0014-2921(87)80010-7

Jenkins, S. P., & Van Kerm, P. (2009). The Measurement of Economic Inequality. In W. Salverda, B. Nolan, & T. M. Smeeding (Eds.), *The Oxford Handbook of Economic Inequality* (pp. 40-67). Auckland: Oxford University Press.

Jones, F. L., & Davis, P. (1986). *Models of Society: Class, Stratification and Gender in Australia and New Zealand*. Sydney: Croom Helm.

Jones, F. L., & Davis, P. (1988). Class Structuration and Patterns of Social Closure in Australia and New Zealand. *Sociology, 22*(2), 271-291. https://doi.org/10.1177/0038038588022002007

Kalb, G., & Scutella, R. (2003). *New Zealand Labour Supply from 1991-2001: An Analysis Based on a Discrete Choice Structural Utility Model*. Wellington: New Zealand Treasury. Retrieved from https://www.treasury.govt.nz/publications/wp/new-zealand-labour-supply-1991-2001-analysis-based-discrete-choice-structural-utility-model-wp-03-23

Karacaoglu, G. (2015). *The New Zealand Treasury's Living Standards Framework - Exploring a Stylised Mode*. Wellington: New Zealand Treasury. Retrieved from https://www.treasury.govt.nz/publications/wp/new-zealand-treasurys-living-standards-framework-exploring-stylised-model-wp-15-12-html

Knapp, M., King, D., Healey, A., & Thomas, C. (2011). Economic outcomes in adulthood and their associations with antisocial conduct, attention deficit and anxiety problems in childhood. *Journal of Mental Health Policy and Economics, 14*(3), 137-147.

Landersø, R., & Heckman, J. J. (2017). The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US. *The Scandinavian Journal of Economics, 119*(1), 178-230. https://doi.org/10.1111/sjoe.12219

Lauder, H., & Hughes, D. (1990a). Social Inequalities and Differences in School Outcomes. *New Zealand Journal of Educational Studies, 25*(1), 37-60.

Lauder, H., & Hughes, D. (1990b). Social Origins, Destinations and Educational Inequality. In J. Codd, R. Harker, & R. Nash (Eds.), *Political Issues in New Zealand Education* (Second ed., pp. 43-60). Palmerston North: The Dunmore Press.

Lee, H., & Lee, J.-W. (2021). Patterns and determinants of intergenerational educational mobility: Evidence across countries. *Pacific Economic Review, 26*(1), 70-90. https://doi.org/10.1111/1468-0106.12342

Lee, S. Y., & Seshadri, A. (2019). On the Intergenerational Transmission of Economic Status. *Journal of Political Economy, 127*(2), 855-921. https://doi.org/10.1086/700765

Lefgren, L., Lindquist, M. J., & Sims, D. (2012). Rich Dad, Smart Dad: Decomposing the Intergenerational Transmission of Income. *Journal of Political Economy, 120*(2), 268-303. https://doi.org/10.1086/666590

Lefranc, A. (2018). Intergenerational Earnings Persistence and Economic Inequality in the Long Run: Evidence from French Cohorts, 1931–75. *Economica, 85*(340), 808-845. https://doi.org/10.1111/ecca.12269

Loomis, T. (1985). *Occupational Mobility, Social Class and Networks among Cook Islanders in New Zealand*. Auckland: The Social Research & Development Trust.

Low, F., Gluckman, P., & Poulton, R. (2021). *Intergenerational Disadvantage: Why Maternal Mental Health Matters*. Auckland: Koi Tū: The Centre for Informed Futures, University of Auckland. Retrieved from https://informedfutures.org/intergenerational-disadvantage-why-maternal-mental-health-matters/

Lundborg, P., & Stenberg, A. (2010). Nature, nurture and socioeconomic policy—What can we learn from molecular genetics? *Economics & Human Biology, 8*(3), 320-330. https://doi.org/10.1016/j.ehb.2010.08.002

Maani, S. (1999). *Private and Public Returns to Investments in Secondary and Higher Education in New Zealand Over Time: 1981-1996*. Wellington: New Zealand Treasury. Retrieved from https://www.treasury.govt.nz/publications/wp/private-and-public-returns-investments-secondary-and-higher-education-new-zealand-over-time-1981-1996-wp-99-02

Maani, S., & Kalb, G. (2007). Academic performance, childhood economic resources, and the choice to leave school at age 16. *Economics of Education Review, 26*(3), 361-374. https://doi.org/10.1016/j.econedurev.2005.12.002

Maani, S., & Maloney, T. (2004). *Returns to Post-school Qualifications: New Evidence Based on the HLFS Income Supplement (1997-2002)*. Auckland: New Zealand Department of Labour. Retrieved from https://www.oecd.org/social/labour/37392513.pdf

Maloney, T. (2004). Are the Outcomes of Young Adults Linked to the Family Income Experienced In Childhood? *Social Policy Journal of New Zealand,* (22), 55-82.

Maloney, T., Maani, S., & Pacheco, G. (2003). Intergenerational Welfare Participation in New Zealand. *Australian Economic Papers, 42*(3), 346-362. https://doi.org/10.1111/1467-8454.00203

Mandelman, S. D., & Grigorenko, E. L. (2011). Intelligence: Genes, Environments, and Their Interactions. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge Handbook of Intelligence* (pp. 85-106). Melbourne: Cambridge University Press. https://doi.org/10.1017/CBO9780511977244.006

Maré, D. (2018). The labor market in New Zealand, 2000−2017. *IZA World of Labor,* (427). https://doi.org/10.15185/izawol.427
Maré, D. (2019). *Occupational drift in New Zealand: 1976-2018*. Wellington: Motu Economic and Public Policy Research. Retrieved from https://www.motu.nz/our-research/population-and-labour/individual-and-group-outcomes/occupational-drift-in-new-zealand-1976-2018/
Maré, D., Morten, M., & Stillman, S. (2007). Settlement patterns and the geographic mobility of recent migrants to New Zealand. *New Zealand Economic Papers, 41*(2), 163-195. https://doi.org/10.1080/00779950709558508
Maré, D., & Stillman, S. (2010). *Passing It On: The intergenerational transmission of human capital in New Zealand families*. Wellington: Families Commission. Retrieved from http://motu.nz/our-work/population-and-labour/individual-and-group-outcomes/passing-it-on-the-intergenerational-transmission-of-human-capital-in-new-zealand-families/
Martin, B. (1997a). *Away From Equality: Change in Personal Incomes, 1951 to 1991*. Hamilton: University of Waikato.
Martin, B. (1997b). *Income Trends Among Individuals and Families, 1976 to 1996*. Hamilton: University of Waikato.
Martin, B. (1999). *Incomes of Individuals and Families in New Zealand, 1951–1991*. University of Waikato, Hamilton.
Martin, B. (2002). Change in the distribution of income of the working-age population, 1976–1996. *Labour Market Bulletin*, 132-156.
Masgoret, A.-M., Merwood, P., & Tausi, M. (2009). *New Faces, New Futures: New Zealand. Findings from the Longitudinal Immigration Survey: New Zealand (LisNZ) – Wave One*. Wellington: Department of Labour. Retrieved from https://www.mbie.govt.nz/dmsdocument/2726-new-faces-new-futures-2-background-characteristics-pdf
Mazumder, B. (2003). *Revised Estimates of Intergenerational Income Mobility in the United States*. Chicago: Federal Reserve Bank of Chicago. Retrieved from https://www.chicagofed.org/publications/working-papers/2003/2003-16
Mazumder, B. (2005). Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data. *The Review of Economics and Statistics, 87*(2), 235-255. https://doi.org/10.2307/40042900
Mazumder, B. (2016). Estimating the Intergenerational Elasticity and Rank Association in the United States: Overcoming the Current Limitations of Tax Data. In L. Cappellari, S. W. Polachek, & K. Tatsiramos (Eds.), S. W. Polachek & K. Tatsiramos, *Inequality: Causes and Consequences* (pp. 83-129): Emerald Group Publishing Limited. https://doi.org/10.1108/S0147-912120160000043012
Mendolia, S., & Siminski, P. (2017). Is education the mechanism through which family background affects economic outcomes? A generalised approach to mediation analysis. *Economics of Education Review, 59*, 1-12. https://doi.org/10.1016/j.econedurev.2017.05.001
Mercante, J., & Mok, P. (2014a). *Estimation of Labour Supply in New Zealand*. Wellington: New Zealand Treasury. Retrieved from https://www.treasury.govt.nz/publications/wp/estimation-labour-supply-new-zealand-wp-14-08
Mercante, J., & Mok, P. (2014b). *Estimation of wage equations for New Zealand*. Wellington: New Zealand Treasury. Retrieved from https://www.treasury.govt.nz/publications/wp/estimation-wage-equations-new-zealand-html
Mitnik, P. A. (2019). *Estimating the Intergenerational Elasticity of Expected Income with Short-Run Income Measures: A Generalized Error-in-Variables Model*. Stanford: Stanford Center on Poverty and Inequality, Stanford University. Retrieved from https://inequality.stanford.edu/publications/media/details/estimating-intergenerational-elasticity-expected-income-short-run-0
Mitnik, P. A., Bryant, V., Weber, M., & Grusky, D. B. (2015). *New Estimates of Intergenerational Mobility Using Administrative Data*. Washington, D.C.: Statistics of Income Division, Internal Revenue Service. Retrieved from https://inequality.stanford.edu/publications/media/details/new-estimates-intergenerational-economic-mobility-using-administrative
Mitnik, P. A., & Grusky, D. B. (2020). The Intergenerational Elasticity of What? The Case for Redefining the Workhorse Measure of Economic Mobility. *Sociological Methodology, 50*(1), 47-95. https://doi.org/10.1177/0081175019886613

Mogstad, M. (2017). The Human Capital Approach to Intergenerational Mobility. *Journal of Political Economy, 125*(6), 1862-1868. https://doi.org/10.1086/694633

Moisio, P., Lorentzen, T., Bäckman, O., Angelin, A., Salonen, T., & Kauppinen, T. (2015). Trends in the Intergenerational Transmission of Social Assistance in the Nordic Countries in the 2000s. *European Societies, 17*(1), 73-93. https://doi.org/10.1080/14616696.2014.968797

Mood, C., Jonsson, J. O., & Bihagen, E. (2012). Socioeconomic Persistence Across Generations: Cognitive and Noncognitive Processes. In J. Ermisch, M. Jäntti, & T. Smeeding (Eds.), *From Parents to Children: The Intergenerational Transmission of Advantage* (pp. 53-83). New York: Russell Sage Foundation.

Morton, S. M. B., Atatoa Carr, P. E., Berry, S. D., Grant, C. C., Bandara, D. K., Mohal, J., & Tricker, P. J. (2014). *Growing Up in New Zealand Residential Mobility Report 1: Moving house in the first 1000 days*. Auckland: University of Auckland. Retrieved from https://www.growingup.co.nz/growing-reports

Morton, S. M. B., Walker, C. G., Gerritsen, S., Smith, A., Cha, J., Atatoa Carr, P. E., . . . Wall, C. (2020). *Now We Are Eight: Life in Middle Childhood*. Auckland: Growing Up In New Zealand. Retrieved from https://www.growingup.co.nz/growing-reports

Muller, S. M. (2010). Another problem in the estimation of intergenerational income mobility. *Economics Letters, 108*(3), 291-295. https://doi.org/10.1016/j.econlet.2010.05.018

Mulligan, C. B. (1999). Galton versus the Human Capital Approach to Inheritance. *Journal of Political Economy, 107*(S6), S184-S224. https://doi.org/10.1086/250108

Munk, M. D., Bonke, J., & Hussain, M. A. (2016). Intergenerational top income persistence: Denmark half the size of Sweden. *Economics Letters, 140*(Supplement C), 31-33. https://doi.org/10.1016/j.econlet.2015.11.038

Murray, C., Clark, R. G., Mendolia, S., & Siminski, P. (2018). Direct Measures of Intergenerational Income Mobility for Australia. *Economic Record, 94*(307), 445-468. https://doi.org/10.1111/1475-4932.12445

Nam, J. (2021). Does Economic Inequality Constrain Intergenerational Economic Mobility? The Association Between Income Inequality During Childhood and Intergenerational Income Persistence in the United States. *Social Indicators Research, 154*(2), 469-488. https://doi.org/10.1007/s11205-020-02579-2

Narayan, A., Van der Weide, R., Cojocaru, A., Lakner, C., Redaelli, S., Gerszon Mahler, D., . . . Thewissen, S. (2018). *Fair Progress? Economic Mobility across Generations around the World*. Washington: International Bank for Reconstruction and Development/The World Bank. Retrieved from http://www.worldbank.org/en/topic/poverty/publication/fair-progress-economic-mobility-across-generations-around-the-world

Nash, R. (1993). *Succeeding Generations: Family Resources and Access to Education in New Zealand*. Auckland: Oxford University Press.

Nash, R., Harker, R., & Charters, H. (1990). Reproduction and Renewal through Education. In J. Codd, R. Harker, & R. Nash (Eds.), *Political Issues in New Zealand Education* (pp. 61-73). Palmerston North: Dunmore Press.

Nel, P. (2021). Mobility rules: why New Zealanders oppose redistribution. *Kōtuitui: New Zealand Journal of Social Sciences Online*, 1-20. https://doi.org/10.1080/1177083X.2021.1912121

New Zealand Treasury. (2018). *Living Standards Framework: Background and Future Work*. Wellington: New Zealand Treasury. Retrieved from https://www.treasury.govt.nz/publications/tp/living-standards-framework-background-and-future-work

Nilsen, Ø. A., Vaage, K., Aakvik, A., & Jacobsen, K. (2012). Intergenerational Earnings Mobility Revisited: Estimates Based on Lifetime Earnings. *The Scandinavian Journal of Economics, 114*(1), 1-23. https://doi.org/10.1111/j.1467-9442.2011.01672.x

Nolan, M. (2007). The Reality and Myth of New Zealand Egalitarianism: Explaining the Pattern of a Labour Historiography at the Edge Of Empires. *Labour History Review, 72*(2), 113-134. https://doi.org/10.1179/174581807X224560

Nolan, M. (2009). Constantly on the Move, but Going Nowhere? Work, Community and Social Mobility. In G. Byrnes (Ed.), *The New Oxford History of New Zealand* (pp. 357-388). Auckland: Oxford University Press.

Nybom, M., & Stuhler, J. (2016). Heterogeneous Income Profiles and Lifecycle Bias in Intergenerational Mobility Estimation. *Journal of Human Resources, 51*(1), 239-268. https://doi.org/10.3368/jhr.51.1.239

Nybom, M., & Stuhler, J. (2017). Biases in Standard Measures of Intergenerational Income Dependence. *Journal of Human Resources, 52*(3), 800-825. https://doi.org/10.3368/jhr.52.3.0715-7290R

O'Dea, D. (2000). *The Changes in New Zealand's Income Distribution*. Wellington: New Zealand Treasury. Retrieved from https://treasury.govt.nz/publications/wp/changes-new-zealands-income-distribution-wp-00-13

OECD. (2008). *Growing Unequal?: Income Distribution and Poverty in OECD Countries*. Paris: OECD Publishing. https://doi.org/10.1787/9789264044197-en

OECD. (2018). *A Broken Social Elevator? How to Promote Social Mobility*. Paris: OECD Publishing. https://doi.org/10.1787/9789264301085-en

Olssen, E., Griffen, C., & Jones, F. (2011). *An Accidental Utopia? Social mobility & the foundations of an egalitarian society, 1880-1940*. Dunedin: Otago University Press.

Österbacka, E. (2001). Family Background and Economic Status in Finland. *Scandinavian Journal of Economics, 103*(3), 467-484. https://doi.org/10.1111/1467-9442.00255

Österberg, T. (2000). Intergenerational Income Mobility in Sweden: What Do Tax-Data Show? *Review of Income and Wealth, 46*(4), 421-436. https://doi.org/10.1111/j.1475-4991.2000.tb00409.x

Pacheco, G., & Maloney, T. (2003). Are the Determinants of Intergenerational Welfare Dependency Gender-Specific? *Australian Journal of Labour Economics, 6*(3), 371-382.

Park, H. (2008). Home literacy environments and children's reading performance: a comparative study of 25 countries. *Educational Research and Evaluation, 14*(6), 489-505. https://doi.org/10.1080/13803610802576734

Parker, R. N., & Fenwick, R. (1983). The Pareto Curve and Its Utility for Open-Ended Income Distributions in Survey Research. *Social Forces, 61*(3), 872-885. https://doi.org/10.2307/2578140

Pearson, D. G. (1980a). *Johnsonville: Continuity and Change in a New Zealand Township*. Sydney: George Allen & Unwin.

Pearson, D. G. (1980b). Small-Town Capitalism and Stratification in New Zealand 1880-1930: A Case Study. *New Zealand Journal of History, 14*(2), 107-131.

Pearson, D. G., & Thorns, D. C. (1983). *Eclipse of Equality: Social Stratification in New Zealand*. Sydney: George Allen & Unwin.

Pearson, D. G., & Thorns, D. C. (1986). A Tale of Two Cities: Marriage and Mobility in New Zealand. *The Australian and New Zealand Journal of Sociology, 22*(2), 208-224. https://doi.org/10.1177/144078338602200203

Pekkarinen, T., Uusitalo, R., & Kerr, S. (2009). School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform. *Journal of Public Economics, 93*(7), 965-973. https://doi.org/10.1016/j.jpubeco.2009.04.006

Perry, B. (2019). *Household incomes in New Zealand: Trends in indicators of inequality and hardship 1982 to 2018*. Wellington: Ministry of Social Development. Retrieved from https://www.msd.govt.nz/about-msd-and-our-work/publications-resources/monitoring/household-incomes/household-incomes-1982-to-2018.html

Pfeffer, F. T. (2008). Persistent Inequality in Educational Attainment and its Institutional Context. *European Sociological Review, 24*(5), 543-565. https://doi.org/10.1093/esr/jcn026

Piketty, T. (2000). Theories of Persistent Inequality and Intergenerational Mobility. In A. B. Atkinson & F. Bourguignon (Eds.), K. J. Arrow & M. D. Intriligator, *Handbook of Income Distribution* (Vol. 1, pp. 429-476). Sydney: Elsevier.

Plomin, R. (2019). *Blueprint: How DNA Makes Us Who We Are*. Cambridge, MA: MIT Press.

Poot, J., & Stillman, S. (2016). Skill composition of immigration flows and the measurement of education-occupation mismatch. *IZA Journal of Migration, 5*(1), 18. https://doi.org/10.1186/s40176-016-0066-y

Poulton, R., Hancox, R., Milne, B., Baxter, J., Scott, K., & Wilson, N. (2006). The Dunedin Multidisciplinary Health and Development Study: are its findings consistent with the overall New Zealand population? *The New Zealand Medical Journal, 119*(1235), 45-55.

Raaum, O., Bratsberg, B., Røed, K., Österbacka, E., Eriksson, T., Jäntti, M., & Naylor, R. A. (2007). Marital Sorting, Household Labor Supply, and Intergenerational Earnings Mobility across Countries. *The B.E. Journal of Economic Analysis & Policy, 7*(2), 1-46.

Rashbrooke, M. (2013). Inequality and New Zealand. In M. Rashbrooke (Ed.), *Inequality: A New Zealand Crisis* (pp. 20-34). Wellington: Bridget Williams Books.

Rashbrooke, M. (2021). *Too Much Money: How Wealth Disparities Are Unbalancing Aotearoa New Zealand*. Wellington: Bridget Williams Books.

Reuben, A., Caspi, A., Belsky, D. W., Broadbent, J., Harrington, H., Sugden, K., . . . Moffitt, T. E. (2017). Association of Childhood Blood Lead Levels With Cognitive Function and Socioeconomic Status at Age 38 Years and With IQ Change and Socioeconomic Mobility Between Childhood and Adulthood. *JAMA, 317*(12), 1244-1251. https://doi.org/10.1001/jama.2017.1712

Robb, J. H., & Cloud, J. (1970). Occupational Mobility in a New Zealand Provincial Borough. *The Australian and New Zealand Journal of Sociology, 6*(1), 49-51. https://doi.org/10.1177/144078337000600105
Robertson, O., Nathan, K., Howden-Chapman, P., Baker, M. G., Atatoa Carr, P., & Pierse, N. (2021). Residential mobility for a national cohort of New Zealand-born children by area socioeconomic deprivation level and ethnic group. *BMJ Open, 11*(1), e039706. https://doi.org/10.1136/bmjopen-2020-039706
Roemer, J. E. (2004). Equal opportunity and intergenerational mobility: Going beyond intergenerational income transition matrices. In M. Corak (Ed.), *Generational Income Mobility in North America and Europe* (pp. 48-57). Cambridge, UK: Cambridge University Press.
Roemer, J. E. (2012). What is the Justification of Studying Intergenerational Mobility of Socioeconomic Status? In J. Ermisch, M. Jäntti, & T. Smeeding (Eds.), *From Parents to Children: The Intergenerational Transmission of Advantage* (pp. 482-487). New York: Russell Sage Foundation.
Rohenkohl, B. (2019). *Intergenerational Income Mobility in the UK: New evidence using the BHPS and Understanding Society*. Sheffield: University of Sheffield. Retrieved from https://www.sheffield.ac.uk/economics/research/serps
Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern Epidemiology* (3rd ed.). Sydney: Lippincott Williams & Wilkins.
Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behaviour*. London: Longman.
Schnitzlein, D. D. (2016). A New Look at Intergenerational Mobility in Germany Compared to the U.S. *Review of Income and Wealth, 62*(4), 650-667. https://doi.org/10.1111/roiw.12191
Singhal, R. (2015). *Adjusting for linkage bias in the New Zealand Longitudinal Census*. Auckland: University of Auckland. Retrieved from https://cdn.auckland.ac.nz/assets/auckland/arts/our-research/research-institutes-centres-groups/compass/social-research-projects/docs/nzlc-bias-weights-report.pdf
Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (2nd ed.). London: Sage Publications.
Solon, G. (1989). Biases in the Estimation of Intergenerational Earnings Correlations. *The Review of Economics and Statistics, 71*(1), 172-174. https://doi.org/10.2307/1928066
Solon, G. (1992). Intergenerational Income Mobility in the United States. *The American Economic Review, 82*(3), 393-408. https://doi.org/10.2307/2117312
Solon, G. (1999). Intergenerational Mobility in the Labor Market. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 3A, Part 6, pp. 1761-1800). Oxford: Elsevier. https://doi.org/10.1016/S1573-4463(99)03010-2
Solon, G. (2002). Cross-Country Differences in Intergenerational Earnings Mobility. *The Journal of Economic Perspectives, 16*(3), 59-66. https://doi.org/10.2307/3216950
Solon, G. (2004). A model of intergenerational mobility variation over time and place. In M. Corak (Ed.), *Generational Income Mobility in North America and Europe* (pp. 38-47). Cambridge, UK: Cambridge University Press.
Solon, G. (2008). Intergenerational Income Mobility. In S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics: Volume 1 – 8* (pp. 3128-3131). London: Palgrave Macmillan UK. Retrieved from https://link.springer.com/referenceworkentry/10.1057%2F978-1-349-95121-5_2081-1
Statistics New Zealand. (1995). *New Zealand Standard Classification of Occupations, 1995*. Wellington: Statistics New Zealand.
Statistics New Zealand. (1999). *New Zealand Now - Incomes: 1999 Edition*. Wellington: Statistics New Zealand. Retrieved from https://cdm20045.contentdm.oclc.org/digital/collection/p20045coll18/id/309/rec/68
Statistics New Zealand. (2001). *New Zealand Standard Classification of Occupations, 1999*. Wellington: Statistics New Zealand.
Statistics New Zealand. (2002). *2001 Census of Population and Dwellings: Incomes*. Wellington: Statistics New Zealand.
Statistics New Zealand. (2009). *Statistical Standard for Income Bands*. Wellington: Statistics New Zealand.
Statistics New Zealand. (2014). *2013 Census data user guide*. Wellington: Statistics New Zealand. Retrieved from https://www.stats.govt.nz/methods/2013-census-data-user-guide
Stats NZ. (2021a). *Infoshare*. Retrieved from https://www.stats.govt.nz/tools/stats-infoshare
Stats NZ. (2021b). *NZ.Stat*. Retrieved from http://nzdotstat.stats.govt.nz/wbos/Index.aspx

Stats NZ. (2021c). *Stats NZ DataInfo+*. Retrieved from https://datainfoplus.stats.govt.nz/
Sternberg, R. J., & Grigorenko, E. L. (1999). Myths in Psychology and Education Regarding the Gene-Environment Debate. *Teachers College Record, 100*(3), 536–553.
Suei, S. (2016). *Comparing income information from census and administrative sources*. Wellington: Statistics New Zealand. Retrieved from https://www.stats.govt.nz/research/comparing-income-information-from-census-and-administrative-sources
Swift, A. (2004). Would Perfect Mobility be Perfect? *European Sociological Review, 20*(1), 1-11. https://doi.org/10.1093/esr/20.1.1
Swift, A. (2005). Justice, Luck, and the Family: The Intergenerational Transmission of Economic Advantage from a Normative Perspective. In S. Bowles, H. Gintis, & M. O. Groves (Eds.), *Unequal Chances: Family Background and Economic Success* (pp. 256-276). New York: Russell Sage Foundation.
Ueda, A. (2009). Intergenerational Mobility of Earnings and Income in Japan. *The B.E. Journal of Economic Analysis & Policy, 9*(1). https://doi.org/10.2202/1935-1682.2203
Violato, M., Petrou, S., Gray, R., & Redshaw, M. (2011). Family income and child cognitive and behavioural development in the United Kingdom: does money matter? *Health Economics, 20*(10), 1201-1225. https://doi.org/10.1002/hec.1665
Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children - Revised*. New York: The Psychological Corporation.
Wilkes, C. (1994). Class. In P. Spoonley, D. Pearson, & I. Shirley (Eds.), *New Zealand Society: A Sociological Introduction* (pp. 66-80). Palmerston North: Dunmore Press.
Zuccollo, J., Maani, S., Kaye-Blake, B., & Zeng, L. (2013). *Private Returns to Tertiary Education: How Does New Zealand Compare to the OECD?* Wellington: New Zealand Treasury. Retrieved from https://www.treasury.govt.nz/publications/wp/private-returns-tertiary-education-how-does-new-zealand-compare-wp-13-10-html

## Appendices

**Appendix Table 1**: *Total personal income classifications and representative values assigned to income bands for 1981 to 2013 censuses*.

| **Census:** | **1981 census** | | | **1986 census** | | | **1991 census** | | | **1996 census** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Source from which medians were estimated[1]:** | **None (mid-points used)** | | | **Household Economic Survey** | | | **Household Economic Survey** | | | **Household Economic Survey** | | |
| | **Classification[2]** | **Mid-point ($)** | **CPI-adjusted mid-point (2012 Q3 $)** | **Classification** | **Median ($)** | **CPI-adjusted median (2012 Q3 $)** | **Classification** | **Median ($)** | **CPI-adjusted median (2012 Q3 $)** | **Classification** | **Median ($)** | **CPI-adjusted median (2012 Q3 $)** |
| **Total personal income categories** | Nil or loss | 0 | 0 | Nil or loss | 0 | 0 | Nil income or loss | 0 | 0 | Loss | -4,285 | -6,289 |
| | $1 - $249 | 125 | 551 | $1,000 or less | 445 | 1,112 | $2,500 or less | 850 | 1,376 | Zero income | 0 | 0 |
| | $250 - $499 | 375 | 1,652 | $1,001 - $2,500 | 1,641 | 4,100 | $2,501 - $5,000 | 3,593 | 5,817 | $1 - $5,000 | 1,675 | 2,458 |
| | $500 - $999 | 750 | 3,306 | $2,501 - $5,000 | 3,626 | 9,059 | $5,001 - $7,500 | 6,208 | 10,051 | $5,001 - $10,000 | 8,559 | 12,562 |
| | $1,000 - $1,999 | 1,500 | 6,615 | $5,001 - $7,500 | 6,252 | 15,620 | $7,501 - $10,000 | 8,761 | 14,185 | $10,001 - $15,000 | 12,528 | 18,388 |
| | $2,000 - $3,499 | 2,750 | 12,128 | $7,501 - $10,000 | 8,650 | 21,611 | $10,001 - $15,000 | 11,571 | 18,734 | $15,001 - $20,000 | 17,281 | 25,364 |
| | $3,500 - $4,999 | 4,250 | 18,745 | $10,001 - $12,500 | 11,255 | 28,119 | $15,001 - $20,000 | 17,436 | 28,230 | $20,001 - $25,000 | 22,347 | 32,800 |
| | $5,000 - $6,499 | 5,750 | 25,362 | $12,501 - $15,000 | 13,811 | 34,505 | $20,001 - $25,000 | 22,420 | 36,299 | $25,001 - $30,000 | 27,370 | 40,172 |
| | $6,500 - $7,999 | 7,250 | 31,979 | $15,001 - $17,500 | 16,163 | 40,381 | $25,001 - $30,000 | 27,321 | 44,234 | $30,001 - $40,000 | 34,360 | 50,432 |
| | $8,000 - $9,999 | 9,000 | 39,698 | $17,501 - $20,000 | 18,626 | 46,534 | $30,001 - $40,000 | 34,545 | 55,930 | $40,001 - $50,000 | 43,934 | 64,484 |
| | $10,000 - $11,999 | 11,000 | 48,520 | $20,001 - $25,000 | 22,196 | 55,453 | $40,001 - $50,000 | 44,019 | 71,269 | $50,001 - $70,000 | 57,426 | 84,287 |
| | $12,000 - $13,999 | 13,000 | 57,343 | $25,001 - $30,000 | 26,888 | 67,176 | $50,001 - $70,000 | 56,467 | 91,423 | $70,001 - $100,000 | 81,524 | 119,657 |
| | $14,000 - $15,999 | 15,000 | 66,165 | $30,001 - $35,000 | 32,148 | 80,317 | $70,001 and over | 90,200 | 146,039 | $100,001 or more | 130,960 | 192,217 |
| | $16,000 - $17,999 | 17,000 | 74,987 | $35,001 - $40,000 | 36,914 | 92,224 | | | | | | |
| | $18,000 - $19,999 | 19,000 | 83,810 | $40,001 - $50,000 | 43,719 | 109,225 | | | | | | |
| | $20,000 - $22,499 | 21,250 | 93,735 | $50,001 and over | 59,918 | 149,696 | | | | | | |
| | $22,500 - $24,999 | 23,750 | 104,763 | | | | | | | | | |
| | $25,000 - $27,499 | 26,250 | 115,791 | | | | | | | | | |
| | $27,500 - $29,999 | 28,750 | 126,818 | | | | | | | | | |
| | $30,000 - $34,999 | 32,500 | 143,360 | | | | | | | | | |
| | $35,000 - $39,999 | 37,500 | 165,416 | | | | | | | | | |
| | $40,000 - $49,999 | 45,000 | 198,500 | | | | | | | | | |
| | $50,000 - $59,999 | 55,000 | 242,611 | | | | | | | | | |
| | $60,000 or over | 81,213[3] | 358,243 | | | | | | | | | |
| **Residual categories** | Not specified | | | Not specified | | | Not specified | | | Not specified | | |

| **Census:** | **2001 census** | | | **2006 census** | | | **2013 census** | | |
|---|---|---|---|---|---|---|---|---|---|
| **Source from which medians were estimated[a]:** | **Tax** | | | **Survey of Family, Income and Employment** | | | **Household Economic Survey** | | |
| | **Classification** | **Median ($)** | **CPI-adjusted median (2012 Q3 $)** | **Classification** | **Median ($)** | **CPI-adjusted median (2012 Q3 $)** | **Classification** | **Median ($)** | **CPI-adjusted median (2012 Q3 $)** |
| **Total personal income categories** | Loss | -2,440 | -3,321 | Loss | -6,998 | -8,426 | Loss | -20,716 | -20,716 |
| | Zero income | 0 | 0 | Zero income | 0 | 0 | Zero income | 0 | 0 |
| | $1 - $5,000 | 1,981 | 2,697 | $1 - $5,000 | 1,604 | 1,931 | $1 - $5,000 | 1,787 | 1,787 |
| | $5,001 - $10,000 | 8,067 | 10,981 | $5,001 - $10,000 | 7,840 | 9,440 | $5,001 - $10,000 | 7,717 | 7,717 |
| | $10,001 - $15,000 | 12,547 | 17,080 | $10,001 - $15,000 | 12,345 | 14,864 | $10,001 - $15,000 | 13,155 | 13,155 |
| | $15,001 - $20,000 | 17,236 | 23,463 | $15,001 - $20,000 | 17,084 | 20,570 | $15,001 - $20,000 | 17,301 | 17,301 |
| | $20,001 - $25,000 | 22,373 | 30,455 | $20,001 - $25,000 | 22,201 | 26,731 | $20,001 - $25,000 | 21,997 | 21,997 |
| | $25,001 - $30,000 | 27,524 | 37,467 | $25,001 - $30,000 | 27,203 | 32,754 | $25,001 - $30,000 | 27,443 | 27,443 |
| | $30,001 - $40,000 | 34,668 | 47,192 | $30,001 - $35,000 | 32,119 | 38,673 | $30,001 - $35,000 | 32,349 | 32,349 |
| | $40,001 - $50,000 | 44,164 | 60,119 | $35,001 - $40,000 | 37,131 | 44,708 | $35,001 - $40,000 | 37,539 | 37,539 |
| | $50,001 - $70,000 | 56,912 | 77,472 | $40,001 - $50,000 | 44,168 | 53,181 | $40,001 - $50,000 | 44,895 | 44,895 |
| | $70,001 - $100,000 | 80,531 | 109,624 | $50,001 - $70,000 | 57,514 | 69,250 | $50,001 - $60,000 | 54,429 | 54,429 |
| | $100,001 or more | 135,948 | 185,060 | $70,001 - $100,000 | 80,725 | 97,197 | $60,001 - $70,000 | 64,800 | 64,800 |
| | | | | $100,001 or more | 135,007 | 162,556 | $70,001 - $100,000 | 81,000 | 81,000 |
| | | | | | | | $100,001 - $150,000 | 120,000 | 120,000 |
| | | | | | | | $150,001 or more | 200,000 | 200,000 |
| **Residual categories** | Not stated | | | Not stated | | | Not stated | | |

**Notes**:

[1] These representative values (mid-points and medians) were supplied to the author by Dr Robert Didham from Stats NZ.

[2] Total personal income in 1981 was derived by adding together the responses to two separate census questions, one on gross income from all social security benefits, and one on gross income from all sources excluding social security benefits.

[3] This value of $81,213 for the highest income category in 1981 is taken from Fawcett, Atkinson, Herd and Blakely (2008, p. 138), who estimated it using Parker and Fenwick's (1983) Pareto curve-based method (see Fawcett, 2005, pp. 99-100).

**Appendix Table 2:** *Calculation of a proxy for lifetime average income by consecutive else non-consecutive multiyear average of income observations, else (for offspring only) a single year observation, all centred on sex-specific proxy ages*.

| | | Census year | | | | | | | Proxies for lifetime average income centred at age 35 ± 1 year (or 1.5 years for some age cohorts) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Year son born | | 1981 | 1986 | 1991 | 1996 | 2001 | 2006 | 2013 | Consecutive multiyear average (number of observations, median age) | Non-consecutive multiyear average (number of observations, median age) | Single year (number of observations, age observed) |
| 1979 | Son's age | 2 | 7 | 12 | 17 | 22 | 27 | 34 | | | 2013 (1 obs, 34.0 yrs) |
| 1978 | | 3 | 8 | 13 | 18 | 23 | 28 | 35 | | | 2013 (1 obs, 35.0 yrs) |
| 1977 | | 4 | 9 | 14 | 19 | 24 | 29 | 36 | | | 2013 (1 obs, 36.0 yrs) |
| 1976 | | 5 | 10 | 15 | 20 | 25 | 30 | 37 | 2006-2013 (2 obs, 33.5 yrs) | | |
| 1975 | | 6 | 11 | 16 | 21 | 26 | 31 | 38 | 2006-2013 (2 obs, 34.5 yrs) | | |
| 1974 | | 7 | 12 | 17 | 22 | 27 | 32 | 39 | 2006-2013 (2 obs, 35.5 yrs) | | |
| 1973 | | 8 | 13 | 18 | 23 | 28 | 33 | 40 | 2001-2013 (3 obs, 34.0 yrs) | 2001+2013 (2 obs, 34.0 yrs) | |
| 1972 | | 9 | 14 | 19 | 24 | 29 | 34 | 41 | 2001-2013 (3 obs, 35.0 yrs) | 2001+2013 (2 obs, 35.0 yrs) | 2006 (1 obs, 34.0 yrs) |
| 1971 | | 10 | 15 | 20 | 25 | 30 | 35 | 42 | 2001-2013 (3 obs, 36.0 yrs) | 2001+2013 (2 obs, 36.0 yrs) | 2006 (1 obs, 35.0 yrs) |
| 1970 | | 11 | 16 | 21 | 26 | 31 | 36 | 43 | 1996-2013 (4 obs, 34.5 yrs) | 1996+2013 (2 obs, 34.5 yrs) | 2006 (1 obs, 36.0 yrs) |
| 1969 | | 12 | 17 | 22 | 27 | 32 | 37 | 44 | 1996-2013 (4 obs, 35.5 yrs) or 2001-2006 (2 obs, 34.5 yrs) | 1996+2013 (2 obs, 35.5 yrs) | |
| 1968 | | 13 | 18 | 23 | 28 | 33 | 38 | 45 | 1991-2013 (5 obs, 34.0 yrs) or 2001-2006 (2 obs, 35.5 yrs) | 1991+2013 (2 obs, 34.0 yrs) | |
| 1967 | | 14 | 19 | 24 | 29 | 34 | 39 | 46 | 1991-2013 (5 obs, 35.0 yrs) or 1996-2006 (3 obs, 34.0 yrs) | 1991+2013 (2 obs, 35.0 yrs) or 1996+2006 (2 obs, 34.0 yrs) | 2001 (1 obs, 34.0 yrs) |

| | | Census year | | | | | | | Proxies for lifetime average income centred at age 40 ± 1 year (or 1.5 years for some age cohorts) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Year daughter born | | 1981 | 1986 | 1991 | 1996 | 2001 | 2006 | 2013 | Consecutive multiyear average (number of observations, median age) | Non-consecutive multiyear average (number of observations, median age) | Single year (number of observations, age observed) |
| 1974 | Daughter's age | 7 | 12 | 17 | 22 | 27 | 32 | 39 | | | 2013 (1 obs, 39.0 yrs) |
| 1973 | | 8 | 13 | 18 | 23 | 28 | 33 | 40 | | | 2013 (1 obs, 40.0 yrs) |
| 1972 | | 9 | 14 | 19 | 24 | 29 | 34 | 41 | | | 2013 (1 obs, 41.0 yrs) |
| 1971 | | 10 | 15 | 20 | 25 | 30 | 35 | 42 | 2006-2013 (2 obs, 38.5 yrs) | | |
| 1970 | | 11 | 16 | 21 | 26 | 31 | 36 | 43 | 2006-2013 (2 obs, 39.5 yrs) | | |
| 1969 | | 12 | 17 | 22 | 27 | 32 | 37 | 44 | 2006-2013 (2 obs, 40.5 yrs) | | |
| 1968 | | 13 | 18 | 23 | 28 | 33 | 38 | 45 | 2001-2013 (3 obs, 39.0 yrs) | 2001+2013 (2 obs, 39.0 yrs) | |
| 1967 | | 14 | 19 | 24 | 29 | 34 | 39 | 46 | 2001-2013 (3 obs, 40.0 yrs) | 2001+2013 (2 obs, 40.0 yrs) | 2006 (1 obs, 39.0 yrs) |

| | | Census year | | | | | | | Proxies for lifetime average income centred at age 35 ± 1 year | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year father born | | 1981 | 1986 | 1991 | 1996 | 2001 | 2006 | 2013 | Consecutive multiyear average (number of observations, median age) | Non-consecutive multiyear average (number of observations, median age) |
| 1964 | Father's age | 17 | 22 | 27 | 32 | 37 | 42 | 49 | 1986-2013 (6 obs, 35.5 yrs) or 1991-2006 (4 obs, 34.5 yrs) or 1996-2001 (2 obs, 34.5 yrs) | 1986+2013 (2 obs, 35.5 yrs) or 1991+2006 (2 obs, 34.5 yrs) |
| 1963 | | 18 | 23 | 28 | 33 | 38 | 43 | 50 | 1981-2013 (7 obs, 34.0 yrs) or 1991-2006 (4 obs, 35.5 yrs) or 1996-2001 (2 obs, 35.5 yrs) | 1981+2013 (2 obs, 34.0 yrs) or 1991+2006 (2 obs, 35.5 yrs) |
| 1962 | | 19 | 24 | 29 | 34 | 39 | 44 | 51 | 1981-2013 (7 obs, 35.0 yrs) or 1986-2006 (5 obs, 34.0 yrs) or 1991-2001 (3 obs, 34.0 yrs) | 1981+2013 (2 obs, 35.0 yrs) or 1986+2006 (2 obs, 34.0 yrs) or 1991+2001 (2 obs, 34.0 yrs) |
| 1961 | | 20 | 25 | 30 | 35 | 40 | 45 | 52 | 1981-2013 (7 obs, 36.0 yrs) or 1986-2006 (5 obs, 35.0 yrs) or 1991-2001 (3 obs, 35.0 yrs) | 1981+2013 (2 obs, 36.0 yrs) or 1986+2006 (2 obs, 35.0 yrs) or 1991+2001 (2 obs, 35.0 yrs) |
| 1960 | | 21 | 26 | 31 | 36 | 41 | 46 | 53 | 1986-2006 (5 obs, 36.0 yrs) or 1991-2001 (3 obs, 36.0 yrs) | 1986+2006 (2 obs, 36.0 yrs) or 1991+2001 (2 obs, 36.0 yrs) |
| 1959 | | 22 | 27 | 32 | 37 | 42 | 47 | 54 | 1981-2006 (6 obs, 34.5 yrs) or 1986-2001 (4 obs, 34.5 yrs) or 1991-1996 (2 obs, 34.5 yrs) | 1981+2006 (2 obs, 34.5 yrs) or 1986+2001 (2 obs, 34.5 yrs) |
| 1958 | | 23 | 28 | 33 | 38 | 43 | 48 | 55 | 1981-2006 (6 obs, 35.5 yrs) or 1986-2001 (4 obs, 35.5 yrs) or 1991-1996 (2 obs, 35.5 yrs) | 1981+2006 (2 obs, 35.5 yrs) or 1986+2001 (2 obs, 35.5 yrs) |
| 1957 | | 24 | 29 | 34 | 39 | 44 | 49 | 56 | 1981-2001 (5 obs, 34.0 yrs) or 1986-1996 (3 obs, 34.0 yrs) | 1981+2001 (2 obs, 34.0 yrs) or 1986+1996 (2 obs, 34.0 yrs) |
| 1956 | | 25 | 30 | 35 | 40 | 45 | 50 | 57 | 1981-2001 (5 obs, 35.0 yrs) or 1986-1996 (3 obs, 35.0 yrs) | 1981+2001 (2 obs, 35.0 yrs) or 1986+1996 (2 obs, 35.0 yrs) |
| 1955 | | 26 | 31 | 36 | 41 | 46 | 51 | 58 | 1981-2001 (5 obs, 36.0 yrs) or 1986-1996 (3 obs, 36.0 yrs) | 1981+2001 (2 obs, 36.0 yrs) or 1986+1996 (2 obs, 36.0 yrs) |
| 1954 | | 27 | 32 | 37 | 42 | 47 | 52 | 59 | 1981-1996 (4 obs, 34.5 yrs) or 1986-1991 (2 obs, 34.5 yrs) | 1981+1996 (2 obs, 34.5 yrs) |
| 1953 | | 28 | 33 | 38 | 43 | 48 | 53 | 60 | 1981-1996 (4 obs, 35.5 yrs) or 1986-1991 (2 obs, 35.5 yrs) | 1981+1996 (2 obs, 35.5 yrs) |
| 1952 | | 29 | 34 | 39 | 44 | 49 | 54 | 61 | 1981-1991 (3 obs, 34.0 yrs) | 1981+1991 (2 obs, 34.0 yrs) |
| 1951 | | 30 | 35 | 40 | 45 | 50 | 55 | 62 | 1981-1991 (3 obs, 35.0 yrs) | 1981+1991 (2 obs, 35.0 yrs) |
| 1950 | | 31 | 36 | 41 | 46 | 51 | 56 | 63 | 1981-1991 (3 obs, 36.0 yrs) | 1981+1991 (2 obs, 36.0 yrs) |
| 1949 | | 32 | 37 | 42 | 47 | 52 | 57 | 64 | 1981-1986 (2 obs, 34.5 yrs) | |
| 1948 | | 33 | 38 | 43 | 48 | 53 | 58 | 65 | 1981-1986 (2 obs, 35.5 yrs) | |

| | | Census year | | | | | | | Proxies for lifetime average income centred at age 40 ± 1 year | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year mother born | | 1981 | 1986 | 1991 | 1996 | 2001 | 2006 | 2013 | Consecutive multiyear average (number of observations, median age) | Non-consecutive multiyear average (number of observations, median age) |
| 1965 | Mother's age | 16 | 21 | 26 | 31 | 36 | 41 | 48 | 1996-2013 (4 obs, 39.5 yrs) | 1996+2013 (2 obs, 39.5 yrs) |
| 1964 | | 17 | 22 | 27 | 32 | 37 | 42 | 49 | 1996-2013 (4 obs, 40.5 yrs) or 2001-2006 (2 obs, 39.5 yrs) | 1996+2013 (2 obs, 40.5 yrs) |
| 1963 | | 18 | 23 | 28 | 33 | 38 | 43 | 50 | 1991-2013 (5 obs, 39.0 yrs) or 2001-2006 (2 obs, 40.5 yrs) | 1991+2013 (2 obs, 39.0 yrs) |
| 1962 | | 19 | 24 | 29 | 34 | 39 | 44 | 51 | 1991-2013 (5 obs, 40.0 yrs) or 1996-2006 (3 obs, 39.0 yrs) | 1991+2013 (2 obs, 40.0 yrs) or 1996+2006 (2 obs, 39.0 yrs) |
| 1961 | | 20 | 25 | 30 | 35 | 40 | 45 | 52 | 1991-2013 (5 obs, 41.0 yrs) or 1996-2006 (3 obs, 40.0 yrs) | 1991+2013 (2 obs, 41.0 yrs) or 1996+2006 (2 obs, 40.0 yrs) |
| 1960 | | 21 | 26 | 31 | 36 | 41 | 46 | 53 | 1986-2013 (6 obs, 39.5 yrs) or 1996-2006 (3 obs, 41.0 yrs) | 1986+2013 (2 obs, 39.5 yrs) or 1996+2006 (2 obs, 41.0 yrs) |
| 1959 | | 22 | 27 | 32 | 37 | 42 | 47 | 54 | 1986-2013 (6 obs, 40.5 yrs) or 1991-2006 (4 obs, 39.5 yrs) or 1996-2001 (2 obs, 39.5 yrs) | 1986+2013 (2 obs, 40.5 yrs) or 1991+2006 (2 obs, 39.5 yrs) |
| 1958 | | 23 | 28 | 33 | 38 | 43 | 48 | 55 | 1981-2013 (7 obs, 39.0 yrs) or 1991-2006 (4 obs, 40.5 yrs) or 1996-2001 (2 obs, 40.5 yrs) | 1981+2013 (2 obs, 39.0 yrs) or 1991+2006 (2 obs, 40.5 yrs) |
| 1957 | | 24 | 29 | 34 | 39 | 44 | 49 | 56 | 1981-2013 (7 obs, 40.0 yrs) or 1986-2006 (5 obs, 39.0 yrs) or 1991-2001 (3 obs, 39.0 yrs) | 1981+2013 (2 obs, 40.0 yrs) or 1986+2006 (2 obs, 39.0 yrs) or 1991+2001 (2 obs, 39.0 yrs) |
| 1956 | | 25 | 30 | 35 | 40 | 45 | 50 | 57 | 1981-2013 (7 obs, 41.0 yrs) or 1986-2006 (5 obs, 40.0 yrs) or 1991-2001 (3 obs, 40.0 yrs) | 1981+2013 (2 obs, 41.0 yrs) or 1986+2006 (2 obs, 40.0 yrs) or 1991+2001 (2 obs, 40.0 yrs) |
| 1955 | | 26 | 31 | 36 | 41 | 46 | 51 | 58 | 1986-2006 (5 obs, 41.0 yrs) or 1991-2001 (3 obs, 41.0 yrs) | 1986+2006 (2 obs, 41.0 yrs) or 1991+2001 (2 obs, 41.0 yrs) |
| 1954 | | 27 | 32 | 37 | 42 | 47 | 52 | 59 | 1981-2006 (6 obs, 39.5 yrs) or 1986-2001 (4 obs, 39.5 yrs) or 1991-1996 (2 obs, 39.5 yrs) | 1981+2006 (2 obs, 39.5 yrs) or 1986+2001 (2 obs, 39.5 yrs) |
| 1953 | | 28 | 33 | 38 | 43 | 48 | 53 | 60 | 1981-2006 (6 obs, 40.5 yrs) or 1986-2001 (4 obs, 40.5 yrs) or 1991-1996 (2 obs, 40.5 yrs) | 1981+2006 (2 obs, 40.5 yrs) or 1986+2001 (2 obs, 40.5 yrs) |
| 1952 | | 29 | 34 | 39 | 44 | 49 | 54 | 61 | 1981-2001 (5 obs, 39.0 yrs) or 1986-1996 (3 obs, 39.0 yrs) | 1981+2001 (2 obs, 39.0 yrs) or 1986+1996 (2 obs, 39.0 yrs) |
| 1951 | | 30 | 35 | 40 | 45 | 50 | 55 | 62 | 1981-2001 (5 obs, 40.0 yrs) or 1986-1996 (3 obs, 40.0 yrs) | 1981+2001 (2 obs, 40.0 yrs) or 1986+1996 (2 obs, 40.0 yrs) |
| 1950 | | 31 | 36 | 41 | 46 | 51 | 56 | 63 | 1981-2001 (5 obs, 41.0 yrs) or 1986-1996 (3 obs, 41.0 yrs) | 1981+2001 (2 obs, 41.0 yrs) or 1986+1996 (2 obs, 41.0 yrs) |
| 1949 | | 32 | 37 | 42 | 47 | 52 | 57 | 64 | 1981-1996 (4 obs, 39.5 yrs) or 1986-1991 (2 obs, 39.5 yrs) | 1981+1996 (2 obs, 39.5 yrs) |
| 1948 | | 33 | 38 | 43 | 48 | 53 | 58 | 65 | 1981-1996 (4 obs, 40.5 yrs) or 1986-1991 (2 obs, 40.5 yrs) | 1981+1996 (2 obs, 40.5 yrs) |
| 1947 | | 34 | 39 | 44 | 49 | 54 | 59 | 66 | 1981-1991 (3 obs, 39.0 yrs) | 1981+1991 (2 obs, 39.0 yrs) |
| 1946 | | 35 | 40 | 45 | 50 | 55 | 60 | 67 | 1981-1991 (3 obs, 40.0 yrs) | 1981+1991 (2 obs, 40.0 yrs) |
| 1945 | | 36 | 41 | 46 | 51 | 56 | 61 | 68 | 1981-1991 (3 obs, 41.0 yrs) | 1981+1991 (2 obs, 41.0 yrs) |
| 1944 | | 37 | 42 | 47 | 52 | 57 | 62 | 69 | 1981-1986 (2 obs, 39.5 yrs) | |
| 1943 | | 38 | 43 | 48 | 53 | 58 | 63 | 70 | 1981-1986 (2 obs, 40.5 yrs) | |

**Appendix Table 3:** *Descriptive statistics of proxy for lifetime average income for sons in son-father and son-mother pairs, in real dollars and natural logs*.

| | | Proxy for lifetime average income (in 2012 Q3 dollars) for sons in son-father pairs | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | Minimum | | P25 | | Median | | P75 | | Maximum | | Mean | | SD | | Mean no. observations used in proxy | Mean age (years) at which proxy is centred |
| | | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | | |
| *All sons in son-father pairs* | 4,617 | 1 | 0.00 | 39,904 | 10.59 | 54,429 | 10.90 | 75,125 | 11.23 | 200,000 | 12.21 | 60,779 | 10.80 | 34,634 | 1.00 | 1.8 | 34.7 |
| *By age in 1981* | | | | | | | | | | | | | | | | | |
| 2 | 720 | 1 | 0.00 | 37,539 | 10.53 | 54,429 | 10.90 | 81,000 | 11.30 | 200,000 | 12.21 | 60,062 | 10.71 | 37,426 | 1.26 | 1.0 | 34.0 |
| 3 | 627 | 1 | 0.00 | 37,539 | 10.53 | 54,429 | 10.90 | 81,000 | 11.30 | 200,000 | 12.21 | 62,529 | 10.73 | 38,398 | 1.39 | 1.0 | 35.0 |
| 4 | 561 | 1 | 0.00 | 37,539 | 10.53 | 54,429 | 10.90 | 81,000 | 11.30 | 200,000 | 12.21 | 63,445 | 10.79 | 37,984 | 1.16 | 1.0 | 36.0 |
| 5 | 537 | 1 | 0.00 | 38,824 | 10.57 | 53,805 | 10.89 | 75,125 | 11.23 | 181,278 | 12.11 | 59,706 | 10.78 | 34,040 | 1.03 | 2.0 | 33.5 |
| 6 | 471 | 966 | 6.87 | 41,123 | 10.62 | 53,805 | 10.89 | 75,125 | 11.23 | 181,278 | 12.11 | 59,443 | 10.84 | 31,175 | 0.62 | 2.0 | 34.5 |
| 7 | 390 | 966 | 6.87 | 41,784 | 10.64 | 54,754 | 10.91 | 75,125 | 11.23 | 181,278 | 12.11 | 64,020 | 10.91 | 36,091 | 0.62 | 2.0 | 35.5 |
| 8 | 288 | 4,870 | 8.49 | 41,550 | 10.63 | 55,057 | 10.92 | 70,123 | 11.16 | 182,539 | 12.11 | 58,802 | 10.87 | 28,239 | 0.51 | 3.0 | 34.0 |
| 9 | 396 | 1,931 | 7.57 | 40,651 | 10.61 | 53,180 | 10.88 | 69,250 | 11.15 | 182,539 | 12.11 | 60,194 | 10.87 | 31,030 | 0.56 | 2.3 | 34.6 |
| 10 | 285 | 1 | 0.00 | 41,181 | 10.63 | 53,180 | 10.88 | 70,123 | 11.16 | 182,539 | 12.11 | 60,868 | 10.83 | 33,510 | 0.87 | 2.3 | 35.6 |
| 11 | 162 | 1 | 0.00 | 42,923 | 10.67 | 53,180 | 10.88 | 69,250 | 11.15 | 162,556 | 12.00 | 57,036 | 10.80 | 26,572 | 0.95 | 2.8 | 35.1 |
| 12 | 84 | 8,540 | 9.05 | 35,110 | 10.47 | 48,905 | 10.80 | 68,713 | 11.14 | 173,808 | 12.07 | 56,778 | 10.79 | 33,131 | 0.57 | 3.3 | 35.1 |
| 13 | 51 | 10,210 | 9.23 | 42,873 | 10.67 | 49,219 | 10.80 | 61,007 | 11.02 | 136,090 | 11.82 | 52,902 | 10.80 | 20,054 | 0.41 | 3.9 | 34.5 |
| 14 | 42 | 17,079 | 9.75 | 42,953 | 10.67 | 57,023 | 10.95 | 65,157 | 11.08 | 177,174 | 12.08 | 60,336 | 10.90 | 30,550 | 0.46 | 3.7 | 34.6 |
| *By census linkage* | | | | | | | | | | | | | | | | | |
| 1981-2001 | 6 | 17,079 | 9.75 | 30,455 | 10.32 | 37,467 | 10.53 | 60,118 | 11.00 | 77,471 | 11.26 | 44,738 | 10.60 | 21,488 | 0.52 | 1.0 | 34.0 |
| 1981-2006 | 339 | 1 | 0.00 | 38,673 | 10.56 | 53,180 | 10.88 | 69,250 | 11.15 | 173,808 | 12.07 | 62,604 | 10.83 | 35,853 | 1.04 | 1.2 | 34.7 |
| 1981-2013 | 4,275 | 1 | 0.00 | 40,312 | 10.60 | 54,429 | 10.90 | 75,125 | 11.23 | 200,000 | 12.21 | 60,660 | 10.80 | 34,548 | 1.00 | 1.8 | 34.7 |
| *By number of income observations used in proxy* | | | | | | | | | | | | | | | | | |
| 1 | 2,223 | 1 | 0.00 | 37,539 | 10.53 | 54,429 | 10.90 | 81,000 | 11.30 | 200,000 | 12.21 | 61,988 | 10.75 | 37,685 | 1.25 | 1.0 | 34.9 |
| 2 | 1,473 | 1 | 0.00 | 40,312 | 10.60 | 53,805 | 10.89 | 75,125 | 11.23 | 181,278 | 12.11 | 60,367 | 10.83 | 33,700 | 0.79 | 2.0 | 34.4 |
| 3 | 726 | 3,147 | 8.05 | 41,712 | 10.64 | 54,391 | 10.90 | 70,123 | 11.16 | 182,539 | 12.11 | 59,170 | 10.87 | 28,548 | 0.50 | 3.0 | 34.8 |
| 4 | 147 | 11,534 | 9.35 | 41,341 | 10.63 | 53,014 | 10.88 | 65,200 | 11.09 | 157,976 | 11.97 | 55,004 | 10.83 | 22,951 | 0.44 | 4.0 | 34.9 |
| 5 | 54 | 15,143 | 9.63 | 45,282 | 10.72 | 52,837 | 10.87 | 67,220 | 11.12 | 177,174 | 12.08 | 59,626 | 10.91 | 26,956 | 0.41 | 5.0 | 34.4 |
| | | Proxy for lifetime average income (in 2012 Q3 dollars) for sons in son-mother pairs | | | | | | | | | | | | | | | |
| | n | Minimum | | P25 | | Median | | P75 | | Maximum | | Mean | | SD | | Mean no. observations used in proxy | Mean age (years) at which proxy is centred |
| | | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | | |
| *All sons in son-mother pairs* | 14,526 | 1 | 0.00 | 39,971 | 10.60 | 54,429 | 10.90 | 75,125 | 11.23 | 200,000 | 12.21 | 61,195 | 10.83 | 34,385 | 0.88 | 2.3 | 34.8 |
| *By age in 1981* | | | | | | | | | | | | | | | | | |
| 2 | 1,068 | 1 | 0.00 | 37,539 | 10.53 | 54,429 | 10.90 | 81,000 | 11.30 | 200,000 | 12.21 | 57,492 | 10.65 | 35,737 | 1.30 | 1.0 | 34.0 |
| 3 | 1,008 | 1 | 0.00 | 37,539 | 10.53 | 54,429 | 10.90 | 81,000 | 11.30 | 200,000 | 12.21 | 62,720 | 10.69 | 41,443 | 1.47 | 1.0 | 35.0 |
| 4 | 1,011 | 1 | 0.00 | 37,539 | 10.53 | 54,429 | 10.90 | 81,000 | 11.30 | 200,000 | 12.21 | 62,892 | 10.77 | 39,362 | 1.16 | 1.0 | 36.0 |
| 5 | 1,056 | 1 | 0.00 | 38,824 | 10.57 | 53,805 | 10.89 | 75,125 | 11.23 | 181,278 | 12.11 | 60,228 | 10.80 | 34,060 | 0.92 | 2.0 | 33.5 |
| 6 | 1,014 | 1 | 0.00 | 41,123 | 10.62 | 53,805 | 10.89 | 75,125 | 11.23 | 181,278 | 12.11 | 60,414 | 10.84 | 32,505 | 0.76 | 2.0 | 34.5 |
| 7 | 1,065 | 1 | 0.00 | 41,123 | 10.62 | 58,990 | 10.99 | 75,125 | 11.23 | 181,278 | 12.11 | 64,426 | 10.87 | 36,494 | 0.87 | 2.0 | 35.5 |
| 8 | 969 | 5,029 | 8.52 | 41,416 | 10.63 | 55,057 | 10.92 | 70,581 | 11.16 | 182,539 | 12.11 | 61,088 | 10.89 | 31,871 | 0.54 | 3.0 | 34.0 |
| 9 | 1,599 | 1,931 | 7.57 | 40,391 | 10.61 | 54,391 | 10.90 | 70,123 | 11.16 | 182,539 | 12.11 | 62,534 | 10.88 | 34,556 | 0.63 | 2.2 | 34.6 |
| 10 | 1,503 | 1 | 0.00 | 40,538 | 10.61 | 55,057 | 10.92 | 73,243 | 11.20 | 182,539 | 12.11 | 62,682 | 10.88 | 34,001 | 0.72 | 2.2 | 35.6 |
| 11 | 1,323 | 1 | 0.00 | 41,810 | 10.64 | 54,568 | 10.91 | 72,438 | 11.19 | 184,958 | 12.13 | 62,824 | 10.88 | 33,369 | 0.77 | 2.9 | 35.0 |
| 12 | 1,134 | 5,491 | 8.61 | 40,894 | 10.62 | 54,540 | 10.91 | 73,051 | 11.20 | 184,958 | 12.13 | 60,508 | 10.87 | 31,865 | 0.55 | 3.3 | 35.1 |
| 13 | 855 | 5,491 | 8.61 | 40,044 | 10.60 | 52,891 | 10.88 | 70,001 | 11.16 | 173,808 | 12.07 | 57,958 | 10.85 | 27,975 | 0.50 | 3.9 | 34.5 |
| 14 | 918 | 1 | 0.00 | 39,243 | 10.58 | 50,561 | 10.83 | 66,677 | 11.11 | 185,060 | 12.13 | 57,058 | 10.79 | 29,807 | 0.81 | 3.5 | 34.6 |
| *By census linkage* | | | | | | | | | | | | | | | | | |
| 1981-2001 | 207 | 1 | 0.00 | 30,455 | 10.32 | 47,192 | 10.76 | 60,118 | 11.00 | 185,060 | 12.13 | 57,036 | 10.60 | 40,144 | 1.46 | 1.0 | 34.0 |
| 1981-2006 | 2,406 | 1 | 0.00 | 38,673 | 10.56 | 53,180 | 10.88 | 69,250 | 11.15 | 173,808 | 12.07 | 63,877 | 10.87 | 36,712 | 0.82 | 1.4 | 34.8 |
| 1981-2013 | 11,916 | 1 | 0.00 | 40,312 | 10.60 | 54,429 | 10.90 | 75,125 | 11.23 | 200,000 | 12.21 | 60,725 | 10.82 | 33,760 | 0.88 | 2.5 | 34.8 |
| *By number of income observations used in proxy* | | | | | | | | | | | | | | | | | |
| 1 | 4,971 | 1 | 0.00 | 37,539 | 10.53 | 54,429 | 10.90 | 81,000 | 11.30 | 200,000 | 12.21 | 62,113 | 10.75 | 38,940 | 1.21 | 1.0 | 34.9 |
| 2 | 3,972 | 1 | 0.00 | 38,824 | 10.57 | 54,754 | 10.91 | 75,125 | 11.23 | 181,278 | 12.11 | 61,142 | 10.84 | 34,222 | 0.81 | 2.0 | 34.6 |
| 3 | 3,033 | 2,475 | 7.81 | 41,962 | 10.64 | 55,622 | 10.93 | 73,211 | 11.20 | 182,539 | 12.11 | 61,564 | 10.90 | 31,398 | 0.54 | 3.0 | 34.9 |
| 4 | 1,539 | 3,141 | 8.05 | 41,810 | 10.64 | 54,540 | 10.91 | 72,385 | 11.19 | 184,958 | 12.13 | 60,200 | 10.89 | 29,283 | 0.51 | 4.0 | 35.0 |
| 5 | 1,014 | 11,244 | 9.33 | 41,039 | 10.62 | 52,310 | 10.86 | 68,404 | 11.13 | 162,220 | 12.00 | 57,311 | 10.86 | 25,552 | 0.45 | 5.0 | 34.5 |

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

**Appendix Table 4:** *Descriptive statistics of proxy for lifetime average income for daughters in daughter-mother and daughter-father pairs, in real dollars and natural logs*.

| | | Proxy for lifetime average income (in 2012 Q3 dollars) for daughters in daughter-mother pairs[1] | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | Minimum | | P25 | | Median | | P75 | | Maximum | | Mean | | SD | | Mean no. observations used in proxy | Mean age (years) at which proxy is centred |
| | | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | NZD ($) | Log | | |
| *All daughters in daughter-mother pairs* | 9,312 | 1 | 0.00 | 17,301 | 9.76 | 32,349 | 10.38 | 51,600 | 10.85 | 200,000 | 12.21 | 37,690 | 9.90 | 30,787 | 2.05 | 1.7 | 39.7 |
| *By age in 1981* | | | | | | | | | | | | | | | | | |
| 7 | 1,500 | 1 | 0.00 | 13,155 | 9.48 | 32,349 | 10.38 | 54,429 | 10.90 | 200,000 | 12.21 | 37,961 | 9.62 | 32,798 | 2.63 | 1.0 | 39.0 |
| 8 | 1,353 | 1 | 0.00 | 17,301 | 9.76 | 32,349 | 10.38 | 54,429 | 10.90 | 200,000 | 12.21 | 38,848 | 9.60 | 33,643 | 2.73 | 1.0 | 40.0 |
| 9 | 1,407 | 1 | 0.00 | 17,301 | 9.76 | 32,349 | 10.38 | 54,429 | 10.90 | 200,000 | 12.21 | 39,216 | 9.74 | 34,748 | 2.46 | 1.0 | 41.0 |
| 10 | 1,329 | 1 | 0.00 | 17,140 | 9.75 | 31,004 | 10.34 | 51,737 | 10.85 | 181,278 | 12.11 | 38,146 | 10.08 | 30,350 | 1.54 | 2.0 | 38.5 |
| 11 | 1,224 | 1 | 0.00 | 17,224 | 9.75 | 32,135 | 10.38 | 49,568 | 10.81 | 181,278 | 12.11 | 36,237 | 10.04 | 26,484 | 1.60 | 2.0 | 39.5 |
| 12 | 1,038 | 1 | 0.00 | 18,431 | 9.82 | 32,135 | 10.38 | 49,568 | 10.81 | 181,278 | 12.11 | 38,617 | 10.18 | 30,416 | 1.24 | 2.0 | 40.5 |
| 13 | 783 | 1 | 0.00 | 18,130 | 9.81 | 29,124 | 10.28 | 45,598 | 10.73 | 182,539 | 12.11 | 35,265 | 10.15 | 26,920 | 1.09 | 3.0 | 39.0 |
| 14 | 678 | 1 | 0.00 | 18,317 | 9.82 | 32,082 | 10.38 | 45,598 | 10.73 | 151,437 | 11.93 | 34,728 | 10.13 | 23,023 | 1.20 | 3.0 | 40.0 |
| *By number of income observations used in proxy* | | | | | | | | | | | | | | | | | |
| 1 | 4,260 | 1 | 0.00 | 17,301 | 9.76 | 32,349 | 10.38 | 54,429 | 10.90 | 200,000 | 12.21 | 38,657 | 9.65 | 33,716 | 2.61 | 1.0 | 40.0 |
| 2 | 3,615 | 1 | 0.00 | 17,224 | 9.75 | 32,135 | 10.38 | 49,568 | 10.81 | 181,278 | 12.11 | 37,529 | 10.09 | 29,090 | 1.49 | 2.0 | 39.4 |
| 3 | 1,440 | 1 | 0.00 | 18,366 | 9.82 | 30,217 | 10.32 | 45,598 | 10.73 | 182,539 | 12.11 | 35,232 | 10.16 | 25,223 | 1.11 | 3.0 | 39.5 |
| | | **Proxy for lifetime average income (in 2012 Q3 dollars) for daughters in daughter-father pairs[1]** | | | | | | | | | | | | | | | |
| | **n** | **Minimum** | | **P25** | | **Median** | | **P75** | | **Maximum** | | **Mean** | | **SD** | | **Mean no. observations used in proxy** | **Mean age (years) at which proxy is centred** |
| | | **NZD ($)** | **Log** | **NZD ($)** | **Log** | **NZD ($)** | **Log** | **NZD ($)** | **Log** | **NZD ($)** | **Log** | **NZD ($)** | **Log** | **NZD ($)** | **Log** | | |
| *All daughters in daughter-father pairs* | 1,944 | 1 | 0.00 | 17,301 | 9.76 | 32,349 | 10.38 | 53,805 | 10.89 | 200,000 | 12.21 | 37,224 | 9.77 | 29,922 | 2.35 | 1.3 | 39.7 |
| *By age in 1981* | | | | | | | | | | | | | | | | | |
| 7 | 552 | 1 | 0.00 | 13,155 | 9.48 | 27,443 | 10.22 | 54,429 | 10.90 | 200,000 | 12.21 | 37,136 | 9.59 | 33,101 | 2.61 | 1.0 | 39.0 |
| 8 | 456 | 1 | 0.00 | 15,228 | 9.62 | 32,349 | 10.38 | 54,429 | 10.90 | 200,000 | 12.21 | 36,630 | 9.49 | 28,942 | 2.88 | 1.0 | 40.0 |
| 9 | 351 | 1 | 0.00 | 17,301 | 9.76 | 32,349 | 10.38 | 54,429 | 10.90 | 200,000 | 12.21 | 38,401 | 9.84 | 32,127 | 2.26 | 1.0 | 41.0 |
| 10 | 264 | 1 | 0.00 | 16,863 | 9.73 | 32,551 | 10.39 | 51,737 | 10.85 | 181,278 | 12.11 | 37,624 | 10.08 | 28,527 | 1.55 | 2.0 | 38.5 |
| 11 | 144 | 1 | 0.00 | 21,284 | 9.97 | 29,540 | 10.29 | 46,551 | 10.75 | 148,599 | 11.91 | 35,612 | 10.14 | 23,447 | 1.41 | 2.0 | 39.5 |
| 12 | 114 | 1 | 0.00 | 23,607 | 10.07 | 33,366 | 10.42 | 49,568 | 10.81 | 126,590 | 11.75 | 39,494 | 10.25 | 25,223 | 1.28 | 2.0 | 40.5 |
| 13 | 42 | 1 | 0.00 | 18,957 | 9.85 | 29,244 | 10.28 | 42,843 | 10.67 | 88,907 | 11.40 | 32,520 | 9.94 | 21,061 | 1.74 | 3.0 | 39.0 |
| 14 | 24 | 894 | 6.80 | 15,947 | 9.68 | 41,363 | 10.63 | 51,994 | 10.86 | 70,123 | 11.16 | 35,984 | 10.15 | 20,561 | 1.09 | 3.0 | 40.0 |
| *By number of income observations used in proxy* | | | | | | | | | | | | | | | | | |
| 1 | 1,359 | 1 | 0.00 | 13,155 | 9.48 | 32,349 | 10.38 | 54,429 | 10.90 | 200,000 | 12.21 | 37,295 | 9.62 | 31,491 | 2.62 | 1.0 | 39.9 |
| 2 | 522 | 1 | 0.00 | 19,943 | 9.90 | 32,343 | 10.38 | 49,568 | 10.81 | 181,278 | 12.11 | 37,284 | 10.12 | 26,534 | 1.47 | 2.0 | 39.2 |
| 3 | 63 | 1 | 0.00 | 20,115 | 9.91 | 34,108 | 10.44 | 47,147 | 10.76 | 88,907 | 11.40 | 35,159 | 10.13 | 20,148 | 1.47 | 3.0 | 39.3 |

**Source**: Author's calculations from New Zealand Longitudinal Census dataset.

**Appendix Table 5:** *Group-specific IGE estimates for ‘Any Māori’ versus ‘No Māori’ for each offspring-parent pair*.

| | Son-father pairs | | | Son-mother pairs | | | Daughter-mother pairs | | | Daughter-father pairs | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Son’s ethnicity[1] | | | Son’s ethnicity[1] | | | Daughter’s ethnicity[1] | | | Daughter’s ethnicity[1] | |
| | Pooled | Any Māori | No Māori | Pooled | Any Māori | No Māori | Pooled | Any Māori | No Māori | Pooled | Any Māori | No Māori |
| Sample size | 4,617 | 786 | 3,831 | 14,526 | 1,953 | 12,573 | 9,312 | 1,266 | 8,046 | 1,944 | 351 | 1,596 |
| Sample shares (%) | 100.0 | 17.0 | 83.0 | 100.0 | 13.4 | 86.6 | 100.0 | 13.6 | 86.4 | 100.0 | 18.1 | 82.1 |
| Mean log proxy income: offspring | 10.80 | 10.86 | 10.95 | 10.83 | 9.84 | 9.79 | 9.90 | 9.83 | 9.79 | 9.77 | 10.83 | 10.94 |
| Mean log proxy income: parents | 10.93 | 10.65 | 10.83 | 9.80 | 10.63 | 10.86 | 9.80 | 9.99 | 9.89 | 10.92 | 9.80 | 9.76 |
| Within-group component of pooled IGE | **0.239** | 0.245 | 0.224 | **0.054** | 0.041 | 0.056 | **0.145** | 0.125 | 0.146 | **0.135** | -0.064 | 0.211 |
| Between-group component of pooled IGE | | 0.077 | 0.003 | | -0.010 | 0.000 | | 0.004 | 0.000 | | -0.019 | -0.001 |
| **Group-specific IGE** | | **0.314** | **0.226** | | **0.023** | **0.058** | | **0.145** | **0.143** | | **-0.089** | **0.202** |
| Bootstrapped 95% confidence interval | | (0.065, 0.563) | (0.146, 0.307) | | (0.045, 0.071) | (-0.025, 0.070) | | (0.006, 0.284) | (0.083, 0.202) | | (-0.598, 0.419) | (-0.084, 0.489) |

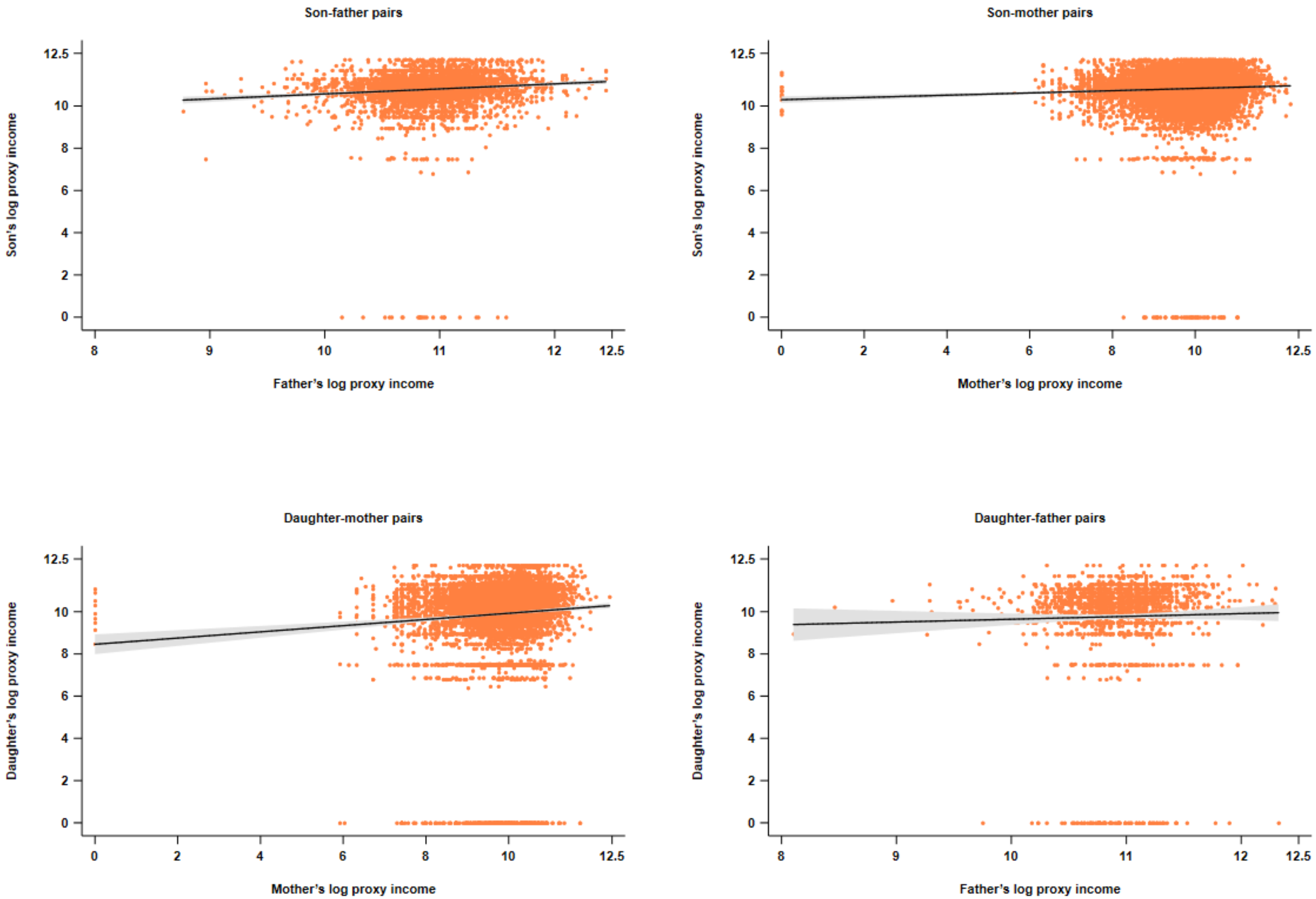
**Notes**:
[1] For this variable, responses to the ethnicity question from each offspring’s latest available census (where the response is non-missing) are dichotomised into ‘Any Māori’ ethnic identification (all offspring who mark the ‘Māori’ ethnic group box, regardless of whether they mark any other ethnic groups) versus ‘No Māori’ (all other offspring, which equates to offspring with no Māori ethnic identification).

**Source**: Author’s calculations from New Zealand Longitudinal Census dataset.

**Appendix Table 6**. *Average weekly gross market income bands used in CHDS parent interviews from child ages 1 to 14 years*.

| Code | 1-year (1978), 2-years (1979) | 3-years (1980) | 4-years (1981) | 5-years (1982), 6-years (1983), 7-years (1984) | 8-years (1985) | 9-years (1986), 10-years (1987) | 11-years (1988), 12-years (1989), 13-years (1990), 14-years (1991) |
|---|---|---|---|---|---|---|---|
| 00 | No income | No income | No income | No income | No income | No income | No income |
| 01 | $1 to $10 | $1 to $20 | $1 to $20 | $1 to $20 | $1 to $25 | $1 to $25 | $1 to $25 |
| 02 | $11 to $20 | $21 to $40 | $21 to $40 | $21 to $40 | $26 to $50 | $26 to $50 | $26 to $50 |
| 03 | $21 to $30 | $41 to $60 | $41 to $60 | $41 to $60 | $51 to $75 | $51 to $75 | $51 to $75 |
| 04 | $31 to $40 | $61 to $80 | $61 to $80 | $61 to $80 | $76 to $100 | $76 to $100 | $76 to $100 |
| 05 | $41 to $50 | $81 to $100 | $81 to $100 | $81 to $100 | $101 to $125 | $101 to $125 | $101 to $125 |
| 06 | $51 to $60 | $101 to $120 | $101 to $120 | $101 to $120 | $126 to $150 | $126 to $150 | $126 to $150 |
| 07 | $61 to $70 | $121 to $140 | $121 to $140 | $121 to $140 | $151 to $175 | $151 to $175 | $151 to $175 |
| 08 | $71 to $80 | $141 to $160 | $141 to $160 | $141 to $160 | $176 to $200 | $176 to $200 | $176 to $200 |
| 09 | $81 to $90 | $161 to $180 | $161 to $180 | $171 to $180 | $201 to $225 | $201 to $225 | $201 to $250 |
| 10 | $91 to $100 | $181 to $200 | $181 to $200 | $181 to $200 | $226 to $250 | $226 to $250 | $251 to $300 |
| 11 | $101 to $110 | $201 to $220 | $201 to $220 | $201 to $220 | $251 to $275 | $251 to $275 | $301 to $350 |
| 12 | $111 to $120 | $221 to $240 | $221 to $240 | $221 to $240 | $276 to $300 | $276 to $300 | $351 to $400 |
| 13 | $121 to $130 | $241 to $260 | $241 to $260 | $241 to $260 | $301 to $325 | $301 to $350 | $401 to $450 |
| 14 | $131 to $140 | $261 to $280 | $261 to $280 | $261 to $280 | $326 to $350 | $351 to $400 | $451 to $500 |
| 15 | $141 to $150 | $281 to $300 | $281 to $300 | $281 to $300 | $351 to $375 | $401 to $450 | $501 to $550 |
| 16 | $151 to $160 | $301 to $320 | $301 to $320 | $301 to $320 | $376 to $400 | $451 to $500 | $551 to $600 |
| 17 | $161 to $170 | $321 to $340 | $321 to $340 | $321 to $340 | $401 to $425 | $501 to $550 | $601 to $650 |
| 18 | $171 to $180 | $341 to $360 | $341 to $360 | $341 to $360 | $426 to $450 | $551 to $600 | $651 to $725 |
| 19 | $181 to $190 | $361 to $380 | $361 to $380 | $361 to $380 | $451 to $475 | $601 to $650 | $726 to $800 |
| 20 | $191 to $200 | $381 to $400 | $381 to $400 | $381 to $400 | $476 to $500 | $651 to $700 | $801 to $875 |
| 21 | $201 to $210 | $401 to $420 | $401 to $420 | $401 to $420 | $501 to $525 | $701 to $775 | $876 to $950 |
| 22 | $211 to $220 | $421 to $440 | $421 to $440 | $421 to $440 | $526 to $550 | $776 to $850 | $951 to $1025 |
| 23 | $221 to $230 | Over $440 | $441 to $460 | $441 to $460 | $551 to $575 | $851 to $925 | $1026 to $1100 |
| 24 | $231 to $240 | | $461 to $480 | $461 to $480 | $576 to $600 | $926 to $1000 | $1101 to $1200 |
| 25 | $241 to $250 | | $481 to $500 | $481 to $500 | $601 to $625 | Over $1000 | $1201 to $1300 |
| 26 | $251 to $260 | | $501 to $520 | $501 to $520 | $626 to $650 | | $1301 to $1400 |
| 27 | $261 to $270 | | Over $520 | $521 to $540 | $651 to $675 | | Over $1400 |
| 28 | $271 to $280 | | | $541 to $560 | $676 to $700 | | |
| 29 | $281 to $290 | | | $561 to $580 | $701 to $725 | | |
| 30 | $291 to $300 | | | $581 to $600 | $726 to $750 | | |
| 31 | Over $300 | | | Over $600 | Over $750 | | |

**Appendix Figure 1**. *Scatterplots of offspring's versus parent's log proxy incomes for each NZLC sample*.




**Source**: Author's calculations from New Zealand Longitudinal Census dataset.