The economic consequences of an exogenous health shock: an empirical study on the labour market outcomes after a road accident.

by

Sandra McCallum

A thesis submitted to Auckland University of Technology in partial fulfillment of the requirement for the degree of Master of Business, Economics

2024

School of Economics

Faculty of Business Economics and Law

Abstract

This thesis examines the impact of an exogenous health shock on labour market outcomes in New Zealand. Assessing the economic impact of a change in health status is challenging due to the endogeneity of health. I take advantage of data available in the Statistics New Zealand's Integrated Data Infrastructure (IDI). I use Accident Compensation Corporation (ACC) data to identify individuals travelling as passengers in a road accident. Next, I link my population of interest with tax data from Inland Revenue (IR) that contains detailed monthly information on employment, earnings and benefit recipiency. To quantify the impact of the health event on the labour market, I exploit the timing variation of the road accident. I apply a newly developed empirical methodology by Callaway and Sant'Anna (2021) in which individuals receive treatment at different times and where there may be heterogeneity in the treatment effects over time. The results show significant decreases in employment and earnings from wages and salary for up to three years after the accident. These effects are exacerbated for health shocks of greater severity and for individuals of non-European ethnicity. These findings indicate that a health shock can cause longer-term economic scars.

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 used artificial intelligence tools or generative artificial intelligence tools (unless it is clearly stated, and referenced, along with the purpose of use), 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

Students signature: Sandra McCallum

Date: 26/2/2024

Student ID: 0006411

Acknowledgements

I would like to thank my supervisors, Alexander Plum, Kabir Dasgupta and Gail Pacheco. Your knowledge and expertise has always inspired me and I feel very fortunate to have had your advice and support during my thesis study. I'd like to thank Kabir for his willingness to see a situation from different angles and suggest alternative ways to interpret each problem. Your broad experience and insights taught me a lot. Most importantly I'd like to thank my main supervisor Alex for his patience and support, particularly in the datalab. I learned so much and gained confidence as you were always happy to answer questions and assist. Your attention to detail and desire for me to reach a high standard of work, helped me produce a much better version of my thesis. And it was always good to have a laugh.

I acknowledge my lecturers from the Economics Department who provided a high standard of teaching and gave support and advice throughout my postgraduate studies. I also want to thank the research team at the New Zealand Work Research Institute who inspired me everyday with their insights, expertise and valuable work. Special thanks to Alice Theadom, Lisa Meehan and Gail Pacheco who employed me as a research assistant on the TBI project. This gave me valuable experience in the IDI and also means I am now a co-author.

I thank my friends who supported me and encouraged me to believe in myself and keep on going. They stood by patiently while I disappeared at times to work on various aspects of my master's degree. This proved to be longer and more difficult than I anticipated. But definitely worth it.

I would like to thank my parents for their love and support on this postgraduate journey. Particularly my mother who understood the importance of education, especially for women. I acknowledge her mother, my grandmother Aline White (nee Wells), who won a scholarship and graduated with a Master of Arts from the University of Canterbury and worked as a journalist. My father's mother Nora McCallum (nee Lawless) who was fiercely independent and a determined hard worker. My Aunty Sheila Downey (nee McCallum) who was an artist and home maker and always supportive of whatever I was doing. My Aunt Krystyna Downey (nee Kolodynska) who came to New Zealand as a refugee from Poland after WW2. She graduated with an Arts Degree from the University of Auckland and spoke 5 languages. Then went on to work as the New Zealand secretary for UNESCO, the first woman to do so. I acknowledge all the inspiring women I have met and who worked so hard so that I could have the choices I've had in my life today.

Finally, I thank Jillian Pennington who employed me as a Data Analyst at Tūtohi/Wild Bamboo part of Wise Group. Working for Wise Group, a not-for-profit organisation, who provide a range of services in the community, is certainly a dream job. And I get to work with a team of supportive individuals who believe in working to make the world a better place.

I dedicate this thesis to Celia and Stephen who died in a car accident at the tender age of 19 while we were studying at university together. They both contributed so much to so many even though they were only here for a short time.

Disclaimer

These results are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI) which is carefully managed by Statistics New Zealand. For more information about the IDI please visit the Statistics New Zealand website.

The opinions, findings, recommendations, and conclusions expressed in this thesis are those of the author, not Statistics NZ.

The results are based in part on tax data supplied by Inland Revenue to Statistics New Zealand under the Tax Administration Act 1994 for statistical purposes. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes and is not related to the data's ability to support Inland Revenue's core operational requirements.

All observation counts have been randomly rounded to base 3 in accordance with Statistics New Zealand's confidentiality rules.

Table of Contents

List of Figures

List of Tables

1. Introduction

Health affects every aspect of our lives, including our ability to work and earn a living. The economic consequences of an exogenous health shock are many. Past research shows that health shocks can reduce employment and incomes. These impacts can be short and long-term (Dano, 2005; Halla & Zweimüller, 2013; Parro & Pohl, 2021). Unemployment reduces the productive capacity of the economy. If unemployment persists, there can be impacts on extended families and children. Children of the long-term unemployed tend to do worse academically. There are links to crime and other social issues after extended periods of unemployment (Nichols, 2013). Then there are the costs to the government funded social security system. This includes medical care, rehabilitation and social security payments; such as the unemployment benefit.

Past literature shows that increased government spending on health leads to better health outcomes, increased labour productivity and higher incomes (Bloom et al., 2004; Hall et al., 2012; Raghupathi & Raghupathi, 2020). Other literature contests that government policy redistributing income via taxation is more effective at improving health and socio-economic status (Deaton, 2002). A key problem with determining effective government policies surrounding health and labour market outcomes is due to the endogeneity of health. The relationship between the two is well established (Smith, 1999). However, it is difficult to measure the impact of specifically health on labour market outcomes. This is due to reverse causality, whereby health status can impact labour market outcomes and also labour market outcomes can impact health. Also, there are a number of confounding factors surrounding health and labour market outcomes such as education, race and socio-economic status (Deaton, 2002).

Thus, to measure the impact of a change in health status on labour market outcomes, it is necessary first to define an exogenous health shock. Past research has used a range of events such as self-assessed health, disease onset and accidents as health shocks. These studies find mixed results about the impact of the health shock on labour market outcomes. These differences may be explained by variations in health shock definitions, the methods used, types of data or the country's social security systems (Dano, 2005; Fadlon & Nielsen, 2021; Halla & Zweimüller, 2013; Parro & Pohl, 2021; Wu, 2003).

I seek to contribute to the literature on the relationship between health and labour market outcomes by answering the following research question: *How does an exogenous health shock affect an individual's future labour market outcomes?*

This study utilises full population administrative data from the Statistics New Zealand Integrated Data Infrastructure (IDI). The population of interest comes from the Accident Compensation Corporation (ACC) data. New Zealand has a unique no-fault-claims system, which means all individuals, if approved, are eligible for medical assistance after an accident. As an exogenous health event, I identify individuals travelling as passengers in a road accident using the ACC data. The assumption is that passengers have no control over the accident and the accident or health shock is truly unexpected or exogenous. Thus, this study's first contribution to the literature is defining a more exogenous health shock.

As a randomised experiment is not possible when measuring health shocks, I exploit the variation in the timing of the road accidents. A newly developed, quasi-experimental method by Callaway and Sant'Anna (2021) is applied. This method allows for treatment effect heterogeneity and variation in treatment timing. A control group of not-yet-treated individuals is used. This means both the treatment and control groups have similar time-invariant unobservable characteristics of individuals undergoing treatment. The second contribution I make to the literature is, to my knowledge, I am the first to use this method to measure the impact of a health shock on labour market outcomes.

In line with previous studies measuring the impact of health shocks, I find that a health shock worsens labour market outcomes. My primary analysis includes employed and unemployed individuals. For the full population of interest, the probability of employment decreases, on average, by 4.9 percentage points and monthly income declines on average by 0.41 log units in the first 34 months after the road accident. The results show an immediate labour market impact, which persists and worsens over the longer term. Individuals who remain employed after the health shock, experience, on average, a 3.3 percent decline in their monthly earnings. These negative effects are greater for more severe health shocks, those more attached to the labour market and for non-Europeans.

The remainder of my thesis proceeds as follows: Section 2 discusses economic theory and then gives a review of the international literature regarding health and labour market outcomes; Section 3 provides the institutional background to the relevant aspects of the New Zealand social security system; Section 4 describes the data used in the study and defines the population of interest and estimation samples; Section 5 details the empirical identification strategy; Section 6 presents the results and robustness checks; and Section 7 concludes.

2. Literature review

This study's main purpose is to examine the impact of a health shock on labour market outcomes. This section discusses economic theory and then gives a review of the international literature regarding health and labour market outcomes.

2.1 Economic Theory

2.1.1 Health Production Model

Like education, health is a form of human capital. In academic literature this concept was theoretically introduced by Becker (1964) and was later formally conceptualised by Grossman (1972). According to Grossman's model, everyone is born with a given stock of health that produces an output of good health and healthy time. This health stock is endogenous and sick days are a source of dis-utility. Dis-utility is experienced when something is undesirable. When purchasing goods and services, demand theory assumes individuals seek to maximise utility or satisfaction rather than dis-utility or dissatisfaction.

The demand for good health is described as a commodity or choice rather than a market good. Individuals make choices about how to allocate their time and resources. This is influenced by the shadow price of health. This in turn is determined by the cost of health promoting inputs, individuals' preferences and the opportunity cost of their time. The demand for health promoting inputs such as health care, healthy eating and exercise is thus a derived demand.

An individual's utility function across time can be expressed as:

$$U = U(\phi_o H_o, \dots, \phi_n H_n, Z_o, \dots, Z_n), (1)$$

Where H_o is the inherited stock of health. This stock increases with investment and decreases (depreciates) with age. Moreover, H_i is the health stock at time i and $h_i (= \phi H_i)$ is the total health promoting inputs consumed in time period i. Z_i is the total consumption of other inputs at time i . N (death) can vary according to investment and depreciation. Death occurs when $H_i = H_{min}$.

Individuals choose whether to invest in exogenous health promoting inputs that increase H_i or other inputs Z_i in order to maximise their utility in equation 1.

$$H_{i+1} - H_i = I_i - \delta_i H_i \ (2)$$

Equation 2 measures an individual's net investment in their health from period H_i to period H_{i+1}. Net investment is gross investment, I_i minus depreciation in period *i*, $\delta_i H_i$. Depreciation, δ_i, increases with age (after some point in the life cycle). If net investment is positive then the health stock increases.

The supply of health stock is fixed at any given time and completely inelastic. Demand for good health is downward sloping due to the diminishing marginal productivity for health capital. The health stock depreciates over time and shifts upwards, reducing the amount of healthy time available. With net investment in health the supply curve shifts downward, increasing the amount of healthy time. The model assumes that more education increases the efficiency of individuals' ability to increase their output of healthy time. Increases in the wage rate and education can shift the demand curve upwards and also increase healthy time.

When healthy time increases, individuals can choose to use it for market or non-market activities, work or leisure. If they work more their incomes increase.

As individuals are assumed to be utility maximisers, they will allocate their resources and time to maximise equation 1. This determines how much they invest in increasing their "good health" or "healthy time".

2.1.2 Other Health Theory

Additionally, poor health is likely to reduce the ability to work effectively, which can result in decreased earnings and reduced labour supply, thereby increasing the opportunity cost of work over leisure. According to Chirikos (1993), the utility from not working may increase with poorer health as people prefer to take time out of the labour force to take care of their health. Individuals may also invest more in health promoting inputs such as medical care. This is expected to reduce labour supply and income.

Alternatively, poor health can cause an increase in the labour supply due to an income effect. Increased costs of healthcare mean people may work more to cover these costs. (Dwyer & Mitchell, 1999).

2.1.2.1 Reverse Causality

The relationship between income and health is one of reverse causality (Currie & Madrian, 1999). Poor health can reduce labour supply and decrease "healthy time" and the ability to work which in turn reduces labour supply and wages. Labour market activity can also worsen health by increasing stress if employment conditions change or incomes drop for example and therefore reducing healthy time.

2.1.2.2 The Health-Income Gradient

An increase in wages means an individual can purchase more health promoting inputs and increase their health stock. With better health individuals tend to be more productive and may choose to substitute work over leisure as their incomes increase. This relationship is often described as a "gradient", as incomes increase so does good health and vice versa. In the broader sense this gradient is also affected by wealth, education, occupation, race and geography or socio-economic status (Deaton, 2002).

Deaton (2002) discusses the gradient between health and income, and presents mechanisms via which government policy can impact health. These are investments in the health care system and promoting healthy lifestyles. The author concludes however that redistribution of income via taxation and benefits for lower income groups is a more effective policy than spending more on the healthcare system. Other researchers argue that health outcomes have greater influence on socio-economic status, making investment in health policies more effective (Smith, 1999). Several studies have shown that increased health spending leads to better health outcomes, increased labour productivity and higher incomes (Bloom et al., 2004; Hall et al., 2012; Raghupathi & Raghupathi, 2020)

2.1.3 Government Policy

Past literature shows the increase in disability, job insurance and pension programmes has had an impact on the labour supply. The main impact being a reduction in the labour supply of older men who may exit the labour market earlier due to generous pension programmes. This may also confound the true impact of a health shock (Chirikos, 1993).

Chirikos (1993) uses a simple model to describe the health-labour market relationship as follows:

$$W_i - \alpha - \beta_h H_i + \Sigma_j \beta_j Y_{ij} + \Sigma_k \beta_k X_{jk} + \varepsilon_i \ (3)$$

Where W is the work effort of individual i measured in say hours or weeks, H is a measure of impaired health; Y_j represents factors that affect work such as wage, pension conditions and health insurance coverage; X_k are exogenous variables such as age and gender and ε_I the normally distributed error term. The coefficient of interest is the β_h. This measures the impact of poor health on work effort and is expected to be negative. In this model the size of β_h may be affected by other factors such as the wage rate and social security programmes. Social security programmes such as disability allowance or pension programmes may provide economic incentives to delay returning to work or to stay out of the labour market altogether. Therefore, it's important to consider the institutional arrangements of a country when determining the impact of a health shock on labour market outcomes.

Thus, Government policy has important implications for health and the relationship between health and labour market outcomes. Government policy can directly and indirectly influence health. First of all, governments can invest in health promoting inputs. This can include a robust health care system; building hospitals, raising incomes for medical staff and supporting rehabilitation programmes. They could also invest in education about healthy living. Secondly, policies that indirectly impact health such as improving families' socio-economic status. This could mean investing in education, addressing income equality via a progressive tax system and improving racial equality. Finally, government programmes such as accident, job insurance and pension schemes which can provide a safety net for those who experience health shocks.

Thus, it is easy to see that the number of different confounding factors surrounding health and income can make it difficult to target effective policies and to measure their impact.

In the next sections I discuss key international literature regarding health shocks and labour market outcomes.

2.2 International Literature to Date:

The past research regarding the impact of health shocks on labour market outcomes has focused on the following areas.

2.2.1 Aging Population

A large body of research to date focuses on the impact of health shocks on the elderly. Most developed countries have an aging population with declining trends of labour force participation rates. Therefore, there is increasing concern about the impact of health shocks on older workers. They often lead to early labour market exits, raising social security costs and reducing labour supply. (Au et al., 2005; Bound et al., 1999; Disney et al., 2006). This could increase the economic burden on countries and also motivates the need for government interventions to support the aging population's health and wellbeing.

2.2.2 Heterogeneity

A second area of interest in past research is heterogeneity in health shocks such as:

- whether different income groups, genders or ethnicity are impacted differently.
- if labour market attachment or social security programmes affect the impact of the health shock.
- whether more severe health shocks lead to worse labour market outcomes (Dano, 2005; Halla & Zweimüller, 2013; Parro & Pohl, 2021)

This past research has documented a variety of findings which are discussed in more detail in the following sections. These disparities may arise from variations in health shock definitions, data sources, methodologies, and institutional contexts (Parro & Pohl, 2021).

2.3 Defining an Exogenous Health Shock

According to Grossman (1972) health is endogenous. The endogeneity of health means there are a number of unobserved factors that may confound the relationship between health and income as discussed above. These include reverse causality and can cause issues when estimating causal inference with regards to the impact of health on income. Therefore, in order to measure the impact of health on labour market outcomes it is necessary to first define an exogenous health shock (Deaton, 2002). This health shock should ideally be uncorrelated with an individual's unobserved characteristics.

Past literature has defined health shocks in a variety of ways. These are discussed in the following sections. I consider first international examples and then local literature from Australasia.

2.3.1 Self-assessed Health

In a study by Lenhart (2019), the author uses panel data and self-assessed health and finds negative impacts on labour market outcomes in the United Kingdom after a health shock are strongest for victims of severe health shocks, males, and those with higher education and in managerial jobs. A propensity score matching difference-in-difference model was applied, using individuals who were actively employed at the time of the reported change in health status. Garcia Gomez and Lopez Nicolas (2006) also use panel data and self-assessed health, and propensity score matching and difference-in-differences methodology. The authors found that a health shock increased the probability of unemployment by 5 percent and labour inactivity by 3.5 percent in Spain. Unemployment includes those individuals who are actively seeking work. Labour inactivity is people who are not actively seeking work such as a student or home maker. In this case they aren't classified as unemployed.

However, using self-assessed health introduces endogeneity concerns due to measurement error. The key concern being rationalisation which can occur when people justify their labour supply position with their health response, potentially biasing results. (Dwyer and Michell, 1999, Bound, 1999, Kreider, 1999). For example, if a person is working less, they may report worse health.

2.3.2 Disease Onset

Another approach has been to use onset of diseases and health disorders such as cancer, diabetes or heart disease diagnosis as the health shock, (Fadlon & Nielsen, 2021; Heinesen & Kolodziejczyk, 2013; Jeon, 2017; Wu, 2003). Wu (2003) used exact coarsened matching and regression methods and found that individuals diagnosed with cancer experienced a 5-percentage point lower employment rate and 10 percentage point lower income after three years than those without the diagnosis in the United States. In a Danish study, Fadlon and Nielsen (2021) use fatal and non-fatal health shocks. The authors use the Death Registry to identify fatal health shocks and the National Patient Registry to identify heart attacks and strokes. Individuals who experience the same event a few years in the future are used as the control group to establish counterfactuals. They find that fatal events increase spousal labour supply although no effect from non-fatal events. These effects are driven by whether or not the person affected by the shock was the main income earner or not. The main earner still tends to be male.

However, these events may not be truly exogenous as symptoms may precede the diagnosis and individuals may have already adjusted their labour market behaviour. Charlies (2003) investigates the effects of a disability on income and finds that incomes have dropped a year before the disability diagnosis.

2.3.3 Accidents

There are also several studies that rely on accidents as exogenous health shocks. A key empirical advantage is that the accidents, unlike other health events, are more likely to be unanticipated and exogenously determined. Dano (2005) compares across individuals who met with a road accident in Denmark in a regression framework and finds significant negative impacts on incomes for older injured persons and those in lower income groups. Moreover, the research identified adverse effects on both short-term and long-term employment for injured men however no effect on employment for women. Importantly, the author did not differentiate between passengers and drivers or consider factors such as whether individuals were convicted of a crime due to the accident, which could also influence future labour market outcomes.

Halla and Zweimüller (2013) use commuting accidents in Austria – accidents to and from work– and find negative impacts on employment and income for injured workers when compared to non-injured workers. The estimated effects seem to be more severe for those who are typically less attached to the labour market (i.e females). The study considered a range of accidents, including tripping or falling as well as being hit by an object or a road accident. The study acknowledges the possibility of selection bias into different types of job due to preferences of living closer to work or using different forms of transport for getting to work. They use a traditional control group of non-injured individuals as does Dano (2005). It's important to note that this control group does not fully account for the unobservable characteristics that may be typical of people who meet with accidents. As such, this could affect the empirical analysis due to selection biases.

Both Dano (2005) and Halla and Zweimüller (2013) utilised propensity score matching and difference-in-differences techniques in their analyses.

Lindeboom et al. (2016) conducted a study using British survey data and an event history model. They found that health shocks, in the form of accidents, significantly increased the onset of a disability, which adversely affected employment. They did not identify a direct impact from the health shock itself on employment status. Effects were stronger for men and lower educated workers.

In a more recent Chilean study Parro and Pohl (2021) use hospital discharge records to identify accidents using the ICD10 codes, including "Injury", "Poisoning", and other consequences of external causes and external causes of morbidity but excluding "self-harm" or "complications from medical and surgical care". This is to ensure they include only individuals experiencing an exogenous health event. These health events occurred in a variety of circumstances and may or may not have been work-related. As all individuals in the study are eventually treated this effectively controls for unobservable characteristics related to accident-prone individuals. To ensure a strong attachment to the labour market, only men who are employed at least 18 months during the period before their accident are included in Parro and Pohl's (2021) analysis. The authors find employment and earnings both decline in the three years after the accident and that economic effects are greater than just healthcare costs.

Parro and Pohl (2021) use event study methods with time and individual fixed effects. The authors incorporate insights from Borusyak and Jaravel (2019) and set two of the pre-periods to zero in order to avoid an under-identification problem. This can occur due to negative weighting when treatment effects are heterogenous.

Finally, I take a look at literature from Australia and New Zealand.

2.4 Australasia

2.4.1 New Zealand

2.4.1.1 ACC Spells

Crichton et al. (2011) use Statistics New Zealand's Linked Employer Employee Database (LEED) from 1999 to 2004 to estimate the impact of injuries, work related and non-work related, on labour market outcomes. This includes wages and salary, accident compensation and benefit receipts of individuals aged 15 to 69. Injured workers are matched to non-injured workers first on observable individual characteristics and in a second specification with individuals from within the same firm. A person is classified as "injured" when they receive an ACC payment. This becomes their injury month and is when they begin an injury or ACC "spell". The authors use this amount of time or "spell" when the individual receives ACC earnings as a proxy for severity of their injuries. Crichton et al. (2011) match individuals on observable characteristics, age, gender, location, number of months employed in the seven months prior to injury, number of employees and industry type. The second specification matches individuals in the same firm, firm location and with similar incomes. The authors control for age and gender but not ethnicity and match on industry but not occupation. The data for ethnicity and occupation was not available. The data starts from 18 months before the injury month and impacts are estimated on outcomes 6, 12 and 18 months after the end of the ACC spell. Using first an event study and then a difference in differences matching approach the authors match non-injured workers to injured workers. The panel nature of the dataset controls for time-invariant unobserved individual heterogeneity.

Crichton et al. (2011) find differences between injured and non-injured workers after shorter ACC spells. However, these may be due to differences already occurring in the labour market before the injury. The authors find clear differences between injured and non-injured workers after longer spells. For those experiencing an ACC spell of 7 to 24 months they have, on average, 9 to 10 percent lower employment rates, 3 to 5 percent lower earnings, and $320 - $370 lower monthly incomes, 6 to 18 months after compensation ends compared to the non-injured group. Crichton et al. (2011) conclude that the NZ institutional framework fails to prevent these negative impacts on future labour market outcomes for injured individuals.

2.4.1.2 Cancer Registration

Another New Zealand-based study authored by Carter et al. (2013) looks into cancer registration or hospitalisation (at least one night's stay) as a health shock. The authors use 7 waves of the longitudinal survey of Family, Income and Employment from 2002 to 2009 (Carter et al., 2013). Participants of the survey, who gave permission, were linked to the cancer registration and public hospital data (N=6,780). A conditional logistic fixed effect regression model is used. The sample is stratified by gender and age group (25-39 years, 40-54 years) as well as gender by age group. The study found that those who experienced the health shock had a significant increased risk in dropping out of the labour force in the future (odds ratio 1.54, 95 percent CI 1.3 – 1.82).

2.4.2 Australia

2.4.2.1 Self-assessed Health

Changes in health may not cause a person to completely exit the labour market but to reduce their labour supply. Cai and Kalb (2006) use six waves of the Household, Income and Labour Dynamics in Australia (HILDA) Survey to estimate the impact of health status and health shocks on hours worked. These are measured separately in the survey. Health shocks in Cai and Kalb's (2006) analysis are measured as injuries or illnesses that occurred in the last 12 months. Health status is a general assessment of a person's health, for example excellent or very good. The authors' population of interest includes men aged 25 to 64 and women aged 25 to 60 years, excluding full-time students (N=22,698). The outcome variable hours worked are total weekly hours from all jobs including paid and unpaid overtime and work outside of their workplace. Each specification includes covariates for age, marital status and education. Using a dynamic random effects tobit model the authors estimate results for health status and health shocks separately and jointly. The authors also stratify each specification by gender.

Cai and Kalb (2006) found a post-health shock reduction in total working hours by an estimated 1 to 2 hours for those who indicated the health shock made their health "worse" and by 7 to 9 hours for "much worse". This reduction in labour supply after a health shock is consistent with other literature (Fadlon & Nielsen, 2021; Garcia Gomez & Lopez Nicolas, 2006). The authors also found that males reduced their hours worked by more than females but that females were more likely to exit the labour market completely.

This study makes two contributions to the literature. Firstly, I restrict my population of interest to individuals travelling as a passenger in a road accident. The assumption is that passengers have no control over the accident. Therefore, the accident (health shock) is unexpected and a more exogenous health shock than previous studies. Finally, to my knowledge I am the first to use the *CS-DID* method to analyse the impact of an exogenous health shock on labor market outcomes.

3. Institutional Background, the New Zealand Social Security System

Deaton (2002) states that to measure health's impact on income, it's necessary to consider the amount of support from the healthcare system and the availability of income support. Institutional arrangements can confound the effect of the causal impact of a health shock. (Chirikos, 1993; Mommaerts et al., 2020).

According to Garcia-Gomez (2011), differences in social security systems can help explain different results regarding recovery from health shocks from different countries. A key factor can be how the social support system provides income support or unemployment benefits and promotes reintegration of victims of negative health events into the work force. European countries such as France and Italy had no significant results in the authors' study. This may be explained by legislation that requires businesses to hire quotas of disabled workers. Countries like Denmark and Greece allow early retirement due to extenuating circumstances. This may encourage people to take early retirement and leave the labour force.

Mommaerts et al. (2020) estimate the impact of hospitalisations on labour market outcomes in the US, China and 13 European countries. They find the economic consequences are far less in European countries who have relatively better institutional support.

Fiebig et al. (2021) investigate whether socio-economic status impacts the use of primary care, specialist care and emergency department (ED) care in Australia. They link survey data to administrative data and find that low- and high-income individuals use different pathways for their care after being diagnosed with heart disease or diabetes. The Australian government fund a range of medical services. Australia's Medicare system is similar to New Zealand's (NZ) where practitioners can charge a co-payment. However, 86 percent of GP care is provided free of cost, or equal to the Medicare rebate, meaning lower income groups rely more on primary care. Specialist visits incur greater co-payments costs and thus higher income groups rely more on this level of healthcare. ED visits are free, and fully funded by the government. The authors found that it would be better for government policy to focus on the full pathway of treatment for individuals. This could ensure more equitable health outcomes.

In the following sections I outline the relevant government support available to individuals who experience health shocks in New Zealand.

3.1 Institutional Background: New Zealand Health System

New Zealand's health system is rooted in Te Tiriti o Waitangi (the Treaty of Waitangi), which specifies the founding principles for our society. These include equitable health outcomes for all groups in society in particular Māori. However, today Māori have worse health outcomes. (Ministry of Health, n.d.).

In 1938, the Social Security Act established New Zealand as the world's first welfare state, extending government assistance. This included a superannuation scheme, support for unemployment, widows, orphans, veterans and disabled individuals, as well as a health care system offering free hospital visits, medicine, a maternity benefit and subsidized doctors' visits (Ministry of Health, n.d)

In the following two sections I discuss the accident and employment insurance systems as these may provide support to individuals in my population of interest.

3.2 Institutional Background: The New Zealand Accident Insurance System: Accident Compensation Corporation (ACC)

In 1900 the Workers Compensation for Accidents Act was passed meaning that workers could receive some compensation in the case where no one was deemed responsible for their accident or injury. In 1972 the Accident Compensation Act was passed extending this coverage to all accidents and the Accident Compensation Commission was established on April 1st 1974 extending this coverage to all people injured in accidents in New Zealand. This unique no fault claims system means anyone seeking medical treatment for injuries is covered by the ACC, including overseas visitors. Therefore, individuals cannot sue for damages other than in exceptional cases. The latest legislation is contained in the Accident Compensation Act 2001 ("Accident Compensation Act," 2001; Accident Compensation Corporation, 2023).[1] Individuals injured in an accident visit their General Practitioner (GP) or other health provider to seek treatment. The provider fills in a claim form and sends it to the ACC. If the claim is accepted, ACC covers part of the payment for the treatment and individuals pay the rest.[2] Parliament approves the health services available for treatment and individuals can also apply to use Rongoa Māori health services (Accident Compensation Corporation, n.d.-c). The ACC can also provide help with a range of other support, transportation to and from work, home help, child care, equipment and a range of other help (Accident Compensation Corporation, n.d.-a) [3]

Those eligible can receive up to 80 percent of their usual income as weekly income compensation from their second week off work.[4] In order to receive income compensation, a person needs a Medical Certificate from their doctor outlining what work they are fit to do if any.[5] If a person is able to do some of their job-related tasks, their employer pays them for those tasks while the ACC tops up the rest. This may be added up to 100 percent of their usual salary. If a person is approved for income compensation from ACC then they are assigned a recovery team who assess their needs and determine what rehabilitation assistance and care they need as well as when they are ready to return to work (Accident Compensation Corporation, n.d.-b).

[1] Anyone who isn't entitled to income compensation and who cannot work may be entitled to an unemployment or sickness benefit which come under the jobseeker benefit.
[2] Individuals with a community services card for those on a lower incomes receive greater part payment by the government.
[3] ACC is funded via five different accounts: The Earners Account collects levies from individuals PAYE and self-employed individuals make direct payments; The non-earners account is financed directly from general taxation; The Work account comes from levies on employers and the self-employed. Petrol levies and license fees fund the motor-vehicle account and the treatment injury account is funded by the earners and non-earners account depending on whether the claimant is working.

[4] If the accident occurred at work, compensation for the first week off work is paid by the employer (note here with more details, need doctors' certificate etc) otherwise via sick leave or annual leave.
[5] Even if a person wasn't working when they had their accident they may be entitled under certain circumstances, for example if they were due to start working within 3 months of the accident or had been working within 28 days before the injury.

3.4 Institutional Background: Jobseeker Support NZ Unemployment Insurance

The Social Security Act 2018 sections 20 to 25 outlines current requirements, as stipulated by the New Zealand government, to qualify for the unemployment insurance or job seeker benefit. These include being at least 18 years old and having experienced a job loss due to redundancy, termination, or voluntary resignation, subject to specific conditions.[6] Furthermore, individuals who become ill or disabled may also be eligible for the job seeker benefit, subject to a medical certificate. The job seeker benefit provides a weekly allowance until the recipient secures alternative employment.[7] The administration of the benefit system falls under the responsibility of Work and Income, an entity operating under the Ministry of Social Development (MSD).[8]

If a person has a permanent or severe injury or disability or are caring for someone with a disability, they are likely to be eligible for the supported living payment, under part 2 sections 34 – 42 of the Social Security Act 2018.[9] There are also other aids available such as an accommodation supplement, rent arrears grant, an emergency grant for food, car repairs and a range of other costs if people meet the criteria.

McAllister et al. (2013) found that those injured and covered by the ACC were more likely to return to work and suffered lower financial losses than those unable to return to work due to illness and eligible for a means tested benefit. The benefit is approximately 50 percent of the

[6] You need to be a New Zealand citizen, permanent resident or have a residence class visa under the immigration Act 2009 and normally resident in NZ.
[7] The minimum weekly amount is $250.63 for a 18 to 19 year old living at home, they can earn up to $519 gross per week before this is reduced to zero support.
[8] If you are still receiving the payment after 12 months you need to reapply.
[9] Supported living payment base amount is $322.61 per week for 16- to 17-year-olds.

minimum wage (Ministry of Social Development, 2010). Even so, Crichton et al. (2011) conclude that the ACC system doesn't prevent employment and income losses.

To investigate if there may be confounding effects due to our social security system, I include benefit recipiency as an outcome variable. Individuals who weren't entitled to the ACC income compensation may use the benefit system if they are experiencing short- or longer-term health problems[10]. Those receiving income compensation that return to work but experience ongoing health problems may also end up unemployed and on a benefit. Due to the availability of income or sickness support, there may be an incentive for individuals who receive the benefit not to return to work. In some cases, if there was no benefit available, some individuals may have returned to work.

3.5 Institutional Background: The NZ Superannuation System

From the age of 65 years, individuals may be eligible for the New Zealand Superannuation payment.[11] This study is restricted to people who have an accident aged 55 or less. It analyses data up to 3 years after the accident therefore it's not possible for individuals to exit the labour market and claim the state pension.

This thesis adds to the international literature by analysing the impact of an exogenous health shock on labour market outcomes in a New Zealand context.

4.Data

4.1 Data Source

---

[10] For those on a sickness benefit they need ongoing medical checks and a certificate from their doctor to stay on the benefit.

[11] If you are a New Zealand citizen, permanent resident, or hold a residence class visa.

This research uses full population administrative data from the Integrated Data Infrastructure (IDI) which is managed by Statistics New Zealand (Stats NZ). The IDI contains deidentified microdata about households and individuals sourced from government agencies, such as the Ministry of Justice and the Inland Revenue Department, and non-government organisations, such as the Auckland City Mission, as well as survey data. A unique confidential number is used to identify individuals and enables linking between different datasets.

In this section I first describe the datasets I use in my analysis, secondly, I explain how I create my population of interest and then I give details about my estimation samples and health shock definitions. Finally, I present descriptive statistics as a precursor to the in-depth analysis using my identification strategy.

4.2 Key Datasets

4.2.1 Accident Compensation Corporation (ACC)

The spine/key dataset used for my research is the Accident Compensation Corporation (ACC) claims dataset which details claims made since 1974 and payments made since 2000 (Statistics New Zealand, n.d.a). I use the ACC records for individuals involved in a road accident between 2005 and 2015.

The data is collated by ACC and is from providers whom clients visit for medical treatment or from direct observations by ACC. New Zealand has a unique no-fault claims system whereby any individual who has an accident and seeks medical treatment is entitled to compensation, if the claim is approved by the ACC, and it is recorded in the ACC claims data. ACC covers all accidents for working, unemployed and retired people, students, visitors to NZ, children and adults (ACC New Zealand, n.d.-a). This unique system enables access to a broader population of interest than previous studies which use hospital data, survey data or social accident insurance covering only employees (Dano, 2005; Garcia Gomez & Lopez Nicolas, 2006; Halla & Zweimüller, 2013; Parro & Pohl, 2021) .

The ACC data contains information about the date of the accident and where it occurred, the type of accident i.e road accident, what the person was doing before the accident happened and whether or not it is work related. I also collate a person's ACC history as the dataset indicates if a person had a serious past injury or road accident. Other key variables in this rich dataset are details about the costs of each person's injury.

4.2.2 Personal Details

The unique confidential identifier allows me to link the spine data from ACC claims to other datasets. First, I link to the personal details data which provides demographic information about gender, age and ethnicity. There are five classifications for ethnicity: Māori, Pasifika, Asian, Middle East and Latin America (Mela), other and European. They are prioritised in this order which means if people classify themselves with more than one ethnicity such as Māori and European then Māori is prioritised. Due to the small sample size, I combine Asian, Mela and other into one category. Gender allows for two classifications, male or female. The age variables are for the month and year of birth which allow me to calculate the age of a person (rounded down to the year) at the time of the accident (Statistics New Zealand, n.d.b).

4.2.3 Overseas Spells

In order to determine any causal impact of the health shock I need to ensure individuals were present in New Zealand during the study period. The overseas spells dataset specifies the date a person left and entered the country whereby I can determine the amount of time they were in New Zealand during the period of interest. I link these details to the ACC spine.

4.2.4 Inland Revenue, Employee Monthly Schedule

The other main dataset is the Inland Revenue Employee Monthly Schedule (IR EMS). This contains all monthly wages and salary information about individuals and businesses engaged in economic activity (excluding self-employed) starting from April 1999 (Statistics New Zealand, 2018a). I use the IR EMS to identify the labour market outcome variables These are explained in detail in section 4.4.1.

4.2.5 Public Hospital Data

Finally, I link the Public Hospital discharges dataset to my ACC spine. I use details in this dataset to provide a robustness check for my main health shock definition. This details the date a person was admitted to hospital (stays for more than 3 hours), the length of stay and whether or not it was due to an accident. A zero-length of stay means a person didn't stay overnight; one indicates an overnight stay in hospital etc.(Statistics New Zealand, 2018b). I match the date the person was admitted to hospital with the date of the road accident from the ACC data to determine if a person was treated in hospital due to their accident.

I am also interested in a person's past hospital stays and create a variable for those who had a past stay of one night or more due to an accident.

4.3 Population of Interest

To evaluate the causal effect of an exogenous health shock on labour market outcomes I use the ACC records for individuals involved in a road accident between 2005 and 2015 (observations: 303,780). I then create my population of interest by imposing the following restrictions.

Firstly, I restrict the sample to those aged 25 to 55 years old at the date of the accident. If individuals have more than one accident during the study period, I use their first accident. My identification strategy uses data from 36 months before and 36 months after the road accident month. This means individuals are aged 22 up to 58 years during the study. I choose this age range as by 22 years old individuals have finished their compulsory schooling and, in many cases, completed their higher education and entered the workforce. Retirement age is 65 years old in New Zealand thus at 58 my population of interest are not eligible for government superannuation. This reduces the likelihood of bias from older participants taking early retirement due to their accident (Chirikos, 1993) (observations: 143,277) .

The ACC data identifies what people were doing before the road accident. One classification is travelling as a passenger (Statistics New Zealand, n.d.a). My second restriction is to select those individuals who were travelling as passengers in a road accident (observations: 10,296). The assumption is that passengers are not responsible for the accident and therefore the accident is unexpected. I can thus define a more exogenous health shock than prior studies that use disease onset or accidents where participants may have caused the accident (Dano, 2005; Halla & Zweimüller, 2013; Jeon, 2017; Parro & Pohl, 2021).

In order to properly assess their labour market activity, a person needs to be present in New Zealand for the majority of the study period. If they are in New Zealand for at least 75 percent of the 73 months I classify them as resident during the study period. Otherwise, they are classified as non-resident and excluded from the study (observations excluded: 1,026).

Next, I exclude those who had a night or more in hospital due to an accident or who were involved in a road accident in the past 3 years (observations excluded: 447). I want to ensure that my population of interest are not accident prone or regularly putting themselves in to risky situations. These are endogenous characteristics that can impact the exogeneity of the health shock. I also exclude any road accidents that occurred while a person is working, for example if they had a car accident during work hours. This avoids selection bias due to occupation type (observations excluded: 222).

Finally, anyone who has a previous serious injury or is deceased during the study is excluded.

This leaves me with a population of interest of 8,292 observations. My population of interest includes employed and unemployed individuals, male and female.

4.4 Data Structure

In this section I describe how I create the full panel of data ready for analysis.

In order to estimate the impact of the health shock on individuals' labour market outcomes I start with my ACC spine dataset of 8,292 observations. I set up a panel with 36 months before the health shock month and 36 months afterwards, a total of 73 months. This is then filled with IR EMS monthly data which includes gross wages and salaries and benefit recipiency[12]. This creates a balanced panel of IR data. Next, I create the outcome variables.

4.4.1 Outcome Variables

The two key outcome variables created are for monthly employment and earnings. First, a binary indicator for employment which is equal to one if they earn positive wages and salary (W & S) in a month. If an individual earns less than $50 in a month this is changed to missing. All missing values are changed to zero as per Parro and Pohl (2021). The second key outcome variable is the total W & S earned in a month. This is adjusted for inflation and converted to the log of total W & S. Finally, a third key outcome variable measures the change in monthly inflation adjusted W & S for those who remain in employment.

In addition to the main outcome variables, I create a binary indicator for whether or not an individual earned positive income from a benefit. There are two classifications for benefit. I am interested in changes in the unemployment, DPB, sickness, invalid, widow and emergency benefit classification. This includes the unemployment and sickness benefit which are the two most likely to change if the health shock has a longer-term impact on the population of interest.

[12] Gross wages and salaries are measured in $NZ and adjusted for inflation.

The other type of benefit indicates income from paid parental leave (PPL), student benefit (STU), pension (PEN) or withholding payment (WHP).

The literature states that the social security system of a country can have a confounding effect on the impact of a health shock. Thus, I include this final outcome variable, benefit recipiency, as a general indicator of whether the New Zealand system may be part of the reason for any changes in employment.

4.5 Estimation Samples

4.5.1 Baseline Sample

I define the health shock as the physical, emotional and psychological impact of the road accident. Anyone who seeks medical attention after the road accident is therefore included in the population of interest. Past studies using accidents have used length of stay in hospital or sick leave spells to determine seriousness (Crichton et al., 2011; Dano, 2005; Halla & Zweimüller, 2013; Parro & Pohl, 2021). Thus, the baseline sample includes N = 8,292 individuals who sought medical attention as a result of a road accident.

4.5.2 Health Shock Severity

Another key aspect of this study is heterogeneity with regards to the seriousness of the health shock. How the seriousness of the health shock impacts labour market outcomes.
It is difficult to compare one person's health shock and recovery to another person's and thus I use a range of seriousness: mild, medium and severe, which allows me to determine whether a more serious health shock, on average, has a greater effect on labour market outcomes. To provide a more robust analysis I use two different definitions of health shock and seriousness.

In the next section I explain the two health shock definitions used in this study.

4.5.2.1 Health Shock Definitions: Total Cost Definition

My main health shock definition uses the total cost variable. The assumption is that a road accident or injury that incurs a greater total cost is more serious as they use more medical and/or rehabilitation services. Each health provider a person visits after their accident fills in a claim form that is then forwarded to the ACC and the provider is reimbursed for the cost of the visit. In some cases, costs are not fully funded by the ACC and individuals pay a co-payment to health providers, therefore there can be an incentive not to seek follow-up treatment. Lower income individuals who are eligible for the community support services card are covered for a greater amount by the ACC (ACC New Zealand, n.d.-b). These costs are contained within two main variables. First, medical costs paid to practitioners, for example General Practitioners (GPs), Physiotherapists and other specialists. Second, the cost of entitlements which includes weekly income compensation, social and vocational rehabilitation. If a person is entitled to income compensation the employer pays the first week of compensation[13]. After the first week income compensation is paid by the ACC.

The total cost of each individual's claim is equal to the medical plus entitlements cost. In the case where a person has been to more than one provider there may be more than one claim number. To account for this, I check for duplicates for each claim number and accident date. I then sum the total medical fees and total entitlements for all unique claim numbers for the same accident date. This gives me the total ACC cost, now referred to as the total cost (Statistics New Zealand, n.d.a).

One caveat with the total cost variable is that emergency services are bulk funded therefore if a person is admitted to hospital after their road accident and has a stay of 3 days this cost won't be attributed to their claim (ACC New Zealand, n.d.-b). However, any follow up treatment in hospital, with medical practitioners and specialists or rehabilitation will be included. Thus, I assume that more serious injuries (health shocks) require follow up care.

[13] A medical certificate from a doctor is required to get income compensation and other assistance.

My identification strategy first estimates the impact of the health shock on the full population, the baseline sample, and then on three different levels of health shock. To define each level of health shock I look at the distribution of the total cost variable. The first 75 percent of the population of interest with a total cost of less than $700 are classified as a mild shock N=6,042. From 75 percent to 95 percent, with a total cost between $700 and $9000, are classified as medium, N=1,827. Those above 95 percent as severe, N=423.

In this study the mild health shock group is significantly larger as it contains people who not only attended hospital but sought medical treatment from other health providers such as their General Practitioner (GP). See Table 1. As a robustness check, in some parts of the analysis I also combine the medium and severe health shock samples, N = 2,250. This increases the sample size and therefore the power of the estimation. The key purpose, in this instance, is to determine the impact of more severe health shocks.

Table 1. Population of interest and estimation samples using the total cost definition

Sample: Health Shock	Total observation counts
Mild	6,042
Medium	1,827
Severe	423
Medium + Severe	2250
Total Population of Interest	8,292

Notes: A mild health shock includes individuals with a total cost of less than $700. A medium health shock those who have a total cost of >= $700 or less than $9000. A severe health shock >= $9000. Due to random rounding not, all totals add up.
Source: Own calculations using Statistics NZ's IDI

4.5.2.2 Health Shock Definition: Length-of-stay in Hospital

As a robustness check I use the public hospital discharges data to create a second health shock definition. This also creates a more direct comparison to previous literature that uses hospital data and which finds longer hospital stays lead to worse labour market outcomes (Dano, 2005; Halla & Zweimüller, 2013; Parro & Pohl, 2021).

I define a mild health shock as those who are not admitted to hospital or have an overnight stay in hospital. This includes all those from the ACC data who are not in the hospital data and all those in the hospital data who have a 0 or 1 length-of-stay. The more severe health shock is defined as a stay of more than 1 night in hospital and most severe more than 5 night's stay in hospital. It's important to note that the sample size gets smaller as the health shock is more severe. However, this is inevitable as using ACC data means a larger population of less severe injuries are included in my study. See table 2.

Table 2. Population of interest and estimation samples using hospital definition

Sample: Health Shock	Total observation counts
Mild	7,800
Medium	498
Severe	234
Total Population of Interest	8,292

Note: A mild health shock includes individuals who were not admitted to hospital or had length-of-stay of 0 or 1 night in hospital. A medium health shock are those individuals who had more than 1 night length-of-stay and severe more than 5 nights length-of-stay. Due to random rounding not all numbers add up.
Source: Own calculations using Statistics NZ's IDI

4.6 Descriptive Statistics

Table 3 below gives basic characteristics of the population of interest and each health shock description. The average age of the population of interest is 39 years old. The majority are female (69 percent), 46 percent are non-European and 54 percent are European. The full New Zealand population is 70.2 percent European according to the 2018 census (Statistics New Zealand, 2020) and 50.4% female (World Bank, 2022). The population of interest in this study therefore has a greater proportion of females and non-European individuals. The large proportion of females indicates some selection bias in to being a passenger. However, I am willing to accept this as the exogeneity of the health shock is the key determining factor in my analysis. Another point of note is that 78 percent of individuals who experienced the most severe health shock were employed in the 12 months before the accident compared to the mean of 69.5 percent. This larger proportion may be partly due to those who receive income compensation having a higher total cost variable. It's not possible to deduct income compensation from the total cost variable. Without the income compensation payments some individuals may move to the previous health shock group level. Therefore, as a robustness check, in some parts of the analysis, I combine the medium and severe cost group in to one. This also creates a larger sample size, N = 2,250 at the more severe health shock level. A further check on the outcomes for those who experience a more severe health shock. See appendix A, table 16 for the descriptive statistics for the hospital definition.

Table 3. Descriptive Statistics

	General Population N = 4,699,795	Total N=8292	Mild N=6042	Medium N=1827	Severe N=423
Age, yrs	38	39	38	39	40
Female %	54.4	69	70	68	59
Maori %	16.5	20	20	21	27
Pasifika %	8.1	7	7	6	6
Other %	16.6	19	19	21	13
European %	70.2	54	54	52	54
Employed %	69.5	65	63	63	78

Note: The general population figures in column 1 are from the 2018 census. Total in column 2 refers to the total population of interest. A mild health shock includes individuals with a total cost of less than $700. A medium health shock those who have a total cost of >= $700 or less than $9000. A severe health shock >= $9000. Due to random rounding not, all totals add up. Age is the mean age in years. All other variables are a percentage. Employed is the percentage of individuals who had positive income in the 12 months before the health shock for the full population figures individuals may classify themselves as more than one ethnicity therefore the total population ethnicity numbers don't add up to 100. The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018).
Source: Own calculations using Statistics NZ's IDI.

For the next part of my descriptive analysis, I plot the monthly employment and average log income for my population of interest across time stretching over 36 months prior to the accident and 36 months post-accident. This allows me to visually inspect any trends before and after the health shock occurs. As treatment occurs at different calendar month times, I normalise the month each person has their accident to month zero. Figure 1 shows the monthly employment pattern for my population of interest in event time. The 36 months before the accident are represented by negative values running from -36 to -1 and the months after the accident are denoted by positive values 1 to 36. The red vertical line indicates when the health shock (accident) or the treatment occurs, at event time zero. The y axis shows the percentage of the population of interest employed in a given month and the x axis the event time. Figure 1 clearly shows monthly employment drops after the health shock occurs at event time zero.

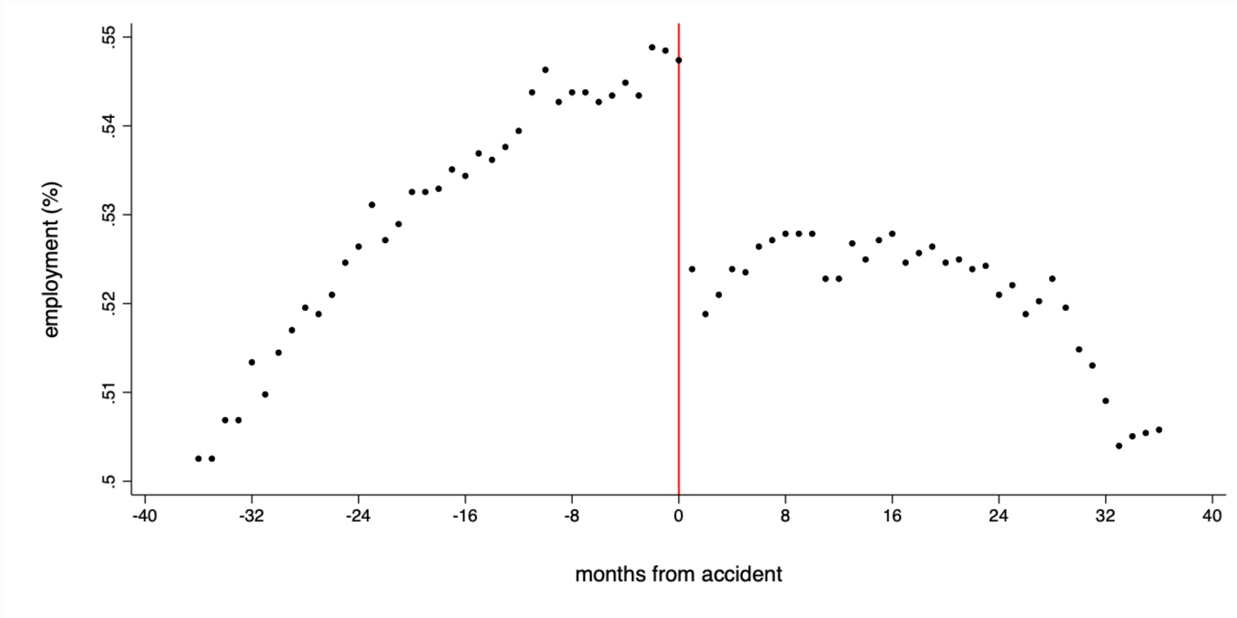
During the first 12 months from month -36 to month -24 monthly employment increases gradually from 50.25 to 52.64 then further to 53.94 percent in month -12. In the 12 months before the accident employment is relatively steady increasing slightly to 54.85 percent of the population of interest employed in the month before the health shock. After the accident occurs employment drops sharply to 52.39 percent in the month after the accident, then further in the second month to 51.88 percent of the population of interest.

In addition to the full population in Figure 1, I plot the monthly employment for the three different health shocks, mild, medium and severe, using the total cost definition (see section 4.5.2.1 for further explanation). The monthly employment pattern clearly shows as the health shock worsens the drop in employment after event time zero is larger (see appendix A figure 5, Table 17,18,19 and 20).

The severe shock sample has the highest level of pre-accident monthly employment at 68 percent employed in the month before the health shock (see appendix A, table 20).

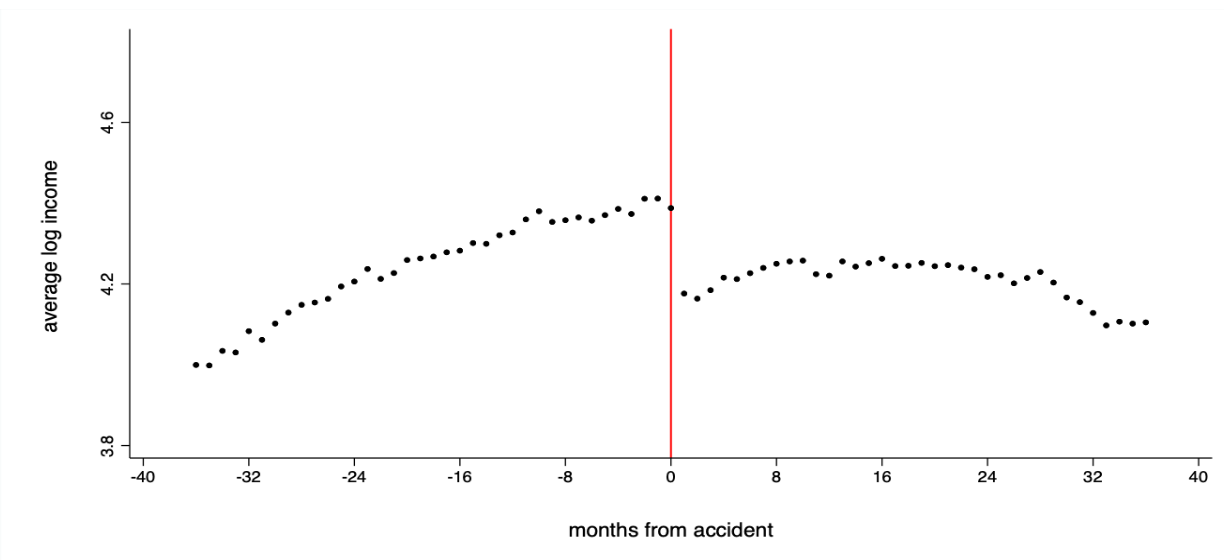
In a similar fashion, as for monthly employment I plot the monthly average income for the population of interest. Figure 2 shows the average monthly log income in the 36 months before the health shock and the 36 months after the health shock. Figure 2 clearly shows that the average monthly income drops sharply the month after the health shock occurs at event time zero. The month before the health shock average log income is 4.41 and the month after the health shock it drops to 4.18. This decline in monthly income after the accident increases as the health shock worsens.

Figure 1: Monthly employment in event time (%)



Notes: Figure 1 depicts the percentage of the population of interest who were employed each month. The monthly employment for the population of interest from 36 months before the health shock to 36 months after the health shock. The thirty-six months before the accident are represented by negative values running from -36 to -1 and the months after the accident are denoted by positive values 1 to 36. The red vertical line indicates when the health shock or the treatment occurs, at event time zero. The y axis shows the percentage of the population of interest employed in a given month and the x axis the event time. The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018).
Source: Own calculations using Statistics NZ's IDI

Figure 2: Monthly log income in event time



Notes: Figure 2 depicts the average monthly log income for the population of interest from 36 months before the health shock to 36 months after the health shock. The thirty-six months before the accident are represented by negative values running from -36 to -1 and the months after the accident are denoted by positive values 1 to 36. The red vertical line indicates when the health shock or the treatment occurs, at event time zero. The y axis shows the average monthly log income of the population of interest in a given month and the x axis the event time. The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018).
Source: Own calculations using Statistics NZ's IDI

In the next section I explain my empirical methodology.

5. Empirical Methodology

The purpose of this research is to identify the causal impact of an exogenous health shock on future labour market outcomes which include employment, wages and salary (W&S) and benefit recipiency.

In order to determine the causal impact of the health shock or treatment we need to compare the labour market outcomes of the same individual with and without the health shock as defined by Rubin (1974).

$\Delta Y_i = Y_i\ (D_i\ = 1) - \ Y_i\ (D_i = 0)$ (4)

where ΔY_i is the causal effect of the treatment on the outcome Y_i for individual *i*. Measured by the difference in the outcome if *i* received the treatment, ($D_i\ = 1$) and if the same individual *i* doesn't receive the treatment, ($D_i\ = 0$).

For the population of interest this can be described as the average treatment effect on the treated or *ATT*:

$ATT = E[Y_{1i}|D_i = 1] - E[\mathrm{Y_{0i}}|\mathrm{D_i} = 1]$ (5)

where Y_{1i} measures the potential average outcomes for individuals after a health shock; and $\mathrm{Y_{0i}}$ measures the potential average outcomes for the same group of individuals without the health shock. The dummy $D_i\ = 1$ indicates all individuals are members of the treatment group.

As we can't observe the counterfactual, the recognised gold standard for determining causal impact is a randomised experiment (Angrist, 2005). It's not possible to randomly assign health shocks to people. As such, we need to find a suitable treatment and comparison group that can allow us to determine the *ATT*. In this study, I use observational data and a quasi-experimental approach.

5.1 Basic Model: Difference-in-Differences Design

A common method to determine causal impact is the canonical difference-in-differences design (*DD*). This method typically involves two groups, a treatment and a control, and two time periods denoted by period 1 or pre-treatment period and period 2 or post-treatment period. In the pre-treatment period neither group is treated and in the second period only one group, the treatment group is treated.

$$ATT = E\left[Y_{i,2} - Y_{i,1} \,\middle|\, D_i = 1\right] - E\left[Y_{i,2} - Y_{i,1} \,\middle|\, D_i = 0\right] \text{ (6)}$$

where $E\left[Y_{i,2} - Y_{i,1} \,\middle|\, D_i = 1\right]$, measures the change between period 1 and 2 for individual *i* in the treatment group, D_i =1, and $E\left[Y_{i,2} - Y_{i,1} \,\middle|\, D_i = 0\right]$ for the control group, $D_i = 0$. The main assumption of the *DD* model is that without treatment the differences between treatment and control groups would remain the same. Thus, subtracting the second term from the first term, we derive the average causal impact of the treatment, or the *ATT* (Roth et al., 2023).

This method works well for a situation with two groups and two time periods and when all individuals receive treatment at the same time. However, in reality there are a range of situations such as more than two time periods and variation in treatment timing. In my study, my population of interest experience their treatment at different times between 2005 to 2015 and I observe their outcomes over a period of 73 months.

5.2 Two-Way Fixed Effects (*TWFE*)

In the case when there are more than two time periods and heterogenous treatment, one of the most common *DD* models is a linear regression with two-way fixed effects (*TWFE*).

This method estimates causal inference by comparing the outcomes of two groups, treatment and control, before and after a specific treatment that happens at event time *e* = 0 to the treatment group. Units can be treated at different times but the treatment time is normalised to time zero. In my study this is the month of the health shock. Therefore, all individuals experience the health shock at event time zero. Event time = -1 is the month before the event and event time +1 is the month after the event. The population of interest pre and post periods are the same in event or relative time but different in real or calendar time. The control group can contain never treated or not-yet-treated individuals. *TWFE* includes time and unit fixed effects that may impact an individual's outcomes and thus potentially bias the estimator. Time fixed effects control for time trends or seasonality and unit fixed effects control for time-constant characteristics.

The basic static specification for TWFE is as follows:

$$Y_{i,t} = \alpha_t + \alpha_i + \beta D_{i,t} + \varepsilon_{i,t} \quad (7)$$

where α_t is a time fixed effect and α_i is a unit fixed effect. The error term, $\varepsilon_{i,t}$ contains factors that are unobservable and vary over time with mean = 0. D_i is a dummy variable where D_i = 1 if treatment has occurred at time *t* and D_i = 0 otherwise. $Y_{i,t}$ is the outcome for individual *i* at time *t*. Once treatment has occurred a unit remains treated, this is called staggered adoption. β is the coefficient of interest which is the average causal impact over all periods for all units once the event has occurred.

$$Y_{i,t} = \alpha_t + \alpha_i + \sum_{e=-K}^{-2} \delta_e D_{i,t}^e + \sum_{e=0}^{l} \beta_e D_{i,t}^e + v_{i,t} \quad (8)$$

is the dynamic specification where the term $\sum_{e=-K}^{-2} \delta_e D_{i,t}^e$ is the pre-trends indicator that measures average outcomes, $Y_{i,t}$, from *-K* to *-2* before the event occurs at time *e*. Period *-1* is used as the baseline period. If the parallel trends assumption is met then this term will be *0*.

$\sum_{e=0}^{l} \beta e D_{i,t}^{e}$, this term measures any causal impact after the event occurs, for each β_e $e \geq 0$ up to period *l* and identifies short- and longer-term effects. The error term is $v_{i,t}$ with mean = *0*.

Recent literature has identified a number of issues with the *TWFE* model. In particular if there is treatment effect heterogeneity. This is where treatment effects may vary for each calendar time period and evolve differently over time. In this case, then effects from one period may contaminate another (Borusyak & Jaravel, 2019; Callaway & Sant'Anna, 2021; Goodman-Bacon, 2021; Sun & Abraham, 2021). The *TWFE* model generates average outcomes for all comparisons. These comparisons include those between individuals who become treated in the current period (*g = t*) with those who are not-yet-treated or never treated as well as comparisons to those who have been treated in an earlier time period. Newly treated individuals are compared to other individuals whose treatment status remains constant in that period. For individuals whose treatment status is already treated, if their treatment effect changes over time this change will be registered as a negative weight and reduces the size of the overall average treatment effect. Therefore, *TWFE* is only robust for treatment effect homogeneity. Or treatment effects that remain constant in future time periods once individuals are treated. Therefore, the interpretation of TWFE relies on highly restrictive assumptions in addition to the parallel trend's assumption criteria.

5.3 Callaway and Sant'Anna (2021) Difference-in-Differences (*CS-DID*).

A number of new estimators have been proposed to deal with these issues. The empirical methodology I follow for my study is a *DD* estimator developed by Callaway and Sant'Anna (2021), hereinafter referred to as *CS-DID*. The *CS-DID* is designed for empirical analyses in which individuals receive treatment at different times and where there may be heterogeneity in the treatment effects over time. Once they are treated, individuals remain treated, this is referred to as staggered treatment adoption.

This method requires two groups, a treatment and control group. The control group may contain people who are never treated or who are not-yet-treated. In my study everyone is eventually treated and I use the not-yet-treated individuals as my control group. This allows for similar unobservable individual characteristics that do not vary over time, such as work effort and ability, of the type of person who travels as a passenger and is involved in a road accident. Fadlon and Nielsen (2021) demonstrates that a treated group follows more similar pre-treatment trends when compared to a control group of individuals who undergo the same treatment at some stage in the future rather than a control group that is never treated. They demonstrate this even after controlling for covariates such as age and gender in the never treated control group. This lends support to the notion that individuals who experience the same treatment at some stage in the future make a more credible control group.

In my study, my population of interest experience their health shock at different times between 2005 to 2015. Before conducting my primary analysis, I create a panel with the data in event time to observe any pre or post trends, before and after the health shock. I use a balanced panel of 36 pre-periods and 36 post-periods, 73 event periods, or months, in total. The treatment time is normalised to time zero. In my study this is the month of the health shock. Therefore, all individuals experience the health shock at event time zero. Event time = -1 is the month before the event and event time +1 is the month after the event. The population of interest pre and post periods are the same in event or relative time but different in real or calendar time. While my panel is balanced in event time it is not balanced in calendar time. As the estimator always needs a g-1 and/or a valid comparison period the first 2 periods and the last 2 periods from my panel are not estimated. Consequently, this leaves me with estimations for 34 periods before and 34 periods after the treatment.

The *CS-DID* uses three different time periods, t is the calendar time, g is the cohort time and e, is the event time. Each individual belongs to a cohort, g, when they receive a treatment at the same time t. In my study t is the month of the health shock. If an individual is treated in the first time period (first month), t=1 then g=1, if they are treated in the second time period then g=2 and so on. Callaway and Sant'Anna (2021) then follow the canonical *DD* set up. All estimations are taken back to the pre and post comparison between g (at time t) and g-1 as each cohort moves through calendar time, t. This ensures there is no contamination from other time periods (Sun & Abraham, 2021). Contamination can occur due to heterogeneity of treatment effects in different calendar time periods. For example, each cohort, g, may have different demographic characteristics. There may also be different calendar time effects such as macro-economic factors like an economic downturn that can influence labour market outcomes. The *CS-DID* base estimator is the ATT(g,t). The average effect on the treated for cohort g at time t.

$$ATT\ (g,t) = E\left[Y_t - Y_{g-1}\middle|G = g\right] - E\left[Y_t - Y_{g-1}\middle|D_t = 0, G \neq g\right] \forall\ t \geq g \quad (9)$$

simply put, this is the average change in outcomes for cohort G=g at time t after the treatment occurs. The first term, $E\left[Y_t - Y_{g-1}\middle|G = g\right]$, is the difference between the average outcome Yt, for cohort G=g at time t and their baseline average outcome the month before treatment occurs at Y_{g-1}. The second term, $E\left[Y_t - Y_{g-1}\middle|D_t = 0, G \neq g\right]$, is the average outcome in period t for those who are not-yet-treated, $D_t = 0$ minus their average outcome in period g-1. D_i is a dummy variable where D_i = 1 if treatment has occurred at time *t* and D_i = 0 otherwise. The second term includes all those who are not a member of cohort g, G ≠ g and who are not yet treated, i.e., $D_t = 0$, for all t ≥ g. Any cohorts that are treated before cohort g, g<t, are dropped out of the regression as are those that become treated as cohort g moves through real time t to T - 1. For all the time periods after cohort G = g becomes treated.

The difference between these two terms is the causal estimate for the effect of treatment on cohort g at time t known as the "group-time average treatment effect" (Callaway & Sant'Anna, 2021). This is the base parameter used in my study and the building block that can be aggregated to form other semi-parametric estimators. As time, t, increases the size of the not-yet-treated group decreases, as more cohorts become treated and drop out of the regression, as does the precision of the estimator. Also, the distance from the pre-period g-1 increases and this is known as a "long distance" (Callaway & Sant'Anna, 2021).

Callaway and Sant'Anna (2021) propose an estimator based on the ATT(g,t) which identifies causal impacts for each cohort and then aggregates the individual ATT(g,t)'s in to a range of semi-parametric causal estimators. In the next section I explain the main specification and aggregated causal estimators used in this study.

5.3.1 Final Estimator and Main Specification

Finally, to evaluate how treatment effects evolve over time the main semi-parametric estimator used in this study aggregates the group time average effects (ATT(g,t) into event time average effects, $\theta_{D(e)}$. The average effect for all groups according to how long they have been exposed to the treatment or the event time. This is similar to the original event study *TWFE* model having removed any cross-period contamination.

In this study I estimate the dynamic treatment effects over the 34 months pre and post treatment. The overall average treatment effect in event time combines the treatment effects for the 34 months before the treatment occurs in to one average parameter. This is the vital parameter that indicates whether there are any significant pre-trends. Then the 34 months (event time) post treatment are combined into one average to establish if there are any significant average treatment effects in the 34 months after the treatment has occurred.

While I perform several robustness checks conditioning on pre-treatment covariates. In my main analysis, I adhere to the methodology proposed by Callaway and Sant'Anna (2021) without covariates, which requires unconditional parallel trends for staggered adoption. I also stratify my estimation samples by different observable characteristics to uncover any trends due to key covariates.

The dynamic specification used in this study can be expressed as:

$\theta_{D(e)}$: = $\sum_{g=2}^{T} 1\{\mathrm{g} + \mathrm{e} \leq T\}\mathrm{ATT}(g, g+e)\, P(G = g | G + \mathrm{e} \leq T\ )$ (10)

, where *g* denotes the month of the health shock and *T* is the total time periods in the study. The variable *P* represents the propensity score, or the likelihood of an individual being in a particular group or cohort. The variable *e* represents event time. The *ATT(g,g+e)* is the average treatment effect on the treated for cohort g at event time, g + e. Basically, equation 10 estimates the average treatment effect for all cohorts or groups that have been exposed to the treatment for e periods.

In order to create credible estimators three main assumptions need to be met as follows.

5.3.1.1 Assumption 1: Staggered Treatment Adoption

Basically, this states that once an individual is treated, they remain treated for the following periods. $D_{i,t}$ is a dummy variable that is equal to 1 if individual i is treated at time t and 0 otherwise. Staggered adoption means that for $t = 1 \ldots. T - 1, D_{i,t} = 1 \Rightarrow D_{i,t+1} = 1.$ T denotes the total number of calendar time periods. In this analysis once an individual has experienced the road accident or health shock this status does not change. $D_{it} = 1$ for individual *i* after the accident. This condition remains true for the remainder of the study period, adhering to the assumption of staggered treatment adoption throughout the analysis.

5.3.1.2 Assumption 2: Parallel Trends

The main assumption for the canonical *DD* is parallel trends. This implies that the average change in outcomes for the treated group would have been the same as for the untreated group if there was no causal impact of the treatment. The two groups would have continued to evolve in parallel if both had remained untreated. For all g ≥ t where g is the period (cohort) individuals become treated. This assumption is expressed in equation 11.

$$E\left[Y_{i,2}(0) - Y_{i,1}(0)\middle|D_i = 1\right] = E[Y_{i,2}(0) - Y_{i,1}(0)|D_i = 0] \text{ (11)}$$

where the first term, $E\left[Y_{i,2}(0) - Y_{i,1}(0)\middle|D_i = 1\right]$ is the average change in outcomes for the treated group, Di =1, between period 1 and 2 if they remain untreated. The second term, $E\left[Y_{i,2}(0) - Y_{i,1}(0)\middle|D_i = 0\right]$ is the average change in outcomes for the control or in this case not-yet-treated group, $D_i = 0$, between periods 1 and 2 if they are untreated. If there was no health shock then the above equation would hold.

In my main analysis, I adhere to the methodology proposed by (Callaway & Sant'Anna, 2021) without covariates, which requires unconditional parallel trends for staggered adoption. My primary specification, detailed in equation 12, estimates the Average Treatment effect on the Treated (ATT) for the full population.

$$ATT\ (g,t) = E\left[Y_t - Y_{g-1}\middle|G = g\right] - E\left[Y_t - Y_{g-1}\middle|D_t = 0, G \neq g\right] \forall\ t \geq g \text{ (12)}$$

Basically, this estimates the average causal effect on the treated group G = g at time t compared to the not-yet-treated group Dt =0 and G≠g. This is the difference between the average outcomes for cohort g before (g-1) and after the health shock/treatment measured at period t minus the difference in average outcomes before and after these same periods for those not-yet-treated, $G_i = 0$ in period t.

This ATT(g,t) is then aggregated into a combined ATT(g,t) for all cohorts. Then finally into dynamic event time semi-parametric estimators, 34 months before and 34 months after the treatment. This core analysis spans 34 months before and 34 months after the treatment. This allows for the dynamic assessment of health shock effects over a 34-month period before and after the shock. Subsequent specifications explore the effects using different levels of health shock severity. Importantly, no significant combined pre-trends are observed, enhancing the credibility of this methodology for the study. This lack of pre-trends suggests that, in the absence of treatment, average outcomes would evolve in parallel.

Parallel trends allow for selection bias as long as this bias is the same in both or all periods. This could include individual characteristics that are not impacted by the treatment and that remain the same in both periods. This implies they are mean-independent of any variables that affect trends in the outcome (Roth et al., 2023)

5.3.1.3 Assumption 3: No Anticipation

This assumption states that individuals do not have any prior knowledge about the treatment, preventing them from adjusting their behaviour in anticipation of it. If individuals had already adjusted their behaviour before treatment occurred then we are not measuring the true causal effect of the treatment (Abbring & Gerard.J. van den Berg., 2003; Malani & Reif, 2015). In the context of this study, the assumption is reasonable as individuals are unlikely to anticipate road accidents and passengers are unlikely to have any control over the accident, they are involved in. To mitigate the possibility of confounding influences from individual-specific, unobserved characteristics, the analysis excludes individuals involved in previous road accidents or with a hospital stay due to an accident in the three years before the health shock. This exclusion helps control for potential risk-seeking individuals or those regularly associated with risk-taking individuals.

In my study this means that $ATT(g,t) = 0$ for all pre-treatment periods when $t < g$. There is no average causal impact before treatment occurs (Callaway & Sant'Anna, 2021).

Next, I proceed with my results and final analysis.

6. Empirical Results

In this section I answer my research question: how does an exogenous health shock affect an individual's labour market outcomes? My analysis presents an empirical study of the labour market outcomes after a road accident. The treatment in this study is a health shock which occurs due to a road accident. All these terms are used interchangeably in the following discussion.

First, I analyse estimations for the outcome variables for employment and earnings. A binary indicator of employment, the log level of monthly wages and salary, and finally the log level of monthly wages and salary for those who are employed. I check for pre-trends in the 34 months before the health shock to empirically test for the underlying identifying assumption of parallel trends criteria. Then, I discuss average and dynamic treatment effects in the 34 months after treatment occurs. I classify the seriousness of the health shock and estimate the outcomes for employment and earnings at each health shock level. Next, I consider treatment effect heterogeneity with regard to gender and ethnicity. Furthermore, I classify my empirical analysis by the extent of pre-accident labour market attachment. To test if potential changes in labour market engagements in post-accident months are simultaneously associated with variation in dependency on governmental benefits, I then examine the effect of the health shock on benefit recipiency. Lastly, I present robustness checks. First, I include time-invariant covariates to control for individual characteristics that do not change over time. Second, I use an alternative definition of health shock severity in additional specifications.

6.1 Primary Methodology: Callaway and Sant'Anna (2021; *CS-DID*)

The primary aim of this research is to identify the causal effect of an exogenous health shock (a road accident) on an individual's labour market outcomes. The key methodology used is the *CS-DID*, a staggered difference in differences technique. Staggered adoption of treatment in the context of the current analysis means individuals have their accident or health shock at different times. Once they have their accident the impacts of the health shock remain. Once a person is treated, they remain treated.

6.1.1 Primary Specification

For my baseline specification, I estimate the *CS-DID* specification without any covariates. However, I also test the robustness of my analysis by including time-invariant individual-level characteristics, as suggested by Callaway and Sant'Anna (2021). These findings are in some cases exactly the same otherwise similar to my main specification.

6.1.1.1 Average Treatment Effects: Employment and Earnings

First, I discuss the average treatment effects for the entire post-accident time window of the key outcome variables employment and earnings for the full population (see Table 4). The average treatment effects on monthly employment are in column 1. Next, log earnings where I set non-employed to 0 in column 2 and finally log earnings of the employed individuals in column 3. The first row of results presents the overall average treatment effect in the 34 months before the health shock. The second row shows the average treatment effect for 34 months after the health shock.

It's imperative first to check the vital parallel trends assumption. Looking at the first row of Table 4, there are no combined significant pre-trends for any of these outcome variables. This lack of pre-trends suggests that average outcomes would evolve in parallel in the absence of treatment.

Next, the post-health shock average trends are presented in the second row of Table 4. The average trends are statistically significant at the 99 percent confidence level. For the full population of interest, the probability of employment decreases on average by 4.90 percentage points in the post-accident period, as reported in column 1. To put this number into perspective: the pre-accident employment rate observed at the month just prior to the accident (t-1) is 54.8 percent. Thus, the 4.9 percentage point drop translates into a 9 percent drop (on average) in employment in the 34 months after the health shock. The log earnings in column 2 decline on average by -0.407 after the health shock. Note that this earnings marker includes non-employed individuals (their earnings is set to zero). To disentangle whether the drop in employment primarily drives the earnings effect, I re-run the regression for those who remain employed. Their monthly log earnings drop on average by -0.033 over 34 months, which corresponds to a 3.3 percent drop on average. This indicates a causal impact of the health shock on employment and earnings.

The results suggest that there is a significant number of people leaving the labour market after experiencing a health shock. For those who remain employed, their earnings tend to be lower on average. This could be attributed to various factors such as fewer working hours or transitioning to a different job within their capabilities as a consequence of ongoing negative effects of the health shock.

Table 4: Average treatment effect estimates using the CS-DID

Full Population	Employment (1)	Log Earnings (2)	Earnings Conditional (3)
Pre-average (34 months)	-0.000 (0.000)	-0.002 (0.002)	0.001 (0.001)
Post-average (34 months)	-0.049*** (0.005)	-0.407*** (0.039)	-0.033*** (0.01)
Sample size	8292	8292	8292

Notes: The above estimates for the full population of interest are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Column (1) is the estimate for a binary indicator of employment, column (2) a continuous measure of average monthly wages and salary, column (3) a continuous measure of average monthly wages and salary for those who are employed. While my panel is balanced in event time it isn't balanced in calendar time. As the estimator always needs a g-1 and/or a valid comparison period the first 2 periods and the last 2 periods from my panel are not estimated. Consequently, this leaves me with estimations for 34 periods before and 34 periods after the treatment. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Source: Own calculations using Statistics NZ's IDI

6.1.1.2 Average Dynamic Treatment Effects: Employment and Earnings

Table 4 provides evidence that a health shock can cause disruptions on employment and earnings. However, from the average effects presented in Table 4, it is unclear how this evolves over time. For example, employment might drop in the first months after the accident while the individual is recovering and return to pre-treatment levels. Another scenario is that the labour market effects build up over time. Figures 3 and 4 show the dynamic pre and post labour market effects for employment and earnings.

Figure 3 below visually represents the dynamic average treatment effects for the 34 months before and after an accident-induced health shock. It depicts the percentage point change in monthly employment propensity on average for all cohorts in event time. The red vertical line at event time zero indicates the month the health shock occurs. Each estimate after the health shock is compared to the baseline period the month before treatment, g-1. The red vertical line drawn at zero allows us to check for any pre-trends. The y-axis measures the percentage point change in the likelihood of employment. The x-axis measures event time. The 95 percent confidence intervals are also displayed.

First, inspecting employment probability for the 34 months before the treatment occurs in Figure 3, we can see there are almost no significant effects (see Appendix B: Table 21). This is apparent by the confidence intervals overlapping zero. There are slightly significant decreases in the likelihood of monthly employment in months -33, -22 and -9 of -0.006, -0.006 and -0.005 respectively before the accident. Along with slight significant increases of +0.005 at -33 months before and +0.005 at -2 months before the accident. However, the average difference of all coefficients in the pre-period (F-statistic) is not statistically significant from zero (see table 4[14]). This provides empirical support to the parallel trends' assumption criteria. In other words, the pre-health shock employment trend is indicative of no anticipatory behaviours in the prior months leading up to the month of the road accident.

Next, I check the post-accident trends. For the full population sample, Figure 3 shows the probability of employment decreases on average 4.9 percentage points over the 34 months after the accident compared to the baseline month, the month before the incidence of the health shock. There is a significant 3 percentage point drop in employment propensity in the month after the health shock. This persists for the next 10 months and then decreases further

[14] See Table 4 column 1 = 0.000

to 4 percentage points from months 11 to 16. It then gradually declines to a 9-percentage point drop in month 34. This demonstrates that the full population on average doesn't recover after the health shock and experiences an ongoing decline in the probability of employment. This could be due to ongoing negative effects of the health shock. Health shocks can have longer term impacts than just the immediate consequences (see Appendix B: Table 21).

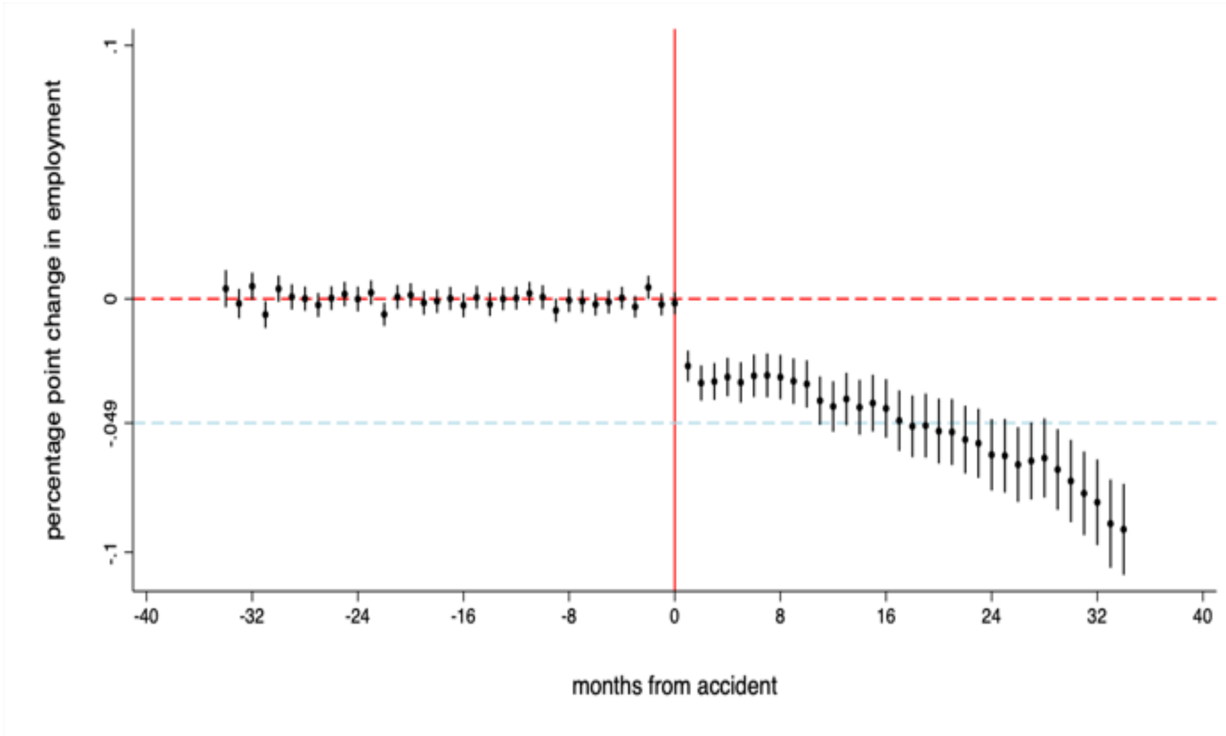
Next, I discuss results for average monthly log earnings seen in Figure 4. We can see, as for average monthly employment, average log earnings decline significantly after the health shock as well. Overall, the population of interest suffers a 0.407 drop in average log earnings in the 34 months after the health shock (see Appendix B: Table 22).

Once again, we see a similar pattern in the dynamic effects of the health shock on monthly average log earnings in Figure 4. With no combined significant pre-trends identified, average monthly earnings drop significantly after the health shock. For the full population in Figure 4, there is a significant drop of -0.255 in average log earnings the first month after the health shock. There is also a significant drop during month zero due to individuals experiencing their accident at different times during the accident month. By month 12 after the accident this drop is -0.343 and by month 34 -0.739 average decline in monthly log earnings (see Appendix B: Table 22).

We can see from Figures 3 and 4 that post-accident labour market outcomes worsen in the 34 months after the accident. This demonstrates ongoing and perhaps delayed impacts of the health shock on physical and mental health. This could be due to factors such as people returning to work before they are ready (Theadom. A et al., 2017). Alternatively, since there is usually a co-payment associated with most follow up care. This could be a disincentive for some individuals to seek ongoing treatment which may impact their recovery and future health and employment.

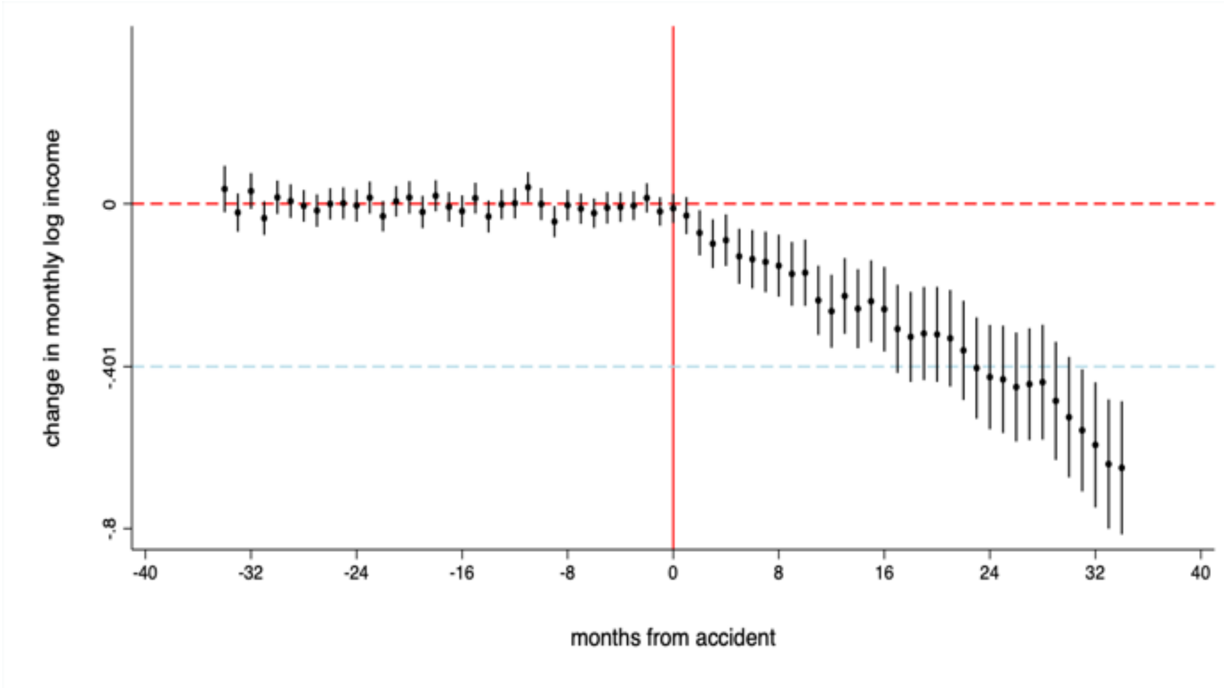
There may also be other patterns in the population of interest. The population of interest is 69 percent female and women are typically less attached to the labour market than men. As of June 2023, 65.4 percent of women aged 15 years and above were employed compared to 74.4 percent of men (Ministry for Women - Manatū Wāhine, 2023). Women’s monthly earnings on average are lower than men and they are less likely to be the main income earner in a household. Fadlon and Nielsen (2021) found that while women were more likely to increase their labour supply after a fatality in the household, men were more likely to decrease labour supply due to supporting one less person in the household. Men still tend to be the main income earners in households. If a person isn’t the main income earner in the household, it may be easier to reduce their labour supply or leave the labour market altogether due to the health shock.

Figure 3: Monthly Employment



Notes: The above estimates for the full population of interest are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Figure 3 visually represents the dynamic average treatment effects for the 34 months before and after an accident-induced health shock. It depicts the percentage point change in monthly employment propensity on average for all cohorts in event time. The red vertical line at event time zero indicates the month the health shock occurs. Each estimate after the health shock is compared to the baseline period the month before treatment, g-1. The red vertical line drawn at zero allows us to check for any pre-trends. The y axis measures the percentage point change in the likelihood of employment. The x axis measures event time. The 95 percent confidence intervals are also displayed. While my panel is balanced in event time it isn't balanced in calendar time. As the estimator always needs a g-1 and/or a valid comparison period the first 2 periods and the last 2 periods from my panel are not estimated. Consequently, this leaves me with estimations for 34 periods before and 34 periods after the treatment.
*** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

Figure 4: Average monthly log income



Notes: The above estimates for the full population of interest are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Figure 4 visually represents the dynamic average treatment effects for the 34 months before and after an accident-induced health shock. It depicts the percentage point change in monthly log income on average for all cohorts in event time. The red vertical line at event time zero indicates the month the health shock occurs. Each estimate after the health shock is compared to the baseline period the month before treatment, g-1. The red vertical line drawn at zero allows us to check for any pre-trends. The y axis measures the percentage point change in monthly log income. The x axis measures event time. The 95 percent confidence intervals are also displayed. While my panel is balanced in event time it isn't balanced in calendar time. As the estimator always needs a g-1 and/or a valid comparison period the first 2 periods and the last 2 periods from my panel are not estimated. Consequently, this leaves me with estimations for 34 periods before and 34 periods after the treatment. *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

6.1.1.3 Health Shock Severity

Another key aim of this study is to investigate the impact of more severe health shocks. The rich administrative data in the IDI allows me to define the seriousness of the health shock in two different ways (see section 4.5.2 for more details). The ACC data details information about the cost of each person's ACC claim. This includes medical and rehabilitation costs (see section 4.5.2 1 for more detailed explanation). Following the assumption that a more serious health shock pertains to a greater cost, I use total cost as my main definition of the health shock.

There are five estimations in total, the full population of interest as reference and three levels of seriousness of the health shock. The fifth estimation combines the medium and severe health shock into one sample. Some individuals, if they meet the criteria, are entitled to income compensation. This is included in the rehabilitation cost variable. If the income compensation amount could be deducted some of the severe shock group may move back to a less severe health shock level. To account for this in parts of my analysis I combine the medium and severe shock in to one category. This also acts as a robustness check with a larger sample size of individuals experiencing a more severe health shock.

The first row of tables 5,6 and 7 presents the average change in employment and income before the health shock. There are no significant values and the pre-trends assumption is met. This indicates that, on average, there were no substantial trends in employment and earnings before the occurrence of the health shock at any level of health shock severity.

In column 2 of table 5, the second row details the post-health shock average change in the likelihood of employment. For the individuals who experience a mild health shock, there is a 3.7 percentage point decrease in the likelihood of being employed. This decline could also be interpreted as an average dip in the likelihood of employment of 6.9 percent relative to the pre-accident employment rate. The medium health shock results in a 6.1 percentage point decrease or an 11 percent average drop, column 3. Finally, the severe health shock causes a decline in monthly employment of 17.9 percentage points on average translating to a 26 percent drop in the likelihood of employment, column 4. The combined group experiencing medium and severe health shocks (column 5) shows an 8.2 percentage point decrease, equivalent to a 14.6 percent drop in the likelihood of employment.

Likewise, row 2 of table 6 shows significant and increasing declines in average monthly log earnings for each health shock level. The mild shock there is a drop of 0.306, medium 0.487 and severe 1.552 in average monthly log earnings. Finally, table 7 shows that for a mild health shock those who remain employed experience a 1.4 percent drop in earnings on average. This decline is 4.6 percent for a medium shock and 25.2 percent for the most severe shock.

In summary, the detailed analysis provides robust evidence that health shocks have escalating effects on both employment and earnings. More severe shocks lead to more pronounced declines. This also underpins what the descriptive statistics demonstrated in section 4.5.

These results are in line with other literature. As the health shock worsens so do labour market outcomes (Crichton et al., 2011; Dano, 2005; Parro & Pohl, 2021)

Table 5: Average treatment effect estimates using the CS-DID Employment

	Full Population (1)	Mild Health Shock (2)	Medium Health Shock (3)	Severe Health Shock (4)	Medium plus Severe (5)
Pre-average (34 months)	0.000 (0.000)	0.000 (0.000)	0.000 (0.001)	0.002 (0.002)	0.000 (0.001)
Post-average (34 months)	-0.049*** (0.005)	-0.037*** (0.006)	-0.061*** (0.011)	-0.179*** (0.024)	-0.082*** (0.01)
Sample size	8292	6042	1827	423	2250

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Column (1) is the estimate for the full population, column (2) those who had a total cost of less than $700 (NZD), column (3) individuals who had a total cost of ≥700 and <9000, column (4) individuals whose total cost was ≥9000, column (5) total cost of ≥700. For further explanation about the total cost variable see section 4.5.2.1 *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

Table 6: Average treatment effects on monthly log earnings

	Full Population (1)	Mild Health Shock (2)	Medium Health Shock (3)	Severe Health Shock (4)
Pre-average (34 months)	0.000 (0.000)	-0.003 (0.002)	0.000 (0.005)	-0.001 (0.001)
Post-average (34 months)	-0.407*** (0.039)	-0.306*** (0.046)	-0.487*** (0.084)	-1.552*** (0.189)
	8292	6042	1827	423

Notes: The above estimates for a continuous measure of average monthly wages and salary are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Column (1) is the estimate for the full population, column (2) those who had a total cost of less than $700 (NZD), column (3) individuals who had a total cost of ≥700 and <9000, column (4) individuals whose total cost was ≥9000, column (5) total cost of ≥700. For further explanation about the total cost variable see section 4.5.2.1 *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

Table 7: The average treatment effect of a health shock on earnings conditional on employment

	Full Population (1)	Mild Health Shock (2)	Medium Health Shock (3)	Severe Health Shock (4)
Pre-average (32 months)	-0.001 (0.001)	0.000 (0.001)	0.002 (0.002)	0.003 (0.003)
Post-average (32 months)	-0.033*** (0.01)	-0.014*** (0.011)	-0.046*** (0.021)	-0.252*** (0.037)
	8292	6042	1827	423

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Column (1) is the estimate for the full population, column (2) those who had a total cost of less than $700 (NZD), column (3) individuals who had a total cost of ≥700 and <9000, column (4) individuals whose total cost was ≥9000, column (5) total cost of ≥700. For further explanation about the total cost variable see section 4.5.2.1 *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

6.1.2 Demographic Heterogeneity: Gender and Ethnicity

Following the main analysis, I stratify my sample by male and female and ethnicity (European versus non-European). This is to check for any significant differences between gender and ethnic groups. First, I estimate the average treatment effects for these groups after the health shock for employment and average monthly log earnings (see Appendix C: Table 23 and 24).

Next, I conduct t tests. I compare the difference between average treatment effects for the 34 months after the health shock between these groups. First, for average monthly employment and second for log earnings. This determines which groups may be more impacted by the health shock. The significant results are presented below in Tables 8 and 9 (see Appendix C: Table 25 and 26 for full table). While gender differences were observed, they did not reach statistical significance. Notably, the most substantial disparity that emerged between males and females was a drop of 20.8 percentage points in the probability of employment for females compared to a 16.3 percentage point for males. Moreover, in the case of severe health shocks, women encountered a decline in average monthly log earnings of 1.74 and males 1.32. These results indicate that there may not be significant differences between male and female attachment to the labour market as discussed above in this instance.

However, for ethnicity, non-Europeans experience significantly worse impacts than Europeans. Table 8 contains the estimations for average monthly employment after the health shock and at each health shock level. The third row shows us the difference and whether these are significant. All results are significant at the 99 percent confidence level. Non-Europeans, for the overall population, experience a 3.1 percentage point greater decline in employment displayed in column 1. For the mild health shock group in column 2 the difference is 1.8 percentage points, 6.8 for medium column 3 and 7.7 for severe column 4. All these differences are significant.

Similarly, in table 9 the decline in average monthly log earnings is significantly greater for non-Europeans than Europeans. An overall 25.5 percent greater decline in average monthly log earnings. The mild shock experiences a 14.2 percent decline, 52.9 percent for the medium and 68.8 percent for the severe health shock group.

This is consistent with changes to the likelihood of employment and average monthly earnings for the full population and at each level of severity. This may be partly explained by a lack of willingness to access healthcare services for non-Europeans. Studies have shown that Māori and Pasifika are less likely to access primary and child healthcare compared to NZ- Europeans (Bourke et al., 2023; Lewycka. S. et al., 2023). Another factor could be co-payments acting as a disincentive to seek follow up care, as non-Europeans earn less on average than Europeans, these co-payments may be less affordable. Type of work may also play a role. Māori are more prevalent in lower skilled, labour intensive industries which can make it more difficult for them to return to work after an accident. (Ministry of Business Innovation and Employment, 2021). They may be physically unable to do the job.

Table 8: Differences in average monthly employment after treatment by ethnicity

	Full Population (1)	Mild Health Shock (2)	Medium Health Shock (3)	Severe Health Shock (4)
European	-0.035***(0.007)	-0.029***(0.008)	-0.029*(0.015)	-0.153***(0.035)
non-European	-0.066***(0.007)	-0.047***(0.009)	-0.097***(0.016)	-0.23***(0.034)
Difference	0.031***(0.01)	0.018*(0.012)	0.068***(0.022)	0.077*(0.049)

Notes: The above estimates in row one and two are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Row 1 are the estimates for the full population of European only individuals. Row 2 are the estimates for the full population of non-European only individuals. Row 3 are the differences between row 1 and 2. Column (1) is the estimate for the full population, column (2) those who had a total cost of less than $700 (NZD), column (3) individuals who had a total cost of ≥700 and <9000, column (4) individuals whose total cost was ≥9000, column (5) total cost of ≥700. For further explanation about the total cost variable see section 4.5.2.1 *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Source: Own calculations using Statistics NZ's IDI

Table 9: Differences in monthly log earnings after treatment by ethnicity

	Full Population (1)	Mild Health Shock (2)	Medium Health Shock (3)	Severe Health Shock (4)
European	-0.289***(0.054)	-0.241***(0.062)	-0.243** (0.114)	-1.233*** (0.253)
non-European	-0.544***(0.058)	-0.383***(0.067)	-0.772*** (0.123)	-1.921*** (0.266)
Difference	0.255***(0.079)	0.142* (0.091)	0.529*** (0.167)	0.688**(0.368)

Notes: The above estimates in row one and two are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Row 1 are the estimates for the full population of European only individuals. Row 2 are the estimates for the full population of non-European only individuals. Row 3 are the idfferences between row 1 and 2. Column (1) is the estimate for the full population, column (2) those who had a total cost of less than $700 (NZD), column (3) individuals who had a total cost of ≥700 and <9000, column (4) individuals whose total cost was ≥9000, column (5) total cost of ≥700. For further explanation about the total cost variable see section 4.5.2.1 *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Source: Own calculations using Statistics NZ's IDI

6.1.3 Labour Market Attachment

In this section I discuss labour market attachment. I compare the full population of interest, which contains employed and unemployed individuals, with those who had positive earnings (classified as employed) in the 12 months before the health shock. Referring first to tables 27,28,29 (see appendix D) there are no combined significant trends in the 34 months before the health shock.

Table 10 below, row 3 highlights the disparity between these two groups and whether or not it is statistically significant. Notably, for individuals employed in the 12 months prior to the health shock, there is a substantial 10.6 percentage-point greater decline in monthly employment propensity (column 1). Additionally, they experience a significantly greater decline of 0.813 log units in average monthly earnings (column 2). However, there is no statistically significant difference in the decline in earnings conditional on employment (column 3).

These findings align with existing literature (Crichton et al., 2011; Parro & Pohl, 2021), highlighting the differential impact of health shocks on labor market outcomes based on the level of prior attachment.

Table 10: Labour market attachment, differences in outcomes

	Employment (1)	Average Log Income (2)	Conditional on Employment (3)
Full Population	-0.049*** (0.005)	-0.407*** (0.039)	-0.033*** (0.01)
Employed	-0.155*** (0.007)	-1.22*** (0.052)	-0.03*** (0.01)
Difference	0.106*** (0.008)	0.813*** (0.064)	-0.003 (0.015)

Notes: The above estimates in row one and two are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Row 1 are the estimates for the full population of interest. Row 2 are the estimates for those individuals who had positive earnings in the 12 months before the accident. Row 3 are the differences between row 1 and 2. Column (1) is the estimate for a binary indicator of employment, column (2) a continuous measure of average monthly wages and salary, column (3) a continuous measure of average monthly wages and salary for those who are employed. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Source: Own calculations using Statistics NZ's IDI

Finally, I discuss the fourth outcome variable, benefit recipiency.

6.1.4 Benefit Recipiency

The final outcome of interest is Benefit Recipiency. This is a dummy indicator that equals 1 if a person received a benefit[15] in a given month and 0 otherwise. This allows me to see if there is a change in the average likelihood of benefit recipiency but not specifically which benefits. This outcome is of interest as New Zealand has a relatively robust social security system. An increase in benefits can indicate ongoing health issues. There can also be a disincentive to going back to work if individuals are able to secure a job seeker benefit (Garcia-Gomez, 2011). This may confound the true impact of a health shock.

[15] unemployment, solo parent (DPB), sickness, invalid, widow and emergency benefit

The ACC no fault claims system removes some barriers that may exist in other countries to securing health care after a health shock. However, Crichton et al. (2011) find that this system may not function as fully as intended. Individuals experiencing longer ACC spells (more severe health shocks) have longer term labour market effects according to this study.

With regards to benefit recipiency once again there are no significant pre-trends before the health shock as recorded in row 1 of table 30 (see appendix D). However, there are significant findings for those who suffered a severe health shock (see Appendix D: row 2 column 4, Table 30). They experience a 5.3 percentage point increase in average benefit recipiency over the 34 months. This could indicate that the system is largely successful at supporting those with less severe health shocks.

After a health shock those who are entitled to ACC income compensation[16], can receive these payments indefinitely[17]. Therefore, this increase in benefits may be partly due to those who are assessed as fit enough to go back to work but continue to experience problems[18]. Alternatively, this may be those who were not entitled to income compensation and struggled to gain or remain in employment. Our jobseeker support provides some financial support to those who are unemployed or sick[19]. However, it is a flat rate, therefore there is an incentive to get back to work to earn a greater income if a person is able. For lower income individuals there may be less incentive to return to work if their incomes were not much greater than the jobseeker benefit.

[16] This can be up to 80% of their income.
[17] If they have a medical certificate
[18] Once a medical professional assesses them as fit for work
[19] The minimum weekly amount is $250.63 for a 18 to 19 year old living at home, they can earn up to $519 gross per week before this is reduced to zero support.

6.2 Robustness Checks

I perform two robustness checks. Firstly, I include time-invariant covariates in an additional specification. Secondly, I use public hospital discharges data to define a second proxy for a health shock.

6.2.1 Regressions Including Covariates (age, ethnicity, gender)

In this supplementary analysis, I include covariates age (as at the accident date), gender and ethnicity. These estimations are presented below in table 11 and 12. Table 11 illustrates the average monthly employment propensity, while table 12 depicts the average monthly log earnings. Notably, the results align closely with those presented in Tables 4, 5 and 6 which did not account for time-invariant covariates.

In both tables, the first row displays the average outcomes in the 34 months preceding the treatment. Importantly, there are no discernible pre-trends in this period before the accident.

Examining table 11, the second row indicates a 4.9 percentage-point decline in monthly employment propensity for the entire population following the health shock. For the mild health shock this decline is 3.7 percentage points. These estimates are the same as the primary specification (see Tables 4 and 5). The effects for medium and severe shocks are 5.9 and 18.6, respectively. These are comparable to the 6.1 and 17.9 percentage points observed in the main specification.

Moving to table 12, average monthly log earnings decline by 0.406, 0.304, 0.473 and 1.594 for the entire population of interest, mild, medium and severe health shocks. These values closely resemble the figures of 0.407, 0.306, 0.487 and 1.552 from the main specification in Tables 4 and 6.

Table 11: Monthly Employment

	Full Population (1)	Mild Health Shock (2)	Medium Health Shock (3)	Severe Health Shock (4)
Pre-average (34 months)	0.000 (0.000)	0.000 (0.000)	0.000 (0.001)	-0.002 (0.002)
Post-average (34 months)	-0.049*** (0.005)	-0.037*** (0.006)	-0.059*** (0.011)	-0.186*** (0.023)
Sample size	8292	6042	1827	423

Notes: The above estimates for a binary indicator of monthly employment are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Column (1) is the estimate for the full population, column (2) those who had a total cost of less than $700 (NZD), column (3) individuals who had a total cost of ≥700 and <9000, column (4) individuals whose total cost was ≥9000. For further explanation about the total cost variable see section 4.5.2.1 *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

Table 12: Monthly log earnings

	Full Population (1)	Mild Health Shock (2)	Medium Health Shock (3)	Severe Health Shock (4)
Pre-average (34 months)	-0.003 (0.002)	-0.008 (0.01)	-0.002 (0.005)	-0.016 (0.011)
Post-average (34 months)	-0.406*** (0.039)	-0.304*** (0.046)	-0.473*** (0.083)	-01.594*** (0.181)
Sample size	8292	6042	1827	423

Notes: The above estimates for average monthly log income are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Column (1) is the estimate for the full population, column (2) those who had a total cost of less than $700 (NZD), column (3) individuals who had a total cost of ≥700 and <9000, column (4) individuals whose total cost was ≥9000. For further explanation about the total cost variable see section 4.5.2.1 *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

6.2.2 Alternative Definition of a Health Shock

As a final robustness check, I employ an alternative definition of a health shock. I use the public hospital discharges data (see section 4.5.2.2. for further details). I use the length of stay in hospital as a second proxy for the severity of the health shock, aligning with established literature (Dano, 2005; Halla & Zweimüller, 2013; Parro & Pohl, 2021). The primary objective of this robustness check is to assess the impact of the health shocks at varying levels of severity. A longer length of stay in hospital serves as a proxy for an increased severity of the health shock.

The outcomes of this analysis are presented in Table 13, 14 and 15. The first row of each table indicates the no pre-trends criteria is met. The second row presents the estimations for the likelihood of employment, average monthly income and average monthly earnings dependent on employment respectively. These results illustrate that employment and earnings outcomes in the 34 months after the road accident worsen as the health shock level worsens.

Specifically in Table 13, row 2, the probability of being employed experiences a decline of 4.6, 9.2 and 12.3 percentage points for mild, medium and severe shocks, respectively. Regarding average monthly log earnings, there is a decrease of 0.385, 0.734 and 0.998 for mild, medium, and severe health shock levels. Lastly, monthly earnings conditional on employment decline by 0.032, 0.09 and 0.166 log units after a mild, medium and severe health shock.

Consistent with the main specification in table 4, 5 and 6, the findings indicate that a health shock, on average, leads to a decline in employment and earnings over the 34 months following the shock. Importantly, the impact on the labour market outcomes intensifies with the severity of the health shock.

Table 13: Estimation of the average effect of a health shock on employment

	Full Population (1)	Mild Health Shock (2)	Medium Health Shock (3)	Severe Health Shock (4)
Pre-average (34 months)	-0.001 (0.00**1**)	0.000 (0.000)	-0.001 (0.002)	0.000 (0.003)
Post-average (34 months)	-0.049*** (0.005)	-0.046*** (0.005)	-0.092*** (0.024)	-0.123*** (0.037)
Sample size	8292	7800	498	234

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month. Column (1) is the estimate for the full population, column (2) those who didn't go to hospital or had 0- or one-night's stay in hospital, column (3) individuals who had more than one night's stay in hospital, column (4) individuals who had more than 5 night's stay in hospital *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

Table 14: Estimation of the average effect of a health shock on average log earnings

	Full Population (1)	Mild Health Shock (2)	Medium Health Shock (3)	Severe Health Shock (4)
Pre-average (34 months)	--0.002 (0.002)	-0.002 (0.002)	-0.008 (0.01)	-0.002 (0.014)
Post-average (34 months)	-0.407*** (0.039)	-0.385*** (0.04)	-0.734*** (0.183)	-0.998*** (0.281)
Sample size	8292	7800	498	234

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month. Column (1) is the estimate for the full population, column (2) those who didn't go to hospital or had a 0-, or one-night's stay in hospital, column (3) individuals who had more than one night's stay in hospital, column (4) individuals who had more than 5 night's stay in hospital *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

Table 15: Estimation of the average effect of a health shock on average log earnings if employed

	Full Population (1)	Mild Health Shock (2)	Medium Health Shock (3)	Severe Health Shock (4)
Pre-average (34 months)	-0.001 (0.001)	0.001 (0.001)	0.001 (0.004)	0.006 (0.005)
Post-average (34 months)	-0.033*** (0.01)	-0.032*** (0.01)	-0.09* (0.05)	-0.166** (0.074)
Sample size	8292	7800	498	234

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month. Column (1) is the estimate for the full population, column (2) those who didn't go to hospital or had 0- or one-night's stay in hospital, column (3) individuals who had more than one night's stay in hospital, column (4) individuals who had more than 5 night's stay in hospital *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

The two robustness checks undertaken provide additional support for the main findings of this study. The descriptive statistics in section 4.6 also mirror these results. Notably, health shocks exhibit a causal and enduring influence on labour market outcomes. The observed impacts demonstrate a worsening trend corresponding to the severity of the health shock. Importantly, these results underscore the robustness of my findings across various model specifications.

7. Conclusion

This research aims to determine the causal effect of an exogenous health shock on labour market outcomes. Due to an endogeneity problem, defining and measuring an exogenous health event is challenging. The unique New Zealand no-fault Accident Insurance system allows me to access a broad population of interest using administrative ACC data. I define an exogenous health shock using ACC data. I identify individuals who were travelling as a passenger in a road accident. Thus, utilising a more exogenous health shock than previous studies. I also use the cost of each person's ACC claim to measure the seriousness of their health shock. Next, I link my sample with tax data from Inland Revenue (IR) that contains detailed information on employment, earnings and benefit recipiency. To quantify the health event's labour market impact, I exploit the timing variation of the road accident.

Due to under-identification issues with the popular TWFE difference-in-differences model, new and improved methods have been developed. I use a method developed by Callaway and Sant'Anna (2021) that allows for treatment heterogeneity and variation in timing of treatment. My control group consists of those individuals who are not-yet-treated.

Individuals receive some support to recover from their health shock and return to the labour market. New Zealand's social security system provides a safety net that encompasses income compensation and jobseeker support, for those who are eligible, after a health shock. The checks in the system[20] and a flat rate for job seeker support reduce the likelihood of individuals staying out of the workforce as a result of this government support. This lowers the risk of overestimating the negative impact of health shocks. In fact, the support available means that estimated impacts may be less severe compared to countries with less support.

My findings in a nutshell are that a health shock, on average, worsens individuals' future labour market outcomes. For the full population of interest, the probability of employment decreases

[20] Individuals need to have a medical certificate to continue receiving income compensation or receive the job seeker benefit due to being unwell.

on average by 4.9 percentage points and monthly income declines on average by 0.41 log units in the first 34 months after the road accident. The results show an immediate labour market impact, which persists and worsens over the longer term. Individuals who remain employed after the health shock experience, on average, a 3.3 percent decline in their monthly earnings. These negative effects are greater for more severe health shocks, those more attached to the labour market and for non-Europeans. I find an increase in benefit recipiency for those who experience a severe health shock.

These findings suggest that health shocks do indeed have a significant and direct impact on both employment status and earnings levels. This effect is immediate and worsens over the longer term. This information could have implications for policymaking or interventions aimed at mitigating the effects of health shocks on individual's economic well-being. In the New Zealand context, these findings align with Crichton et al. (2011), indicating that institutional support does not fully prevent negative labour market outcomes after a health shock.

Limitations in this study exist because some healthcare costs are bulk funded. Therefore, the total cost variable may not include all individual healthcare costs. The small sample size means it's not possible to separate individuals further into groups by injury type. There may also be mental health impacts that are not measured in this study but could be interesting to investigate in the future.

8.Reference List

Abbring, J. H., & Gerard.J. van den Berg. (2003). The nonparametric identification of treatment effects in duration models. *Econemetrica, 71*, 1491-1517.

ACC New Zealand. (n.d.-a). *Injuries we cover*. https://www.acc.co.nz/im-injured/what-we-cover/injuries-we-cover/#what-is-no-fault-cover

ACC New Zealand. (n.d.-b). *Treatment we pay for*. https://www.acc.co.nz/im-injured/what-we-cover/treatment-we-pay-for/#paying-for-your-treatment

Accident Compensation Act. (2001).

Accident Compensation Corporation. (2023). *Our History*. https://www.acc.co.nz/about-us/who-we-are/our-history/

Accident Compensation Corporation. (n.d.-a). Types of ongoing support. https://www.acc.co.nz/im-injured/types-of-ongoing-support/

Accident Compensation Corporation. (n.d.-b). *Weekly compensation*. https://www.acc.co.nz/im-injured/financial-support/weekly-compensation/

Accident Compensation Corporation. (n.d.-c). *What we cover*. https://www.acc.co.nz/im-injured/what-we-cover

Angrist, J. (2005). Instrumental variables methods in experimental criminology research: What, why and how. *Journal of Experimental Criminology, 2*, 1-22.

Au, D., Crossley, T., & Schellhorn, M. (2005). The effects of health shocks and long-term health on the work of older Canadians. *Health Economics, 14*, 999-1018.

Becker, G. S. (1964). Human Capital. *Columbia University Press*.

Bloom, D., Canning, D., & Sevilla, J. (2004). The effect of health on economic growth: A production function approach. *World Development, 32*.

Borusyak, K., & Jaravel, X. (2019). Revisiting event study designs, with an application to the estimation of the marginal propensity to consume. 2017.

Bound, J., Schoenbaum, M., Stinebrickner, T. R., & Waidmann, T. (1999). The dynamic effects of health on the labour force transitions of older workers. *Labour Economics, 6*, 179-202.

Bourke, J. A., Owen, H. E., Derrett, S., & Wyeth, E. H. (2023). Disrupted mana and systemic abdication: Mā0ri qualitative experiences accessing healthcare in the 12 years post-injury. *BMC health services research, 23*, 1-9.

Cai, L., & Kalb, G. (2006, Mar). Health status and labour force participation: evidence from Australia. *Health Econ, 15*(3), 241-261. https://doi.org/10.1002/hec.1053

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics, 225*(2), 200-230. https://doi.org/10.1016/j.jeconom.2020.12.001

Carter, K. N., Gunasekara, F. I., Blakely, T., & Richardson, K. (2013, Jun). Health shocks adversely impact participation in the labour force in a working age population: a longitudinal analysis. *Aust N Z J Public Health, 37*(3), 257-263. https://doi.org/10.1111/1753-6405.12068

Charlies, K. K. (2003). The longitudinal structure of earnings losses among work-limited disabled workers. *Journal of Human Resources, 3*, 618-646.

Chirikos, T. N. (1993). The relationship between health and labour market status. *Annual Review Public Health*(14), 293-312.

Crichton, S., Stillman, S., & Hyslop, D. (2011). Returning to work from injury: longitudinal evidence on employment and earnings. *ILR Review, 64*(4), 765-785.

Currie, J., & Madrian, B. C. (1999). Health, health insurance and the labor market. *Handbook of labor economics, 3*, 3309-3416.

Dano, A. M. (2005, Sep). Road injuries and long-run effects on income and employment. *Health Econ, 14*(9), 955-970. https://doi.org/10.1002/hec.1045

Deaton, A. (2002). Policy implications of the gradient of health and wealth. *Health affairs, 21*(2), 13-30.

Disney, R., Emmerson, C., & Wakefiled, M. (2006). Ill-health and retirement in britain: a panel data-based analysis. *Health Economics*(25), 621-649.

Dwyer, D. S., & Mitchell, O. S. (1999). Health problems as determinants of retirement: are self rated measures endogenous? *Journal Health Economics*(18), 173-193.

Fadlon, I., & Nielsen, T. H. (2021). Family Labor Supply Responses to Severe Health Shocks: Evidence from Danish Administrative Records. *American Economic Journal: Applied Economics, 13*(3), 1-30.

Fiebig, D. G., van Gool, K., Hall, J., & Mu, C. (2021, Dec). Health care use in response to health shocks: Does socio-economic status matter? *Health Econ, 30*(12), 3032-3050. https://doi.org/10.1002/hec.4427

Garcia Gomez, P., & Lopez Nicolas, A. (2006, Sep). Health shocks, employment and income in the Spanish labour market. *Health Econ, 15*(9), 997-1009. https://doi.org/10.1002/hec.1151

Garcia-Gomez, P. (2011, Jan). Institutions, health shocks and labour market outcomes across Europe. *J Health Econ, 30*(1), 200-213. https://doi.org/10.1016/j.jhealeco.2010.11.003

Goodman-Bacon, A. (2021). Defference-in-differences with variation in treatment timing. *Journal of Econometrics, 225*, 254-277.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal Political Economy*(80), 223-225.

Hall, S. G., Swamy, P. A. V. B., & Tavlas, G. S. (2012). Generalised conintegration: A new concept with an application to health expenditure and health outcomes. *Empirical Economics, 42*(2), 603-618.

Halla, M., & Zweimüller, M. (2013). The effect of health on earnings: Quasi-experimental evidence from commuting accidents. *Labour Economics, 24*, 23-38. https://doi.org/10.1016/j.labeco.2013.04.006

Heinesen, E., & Kolodziejczyk, C. (2013). Effects of breast and colorectal cancer on labour market outcomes-average effects and educational gradients. *Journal of Health Economics, 32*(6), 1028-1042.

Jeon, S. H. (2017). The long-term effects of cancer on employment and earnings. *Health Economics, 26*(5), 671-684.

Lenhart, A. (2019). The effects of health shocks on labor market outcomes: evidence from UK panel data. *The European Journal of Health Economics*(20), 83-98. https://doi.org/https://doi.org/10.1007/s10198-018-0985-z

Lewycka. S., Dasgupta. K., Plum. A., Clark. T., Hedges., M., & Pacheco. G. (2023). International Journal for Equity in Health. *22, 1*, 1-15.

Lindeboom, M., Llena-Nozal, A., & van der Klaauw, B. (2016). Health shocks, disability and work. *Labour Economics, 43*, 186-200. https://doi.org/10.1016/j.labeco.2016.06.010

Malani, A., & Reif, J. (2015). Interpreting pre-trends as anticipation: Impact on estimated treatment effects from tort reform. *Journal of Public Economics, 2015*, 1-17.

McAllister, S., Derrett, S., Audas, R., Herbison, P., & Paul, C. (2013, May). Do different types of financial support after illness or injury affect socio-economic outcomes? A natural experiment in New Zealand. *Soc Sci Med, 85*, 93-102. https://doi.org/10.1016/j.socscimed.2013.02.041

Ministry for Women - Manatū Wāhine. (2023). Labour market participation. https://women.govt.nz/women-and-work/labour-market-participation#:~:text=As%20at%20June%202023%20in%20Aotearoa%20New%20Zealand%2C,to%2067.7%25%2C%20the%20highest%20participation%20rate%20since%201986.

Ministry of Business Innovation and Employment. (2021). Maori labour market trends. https://www.mbie.govt.nz/business-and-employment/employment-and-skills/labour-market-reports-data-and-analysis/other-labour-market-reports/maori-labour-market-trends/

Ministry of Health. (n.d). *Chronology of the New Zealand Health System 1840 - 2017*. www.health.govt.nz/system/files/documents/pages/chronology-of-the-new-zealand-health-system-1840-to-2017_0.pdf

Ministry of Health. (n.d.). *Te Tiriti o Waitangi*. https://www.health.govt.nz/our-work/populations/maori-health/te-tiriti-o-waitangi

Ministry of Social Development. (2010).

Mommaerts, C., Raza, S. H., & Zheng, Y. (2020). The economic consequences of hospitalisations for older workers across countries. *The Journal of the Economics of Ageing, 16*.

Nichols, A., Mitchell, J. & Lindner, S. . (2013). Consequences of Long-Term Unemployment. *Washington, Dc: The Urban Institute*.

Parro, F., & Pohl, R. V. (2021, May). The effect of accidents on labor market outcomes: Evidence from Chile. *Health Econ, 30*(5), 1015-1032. https://doi.org/10.1002/hec.4230

Raghupathi, V., & Raghupathi, W. (2020). Healthcare expenditure and economic performance: Insights from the United States data. *Frontiers in Public Health*(8).

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*.

Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*(5), 688-701.

Smith, J. P. (1999). Healthy bodies and thick wallets: the dual relation between health and economic status. *Journal of Economic Perspectives, 2*, 145-166.

Statistics New Zealand. (2018a). IDI IR Restrict data dictionary. In www.stats.govt.nz

Statistics New Zealand. (2018b). IDI MOH Public Hospital Discharges data dictionary. www.stats.govt.nz

Statistics New Zealand. (2020). *Ethnic group summaries reveal New Zealand's multi-cultural make-up*. Retrieved 10 November from https://www.stats.govt.nz/news/ethnic-group-summaries-reveal-new-zealands-multicultural-make-up

Statistics New Zealand. (n.d.a). IDI ACC Injury Claims Data Dictionary. In www.stats.govt.nz

Statistics New Zealand. (n.d.b). IDI Derived Population data dictionary. In www.stats.govt.nz

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics, 225*(2), 175-199. https://doi.org/10.1016/j.jeconom.2020.09.006

Theadom. A, Barker Collo. S, Jones. K, Kahan. M, Te Ao. B, McPherson. K, Starkey. N, & Feigin. V. (2017). Work Limitations 4 Years After mild Traumatic Brain Injury: A Cohort Study. *Archives of Physical Medicine and Rehabilitation, 98*(8), 1560 - 1566.

World Bank. (2022). *Indicator*. Retrieved 10/11 from https://data.worldbank.org/indicator/SP.POP.TOTL.FE.ZS?locations=NZ

Wu, S. (2003). The effects of health events on the economic status of married couples. *Journal of Human Resources, 1*, 219-230.

Appendices

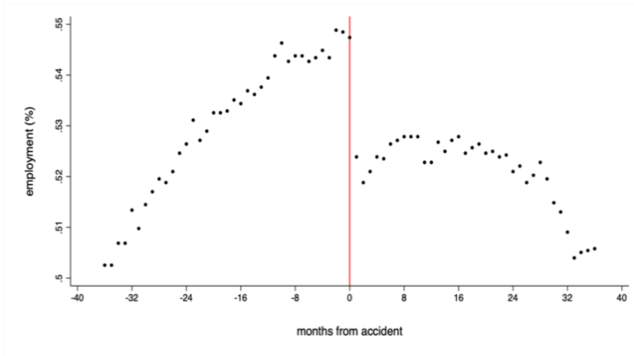
Appendix A: Descriptive Statistics

Table 16: Summary statistics of the population of interest and estimation samples

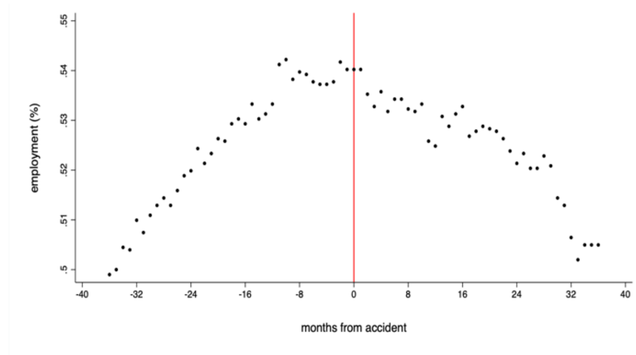
	Total Population of Interest (2) N=8292	Mild (cost) (3) N=6042	Medium (cost) (4) N=1827	Severe (cost) (5) N=423	Hospital mild (6) N=7800	Hospital medium (7) N=498	Hospital severe (8) N=234
Age, yrs mean	39	38	39	40	39	38	38
Female %	69	70	68	59	69	57	55
Maori %	20	20	21	27	19	37	33
Pasifika %	7	7	6	6	7	7	6
Other %	19	19	21	13	20	10	12
European %	54	54	52	54	54	46	49
Employed % 12 months before	65	63	63	78	64	68	71

Notes: Age is the mean age in years. All other variables are a percentage. Employed is the percentage of individuals who had positive income in the 12 months before the health shock. Column (2) gives statistics for the full population of interest. Column (3) mild (cost) health shock includes individuals with a total cost of less than $700. Column (4) medium (cost) health shock those who have a total cost of >= $700 or less than $9000. Column (5) severe (cost) health shock ≥$9000. Column (6) includes individuals who were not admitted to hospital or had length of stay of 0 or 1 night in hospital. Column (7) those individuals who had more than 1 night length of stay and column (8), severe more than 5 nights length of stay. Due to random rounding, not all numbers add up. The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018).
Source: Own calculations using Statistics NZ's IDI.

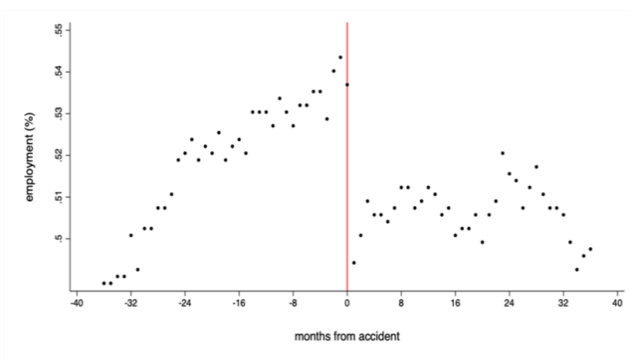
Figure 5: Monthly employment in event time (%)



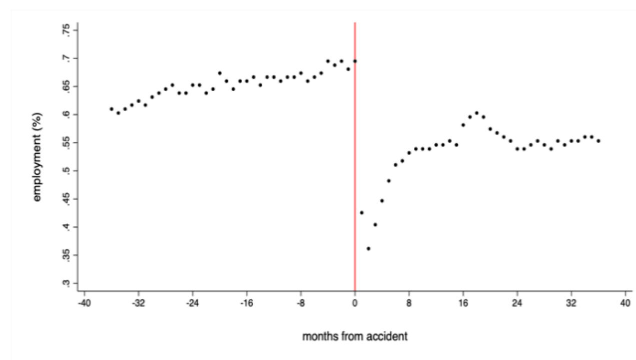
Full population of interest



Mild health shock sample (cost)



Medium health shock sample (cost)



Severe health shock sample

Notes: Figure 5 depicts the percentage of the population of interest who were employed each month. The monthly employment for the population of interest from 36 months before the health shock to 36 months after the health shock. The thirty-six months before the accident are represented by negative values running from -36 to -1 and the months after the accident are denoted by positive values 1 to 36. The red vertical line indicates when the health shock or the treatment occurs, at event time zero. The y axis shows the percentage of the population of interest employed in a given month and the x axis the event time. The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Mild health shock includes individuals with a total cost of less than $700. Medium health shock those who have a total cost of >= $700 or less than $9000. Severe health shock ≥ $9000.
Source: Own calculations using Statistics NZ's IDI

Table 17:Monthly employment in event time for the full population of interest

Event time (1)	Not Employed (2)	Employed (3)	Total (4)	Percentage Employed (5)
-36	4,125	4,167	8,292	0.502533
-35	4,125	4,167	8,292	0.502533
-34	4,089	4,203	8,292	0.506874
-33	4,089	4,203	8,292	0.506874
-32	4,035	4,257	8,292	0.513386
-31	4,065	4,227	8,292	0.509768
-30	4,026	4,266	8,292	0.514472
-29	4,005	4,287	8,292	0.517004
-28	3,984	4,308	8,292	0.519537
-27	3,990	4,302	8,292	0.518813
-26	3,972	4,320	8,292	0.520984
-25	3,942	4,350	8,292	0.524602
-24	3,927	4,365	8,292	0.526411
-23	3,888	4,404	8,292	0.531114
-22	3,921	4,371	8,292	0.527135
-21	3,906	4,386	8,292	0.528944
-20	3,876	4,416	8,292	0.532562
-19	3,876	4,416	8,292	0.532562
-18	3,873	4,419	8,292	0.532923
-17	3,855	4,437	8,292	0.535094
-16	3,861	4,431	8,292	0.53437
-15	3,840	4,452	8,292	0.536903
-14	3,846	4,446	8,292	0.536179
-13	3,834	4,458	8,292	0.537627
-12	3,819	4,473	8,292	0.539436
-11	3,783	4,509	8,292	0.543777

-10	3,762	4,530	8,292	0.54631
-9	3,792	4,500	8,292	0.542692
-8	3,783	4,509	8,292	0.543777
-7	3,783	4,509	8,292	0.543777
-6	3,792	4,500	8,292	0.542692
-5	3,786	4,506	8,292	0.543415
-4	3,774	4,518	8,292	0.544863
-3	3,786	4,506	8,292	0.543415
-2	3,741	4,551	8,292	0.548842
-1	3,744	4,548	8,292	0.54848
0	3,753	4,539	8,292	0.547395
1	3,948	4,344	8,292	0.523878
2	3,990	4,302	8,292	0.518813
3	3,972	4,320	8,292	0.520984
4	3,948	4,344	8,292	0.523878
5	3,951	4,341	8,292	0.523517
6	3,927	4,365	8,292	0.526411
7	3,921	4,371	8,292	0.527135
8	3,915	4,377	8,292	0.527858
9	3,915	4,377	8,292	0.527858
10	3,915	4,377	8,292	0.527858
11	3,957	4,335	8,292	0.522793
12	3,957	4,335	8,292	0.522793
13	3,924	4,368	8,292	0.526773
14	3,939	4,353	8,292	0.524964
15	3,921	4,371	8,292	0.527135
16	3,915	4,377	8,292	0.527858
17	3,942	4,350	8,292	0.524602
18	3,933	4,359	8,292	0.525687

19	3,927	4,365	8,292	0.526411
20	3,942	4,350	8,292	0.524602
21	3,939	4,353	8,292	0.524964
22	3,948	4,344	8,292	0.523878
23	3,945	4,347	8,292	0.52424
24	3,972	4,320	8,292	0.520984
25	3,963	4,329	8,292	0.522069
26	3,990	4,302	8,292	0.518813
27	3,978	4,314	8,292	0.52026
28	3,957	4,335	8,292	0.522793
29	3,984	4,308	8,292	0.519537
30	4,023	4,269	8,292	0.514834
31	4,038	4,254	8,292	0.513025
32	4,071	4,221	8,292	0.509045
33	4,113	4,179	8,292	0.50398
34	4,104	4,188	8,292	0.505065
35	4,101	4,191	8,292	0.505427
36	4,098	4,194	8,292	0.505789

Notes: Table A 2 presents the number of employed and not employed each month for the full population. Column 5 presents the percentage of the population of interest who were employed each month. The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). The monthly employment for the population of interest from 36 months before the health shock to 36 months after the health shock. The thirty-six months before the accident are represented by negative values running from -36 to -1 and the months after the accident are denoted by positive values 1 to 36.
Source: Own calculations using Statistics NZ's IDI

Table 18: Monthly employment in event time for the mild health shock group

Event time (1)	Not employed (2)	Employed (3)	Total (4)	Percentage Employed (5)
-36	3,027	3,015	6,042	0.499007
-35	3,021	3,021	6,042	0.5
-34	2,994	3,048	6,042	0.504469
-33	2,997	3,045	6,042	0.503972
-32	2,961	3,081	6,042	0.50993
-31	2,976	3,066	6,042	0.507448
-30	2,955	3,087	6,042	0.510924
-29	2,943	3,099	6,042	0.51291
-28	2,934	3,108	6,042	0.514399
-27	2,943	3,099	6,042	0.51291
-26	2,925	3,117	6,042	0.515889
-25	2,907	3,135	6,042	0.518868
-24	2,901	3,141	6,042	0.519861
-23	2,874	3,168	6,042	0.52433
-22	2,892	3,150	6,042	0.521351
-21	2,880	3,162	6,042	0.523337
-20	2,862	3,180	6,042	0.526316
-19	2,865	3,177	6,042	0.525819
-18	2,844	3,198	6,042	0.529295
-17	2,838	3,204	6,042	0.530288
-16	2,844	3,198	6,042	0.529295
-15	2,820	3,222	6,042	0.533267
-14	2,838	3,204	6,042	0.530288
-13	2,832	3,210	6,042	0.531281
-12	2,820	3,222	6,042	0.533267
-11	2,772	3,270	6,042	0.541212

-10	2,766	3,276	6,042	0.542205
-9	2,790	3,252	6,042	0.538232
-8	2,781	3,261	6,042	0.539722
-7	2,784	3,258	6,042	0.539225
-6	2,793	3,249	6,042	0.537736
-5	2,796	3,246	6,042	0.537239
-4	2,796	3,246	6,042	0.537239
-3	2,793	3,249	6,042	0.537736
-2	2,769	3,273	6,042	0.541708
-1	2,778	3,264	6,042	0.540218
0	2,778	3,264	6,042	0.540218
1	2,778	3,264	6,042	0.540218
2	2,808	3,234	6,042	0.535253
3	2,823	3,219	6,042	0.532771
4	2,805	3,237	6,042	0.53575
5	2,829	3,213	6,042	0.531778
6	2,814	3,228	6,042	0.53426
7	2,814	3,228	6,042	0.53426
8	2,826	3,216	6,042	0.532274
9	2,829	3,213	6,042	0.531778
10	2,820	3,222	6,042	0.533267
11	2,865	3,177	6,042	0.525819
12	2,871	3,171	6,042	0.524826
13	2,835	3,207	6,042	0.530785
14	2,847	3,195	6,042	0.528798
15	2,832	3,210	6,042	0.531281
16	2,823	3,219	6,042	0.532771
17	2,859	3,183	6,042	0.526812
18	2,853	3,189	6,042	0.527805

19	2,847	3,195	6,042	0.528798
20	2,850	3,192	6,042	0.528302
21	2,853	3,189	6,042	0.527805
22	2,862	3,180	6,042	0.526316
23	2,877	3,165	6,042	0.523833
24	2,892	3,150	6,042	0.521351
25	2,880	3,162	6,042	0.523337
26	2,898	3,144	6,042	0.520357
27	2,898	3,144	6,042	0.520357
28	2,883	3,159	6,042	0.52284
29	2,895	3,147	6,042	0.520854
30	2,934	3,108	6,042	0.514399
31	2,943	3,099	6,042	0.51291
32	2,982	3,060	6,042	0.506455
33	3,009	3,033	6,042	0.501986
34	2,991	3,051	6,042	0.504965
35	2,991	3,051	6,042	0.504965
36	2,991	3,051	6,042	0.504965

Notes: Table A 2.1 presents the number of employed and not employed each month for individuals who experience a mild health shock. Column 5 presents the percentage of the population of interest who were employed each month. The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). The monthly employment for the population of interest from 36 months before the health shock to 36 months after the health shock. The thirty-six months before the accident are represented by negative values running from -36 to -1 and the months after the accident are denoted by positive values 1 to 36. A mild health shock includes individuals with a total cost of less than $700. Due to random rounding not all totals add up.
Source: Own calculations using Statistics NZ's IDI

Table 19: Monthly employment in event time for the medium health shock group

Event time	Not employed	Employed	Total	Percent
-36	933	894	1827	0.489327
-35	933	894	1827	0.489327
-34	930	897	1827	0.490969
-33	930	897	1827	0.490969
-32	912	915	1827	0.500821
-31	927	900	1827	0.492611
-30	909	918	1827	0.502463
-29	909	918	1827	0.502463
-28	900	927	1827	0.507389
-27	900	927	1827	0.507389
-26	894	933	1827	0.510673
-25	879	948	1827	0.518883
-24	876	951	1827	0.520525
-23	870	957	1827	0.52381
-22	879	948	1827	0.518883
-21	873	954	1827	0.522167
-20	876	951	1827	0.520525
-19	867	960	1827	0.525452
-18	879	948	1827	0.518883
-17	873	954	1827	0.522167
-16	870	957	1827	0.52381
-15	876	951	1827	0.520525
-14	858	969	1827	0.530378
-13	858	969	1827	0.530378
-12	858	969	1827	0.530378
-11	864	963	1827	0.527094
-10	852	975	1827	0.533662

-9	858	969	1827	0.530378
-8	864	963	1827	0.527094
-7	855	972	1827	0.53202
-6	855	972	1827	0.53202
-5	849	978	1827	0.535304
-4	849	978	1827	0.535304
-3	861	966	1827	0.528736
-2	840	987	1827	0.54023
-1	834	993	1827	0.543514
0	846	981	1827	0.536946
1	924	903	1827	0.494253
2	912	915	1827	0.500821
3	897	930	1827	0.509031
4	903	924	1827	0.505747
5	903	924	1827	0.505747
6	906	921	1827	0.504105
7	900	927	1827	0.507389
8	891	936	1827	0.512315
9	891	936	1827	0.512315
10	900	927	1827	0.507389
11	897	930	1827	0.509031
12	891	936	1827	0.512315
13	894	933	1827	0.510673
14	903	924	1827	0.505747
15	900	927	1827	0.507389
16	912	915	1827	0.500821
17	909	918	1827	0.502463
18	909	918	1827	0.502463
19	903	924	1827	0.505747

20	915	912	1827	0.499179
21	903	924	1827	0.505747
22	897	930	1827	0.509031
23	876	951	1827	0.520525
24	885	942	1827	0.515599
25	888	939	1827	0.513957
26	900	927	1827	0.507389
27	891	936	1827	0.512315
28	882	945	1827	0.517241
29	894	933	1827	0.510673
30	900	927	1827	0.507389
31	900	927	1827	0.507389
32	903	924	1827	0.505747
33	915	912	1827	0.499179
34	927	900	1827	0.492611
35	921	906	1827	0.495895
36	918	909	1827	0.497537

Notes: Table A 2.2 presents the number of employed and not employed each month for individuals who experience a medium health shock. Column 5 presents the percentage of the population of interest who were employed each month. The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018).The monthly employment for the population of interest from 36 months before the health shock to 36 months after the health shock. The thirty-six months before the accident are represented by negative values running from -36 to -1 and the months after the accident are denoted by positive values 1 to 36. A medium health shock those who have a total cost of >= $700 or less than $9000. Due to random rounding not, all totals add up.
Source: Own calculations using Statistics NZ's IDI

Table 20: Monthly employment in event time for the severe health shock group

Event time (1)	Not Employed (2)	Employed (3)	Total (4)	Percentage Employed (5)
-36	165	258	423	0.609929
-35	168	255	423	0.602837
-34	165	258	423	0.609929
-33	162	261	423	0.617021
-32	159	264	423	0.624113
-31	162	261	423	0.617021
-30	156	267	423	0.631206
-29	153	270	423	0.638298
-28	150	273	423	0.64539
-27	147	276	423	0.652482
-26	153	270	423	0.638298
-25	153	270	423	0.638298
-24	147	276	423	0.652482
-23	147	276	423	0.652482
-22	153	270	423	0.638298
-21	150	273	423	0.64539
-20	138	285	423	0.673759
-19	144	279	423	0.659574
-18	150	273	423	0.64539
-17	144	279	423	0.659574
-16	144	279	423	0.659574
-15	141	282	423	0.666667
-14	147	276	423	0.652482
-13	141	282	423	0.666667
-12	141	282	423	0.666667
-11	144	279	423	0.659574

-10	141	282	423	0.666667
-9	141	282	423	0.666667
-8	138	285	423	0.673759
-7	144	279	423	0.659574
-6	141	282	423	0.666667
-5	138	285	423	0.673759
-4	129	294	423	0.695035
-3	132	291	423	0.687943
-2	129	294	423	0.695035
-1	135	288	423	0.680851
0	129	294	423	0.695035
1	243	180	423	0.425532
2	270	153	423	0.361702
3	252	171	423	0.404255
4	234	189	423	0.446809
5	219	204	423	0.48227
6	207	216	423	0.510638
7	204	219	423	0.51773
8	198	225	423	0.531915
9	195	228	423	0.539007
10	195	228	423	0.539007
11	195	228	423	0.539007
12	192	231	423	0.546099
13	192	231	423	0.546099
14	189	234	423	0.553191
15	192	231	423	0.546099
16	177	246	423	0.58156
17	171	252	423	0.595745
18	168	255	423	0.602837

19	171	252	423	0.595745
20	180	243	423	0.574468
21	183	240	423	0.567376
22	186	237	423	0.560284
23	189	234	423	0.553191
24	195	228	423	0.539007
25	195	228	423	0.539007
26	192	231	423	0.546099
27	189	234	423	0.553191
28	192	231	423	0.546099
29	195	228	423	0.539007
30	189	234	423	0.553191
31	192	231	423	0.546099
32	189	234	423	0.553191
33	189	234	423	0.553191
34	186	237	423	0.560284
35	186	237	423	0.560284
36	189	234	423	0.553191

Notes: Table A 2.3 presents the number of employed and not employed each month for individuals who experience a severe health shock. Column 5 presents the percentage of the population of interest who were employed each month. The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018).The monthly employment for the population of interest from 36 months before the health shock to 36 months after the health shock. The thirty-six months before the accident are represented by negative values running from -36 to -1 and the months after the accident are denoted by positive values 1 to 36. A severe health shock total cost ≥ $9000. Due to random rounding not, all totals add up.
Source: Own calculations using Statistics NZ's IDI

Appendix B: Results: treatment effect estimations using the CS-DID for employment and log-income.

Table 21:Full population: average treatment effects on employment

Event time	Coefficient	std error	Z value	P value	95% confidence interval	
Pre_average (34 months)	-0.00028	0.000282	-0.98	0.329	-0.00083	0.000277
Post_average (34 months)	-0.0492	0.00505	-9.74	0	-0.0591	-0.03931
tm34	0.004003	0.003752	1.07	0.286	-0.00335	0.011357
tm33	-0.00189	0.002978	-0.64	0.525	-0.00773	0.003946
tm32	0.00496	0.002748	1.8	0.071	-0.00043	0.010346
tm31	-0.00622	0.002693	-2.31	0.021	-0.0115	-0.00094
tm30	0.003982	0.002643	1.51	0.132	-0.0012	0.009161
tm29	0.000757	0.002554	0.3	0.767	-0.00425	0.005762
tm28	7.75E-05	0.002442	0.03	0.975	-0.00471	0.004863
tm27	-0.00243	0.002431	-1	0.317	-0.0072	0.00233
tm26	0.000289	0.002384	0.12	0.904	-0.00438	0.004962
tm25	0.001855	0.002474	0.75	0.453	-0.00299	0.006704
tm24	-7.1E-05	0.002511	-0.03	0.978	-0.00499	0.004851
tm23	0.002456	0.002463	1	0.319	-0.00237	0.007284
tm22	-0.00611	0.002333	-2.62	0.009	-0.01069	-0.00154
tm21	0.000738	0.002408	0.31	0.759	-0.00398	0.005457
tm20	0.001484	0.002392	0.62	0.535	-0.0032	0.006173
tm19	-0.00155	0.00242	-0.64	0.522	-0.00629	0.003195
tm18	-0.00093	0.002386	-0.39	0.695	-0.00561	0.003741
tm17	0.000164	0.002327	0.07	0.944	-0.0044	0.004724
tm16	-0.00255	0.002427	-1.05	0.293	-0.00731	0.002202
tm15	0.000608	0.002304	0.26	0.792	-0.00391	0.005122
tm14	-0.00214	0.00239	-0.89	0.371	-0.00682	0.002547
tm13	7.08E-05	0.002329	0.03	0.976	-0.00449	0.004635
tm12	0.000279	0.002319	0.12	0.904	-0.00427	0.004823
tm11	0.00217	0.002302	0.94	0.346	-0.00234	0.006682
tm10	0.000671	0.002414	0.28	0.781	-0.00406	0.005403
tm9	-0.00458	0.002367	-1.93	0.053	-0.00922	6.02E-05
tm8	-0.00058	0.002368	-0.25	0.805	-0.00523	0.004058
tm7	-0.00097	0.002308	-0.42	0.675	-0.00549	0.003554
tm6	-0.00218	0.002255	-0.97	0.334	-0.0066	0.002242
tm5	-0.00126	0.002326	-0.54	0.588	-0.00582	0.003297
tm4	0.000361	0.002249	0.16	0.872	-0.00405	0.004769

tm3	-0.00312	0.002192	-1.42	0.154	-0.00742	0.001174
tm2	0.004532	0.002338	1.94	0.053	-5E-05	0.009115
tm1	-0.00221	0.002224	-0.99	0.321	-0.00657	0.002151
tp0	-0.00174	0.002217	-0.78	0.434	-0.00608	0.002609
tp1	-0.02651	0.003118	-8.5	0	-0.03262	-0.0204
tp2	-0.0332	0.003544	-9.37	0	-0.04015	-0.02626
tp3	-0.03262	0.003715	-8.78	0	-0.0399	-0.02534
tp4	-0.03084	0.003878	-7.95	0	-0.03844	-0.02324
tp5	-0.03298	0.004064	-8.12	0	-0.04095	-0.02502
tp6	-0.03036	0.004279	-7.1	0	-0.03875	-0.02198
tp7	-0.0302	0.004416	-6.84	0	-0.03886	-0.02155
tp8	-0.03089	0.004484	-6.89	0	-0.03968	-0.0221
tp9	-0.03248	0.004591	-7.08	0	-0.04148	-0.02348
tp10	-0.03361	0.004761	-7.06	0	-0.04294	-0.02428
tp11	-0.04019	0.004877	-8.24	0	-0.04974	-0.03063
tp12	-0.04247	0.005091	-8.34	0	-0.05245	-0.03249
tp13	-0.03954	0.005289	-7.48	0	-0.0499	-0.02917
tp14	-0.04281	0.005525	-7.75	0	-0.05364	-0.03198
tp15	-0.04114	0.005715	-7.2	0	-0.05234	-0.02994
tp16	-0.04323	0.005889	-7.34	0	-0.05477	-0.03169
tp17	-0.04806	0.006088	-7.89	0	-0.06	-0.03613
tp18	-0.0503	0.006219	-8.09	0	-0.06249	-0.03811
tp19	-0.05001	0.006433	-7.77	0	-0.06262	-0.0374
tp20	-0.0522	0.006527	-8	0	-0.06499	-0.0394
tp21	-0.05255	0.006651	-7.9	0	-0.06559	-0.03952
tp22	-0.05554	0.00682	-8.14	0	-0.0689	-0.04217
tp23	-0.05702	0.006964	-8.19	0	-0.07067	-0.04337
tp24	-0.06155	0.007199	-8.55	0	-0.07566	-0.04744
tp25	-0.06191	0.007412	-8.35	0	-0.07644	-0.04738
tp26	-0.06541	0.007527	-8.69	0	-0.08016	-0.05065
tp27	-0.06398	0.00776	-8.24	0	-0.07919	-0.04877
tp28	-0.0628	0.007966	-7.88	0	-0.07841	-0.04719
tp29	-0.06736	0.008134	-8.28	0	-0.0833	-0.05141
tp30	-0.0719	0.00829	-8.67	0	-0.08815	-0.05565
tp31	-0.07676	0.008429	-9.11	0	-0.09328	-0.06023
tp32	-0.08028	0.008634	-9.3	0	-0.0972	-0.06336
tp33	-0.08873	0.00889	-9.98	0	-0.10615	-0.0713
tp34	-0.09097	0.009174	-9.92	0	-0.10895	-0.07299

Notes: The above estimates for the full population of interest are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months

after the accident month (from 2002 to 2018). Column (1) records the event time. Tm denotes time periods before the treatment and Tp time periods post treatment in event time. Row 1, column (2) the combined average treatment effect for employment for the 34 months before the treatment month. Row 2, column (2) the combined average treatment effect for employment for the 34 months after the treatment month. Column (2) contains the coefficients for the estimations. While my panel is balanced in event time it isn't balanced in calendar time. As the estimator always needs a g-1 and/or a valid comparison period the first 2 periods and the last 2 periods from my panel are not estimated. Consequently, this leaves me with estimations for 34 periods before and 34 periods after the treatment. *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

Table 22: All passengers: estimations of the treatment effect of a health shock on log-income

Event time	Coefficient	std error	Z value	P value	95% confidence interval	
Pre_average	-0.00233	0.002086	-1.12	0.265	-0.00641	0.001762
Post_average	-0.40671	0.039409	-10.32	0	-0.48395	-0.32947
tm34	0.032609	0.026411	1.23	0.217	-0.01916	0.084375
tm33	-0.02198	0.02126	-1.03	0.301	-0.06364	0.019692
tm32	0.04089	0.019488	2.1	0.036	0.002694	0.079086
tm31	-0.0425	0.019141	-2.22	0.026	-0.08001	-0.00498
tm30	0.030968	0.018754	1.65	0.099	-0.00579	0.067726
tm29	0.011294	0.018196	0.62	0.535	-0.02437	0.046956
tm28	0.000836	0.017437	0.05	0.962	-0.03334	0.035011
tm27	-0.01322	0.017517	-0.75	0.45	-0.04755	0.021113
tm26	-0.00741	0.01723	-0.43	0.667	-0.04117	0.026364
tm25	0.011415	0.017584	0.65	0.516	-0.02305	0.045879
tm24	-0.00435	0.01793	-0.24	0.808	-0.03949	0.030791
tm23	0.016062	0.017458	0.92	0.358	-0.01816	0.05028
tm22	-0.04527	0.016667	-2.72	0.007	-0.07793	-0.0126
tm21	0.005432	0.017002	0.32	0.749	-0.02789	0.038755
tm20	0.015287	0.017403	0.88	0.38	-0.01882	0.049396
tm19	-0.01452	0.017204	-0.84	0.399	-0.04824	0.019195
tm18	-0.00447	0.017059	-0.26	0.793	-0.03791	0.028963
tm17	-0.00518	0.016608	-0.31	0.755	-0.03774	0.027367
tm16	-0.01422	0.017381	-0.82	0.413	-0.04828	0.019847
tm15	0.003082	0.016777	0.18	0.854	-0.0298	0.035964
tm14	-0.01404	0.017347	-0.81	0.418	-0.04804	0.019955
tm13	0.005157	0.016632	0.31	0.757	-0.02744	0.037755
tm12	-0.00503	0.01647	-0.31	0.76	-0.03731	0.027251
tm11	0.015303	0.016551	0.92	0.355	-0.01714	0.047742
tm10	0.003367	0.017452	0.19	0.847	-0.03084	0.037571
tm9	-0.03505	0.01704	-2.06	0.04	-0.06845	-0.00165
tm8	-0.00916	0.017034	-0.54	0.591	-0.04255	0.024225
tm7	-0.00516	0.016491	-0.31	0.754	-0.03748	0.02716
tm6	-0.01651	0.016344	-1.01	0.312	-0.04855	0.01552

tm5	-0.00265	0.016955	-0.16	0.876	-0.03588	0.030577
tm4	0.001843	0.016176	0.11	0.909	-0.02986	0.033546
tm3	-0.02441	0.0161	-1.52	0.129	-0.05597	0.007144
tm2	0.026534	0.016597	1.6	0.11	-0.006	0.059064
tm1	-0.01402	0.015679	-0.89	0.371	-0.04475	0.016711
tp0	-0.03139	0.015996	-1.96	0.05	-0.06274	-3.5E-05
tp1	-0.25484	0.023262	-10.96	0	-0.30044	-0.20925
tp2	-0.28197	0.026674	-10.57	0	-0.33425	-0.22969
tp3	-0.27465	0.028204	-9.74	0	-0.32992	-0.21937
tp4	-0.25738	0.029469	-8.73	0	-0.31514	-0.19962
tp5	-0.27501	0.031044	-8.86	0	-0.33586	-0.21417
tp6	-0.26678	0.03268	-8.16	0	-0.33083	-0.20273
tp7	-0.25969	0.033786	-7.69	0	-0.32591	-0.19347
tp8	-0.25844	0.034396	-7.51	0	-0.32585	-0.19102
tp9	-0.2676	0.035293	-7.58	0	-0.33677	-0.19843
tp10	-0.27524	0.036542	-7.53	0	-0.34686	-0.20362
tp11	-0.32112	0.03769	-8.52	0	-0.39499	-0.24725
tp12	-0.34272	0.039487	-8.68	0	-0.42011	-0.26532
tp13	-0.31905	0.041072	-7.77	0	-0.39955	-0.23856
tp14	-0.3426	0.042942	-7.98	0	-0.42677	-0.25844
tp15	-0.3395	0.044324	-7.66	0	-0.42638	-0.25263
tp16	-0.35559	0.045833	-7.76	0	-0.44542	-0.26576
tp17	-0.38536	0.04742	-8.13	0	-0.4783	-0.29242
tp18	-0.40981	0.048531	-8.44	0	-0.50493	-0.31469
tp19	-0.41161	0.05021	-8.2	0	-0.51002	-0.3132
tp20	-0.42437	0.051024	-8.32	0	-0.52437	-0.32436
tp21	-0.42974	0.051991	-8.27	0	-0.53164	-0.32784
tp22	-0.45069	0.053293	-8.46	0	-0.55514	-0.34624
tp23	-0.46983	0.054445	-8.63	0	-0.57654	-0.36312
tp24	-0.50102	0.056202	-8.91	0	-0.61118	-0.39087
tp25	-0.51011	0.057855	-8.82	0	-0.6235	-0.39671
tp26	-0.53579	0.058887	-9.1	0	-0.65121	-0.42038
tp27	-0.53004	0.060542	-8.75	0	-0.6487	-0.41138
tp28	-0.52257	0.062166	-8.41	0	-0.64441	-0.40073
tp29	-0.56685	0.063602	-8.91	0	-0.69151	-0.4422
tp30	-0.6026	0.064825	-9.3	0	-0.72966	-0.47555
tp31	-0.63828	0.065802	-9.7	0	-0.76725	-0.50931
tp32	-0.66169	0.06747	-9.81	0	-0.79393	-0.52945
tp33	-0.72216	0.069364	-10.41	0	-0.85812	-0.58621

tp34	-0.73878	0.071266	-10.37	0	-0.87846	-0.5991

Notes: The above estimates for the full population of interest are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month (from 2002 to 2018). Column (1) records the event time. Tm denotes time periods before the treatment and Tp time periods post treatment in event time. Row 1, column (2) the combined average treatment effect on log-income for the 34 months before the treatment month. Row 2, column (2) the combined average treatment effect for employment for the 34 months after the treatment month. Column (2) contains the coefficients for the estimations. While my panel is balanced in event time it isn't balanced in calendar time. As the estimator always needs a g-1 and/or a valid comparison period the first 2 periods and the last 2 periods from my panel are not estimated. Consequently, this leaves me with estimations for 34 periods before and 34 periods after the treatment. *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations using Statistics NZ's IDI

Appendix C: Heterogeneity

Table 23: Estimations of the effect of a health shock on employment by gender and ethnicity

	Full Population (1)	Mild Shock (2)	Medium Shock (3)	Severe Shock (4)
Full population	-0.049*** (0.005)	-0.037*** (0.006)	-0.061*** (0.011)	-0.179*** (0.024)
	8292	6042	1827	423
Female	-0.05*** (0.006)	-0.04*** (0.007)	-0.056*** (0.013)	-0.208*** (0.034)
	5691	4206	1236	249
Male	-0.046*** (0.009)	-0.03*** (0.01)	-0.067*** (0.02)	-0.163*** (0.035)
	2601	1836	591	174
European	-0.035*** (0.007)	-0.029*** (0.008)	-0.029** (0.015)	-0.153*** (0.035)
	4449	3267	957	228
Non-European	-0.066*** (0.007)	-0.047*** (0.009)	-0.097*** (0.016)	-0.23*** (0.034)
	3843	2775	870	195

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month. These estimates are the average treatment effect for the 34 months after the health shock. Non-European includes Maori, Pasifika, Asian, Middle Eastern, Latin American and other. *** p<0.01, ** p<0.05, * p<0.1.

Table 24: Estimations of the effect of a health shock on log-income by gender and ethnicity.

	Full Population (1)	Mild Shock (2)	Medium Shock (3)	Severe Shock (4)
Full population	-0.407*** (0.039)	-0.3061*** (0.046)	-0.487*** (0.084)	-1.552*** (0.189)
	8292	6042	1827	423
Female	-0.413*** (0.047)	-0.329*** (0.054)	-0.459*** (0.098)	-1.318*** (0.257)
	5691	4206	1236	249
Male	-0.384*** (0.071)	-0.248*** (0.083)	-0.528*** (0.159)	-0.163*** (0.035)
	2601	1836	591	174
European	-0.289*** (0.054)	-0.241*** (0.062)	-0.243** (0.114)	-1.233*** (0.253)
	4449	3267	957	228
Non-European	-0.544*** (0.058)	-0.383*** (0.067)	-0.772*** (0.123)	-1.921*** (0.266)
	3843	2775	870	195

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month. Non-European includes Maori, Pasifika, Asian, Middle Eastern, Latin American and other. These estimates are the average treatment effect for the 34 months after the health shock *** p<0.01, ** p<0.05, * p<0.1.

Table 25: Checking if there are significant differences in treatment effects for employment between genders and ethnicities

	Post-treatment mean: all passengers	Post-treatment mean: mild health shock	Post-treatment mean: medium health shock	Post-treatment mean: severe health shock
Male	-0.046***(0.009)	-0.03***(0.01)	-0.067***(0.02)	-0.163***(0.035)
Female	-0.05*** (0.006)	-0.04***(0.007)	-0.056***(0.013)	-0.208***(0.034)
Difference	-0.004 (0.011)	.01 (0.012)	.011 (0.023)	.045 (0.050)
European	-0.035***(0.007)	-0.029***(0.008)	-0.029*(0.015)	-0.153***(0.035)sig pre
non-European	-0.066***(0.007)	-0.047***(0.009)	-0.097***(0.016)	-0.23***(0.034)
Difference	.031***(0.01)	.018*(0.012)	.068***(0.022)	.077*(0.049)

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month. The difference between genders and ethnicities is calculated for gender in row 4 and ethnicity in row 7. I then conduct a t-test to determine if the difference is significant. *** p<0.01, ** p<0.05, * p<0.1.

Table 26: Checking if there are significant differences in treatment effects for log-income between genders and ethnicities

	Post-treatment mean (all pass)	Post-treatment mean (cost one)	Post-treatment mean (cost two)	Post-treatment mean (cost three)
Male	-0.384*** (0.071)	-0.248*** (0.083)	-0.528** (0.159)	-1.318***(0.257)
Female	-0.413*** (0.047)	-0.329***(0.054)pre	-0.459*** (0.098)	-1.74*** (0.26)
difference	.029 (0.084)	.081 (0.098)	.069 (0.179)	.422 (0.378)
European	-0.289***(0.054)	-0.241***(0.062)	-0.243** (0.114)	-1.233*** (0.253)
Non-European	-0.544***(0.058)	-0.383***(0.067)	-0.772*** (0.123)	-1.921*** (0.266)
Difference	.255***(0.079)	.142* (0.091)	.529*** (0.167)	.688**(0.368)

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month. The difference between genders and ethnicities is calculated for gender in row 4 and ethnicity in row 7. I then conduct a t-test to determine if the difference is significant. *** p<0.01, ** p<0.05, * p<0.1.

Appendix D: Labour Market Attachment

Table 27: Estimations of the treatment effect on employment for the full population and for those who had positive employment in the 12 months before treatment

	Full Population (1)	Positive earnings in the 12 months before treatment (2)
Pre average (34 months)	-0.00028 (0.000)	0.000 (0.000)
Post average 34 months	-0.049*** (0.005)	-0.155*** (0.007)
Sample size	8292	5325

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month. Column (1) is the estimate for the full population, column (2) individuals who had positive wages and salary (employment) in at least one month in the 12 months before the health shock *** p<0.01, ** p<0.05, * p<0.1. Standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.
Source: own calculations use Statistics New Zealand's IDI

Table 28: Estimations of the treatment effect of log-income for the full population and for those who had positive employment in the 12 months before treatment

	Full Population (1)	Positive earnings 12 months before (2)
Pre-average (34 months)	-0.002 (0.002)	-0.003 (0.003)
Post-average 34 months	-0.407*** (0.039)	-1.22*** (0.052)
Sample size	8292	5325

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month. Column (1) is the estimate for the full population, column (2) individuals who had positive wages and salary (employment) in at least one month in the 12 months before the health shock *** p<0.01, ** p<0.05, * p<0.1. Standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.
Source: own calculations use Statistics New Zealand's IDI

Table 29: Estimations of the treatment effect on log-income for those who remain in employment.

	Full Population (1)	Positive earnings 12 months before (2)
Pre-average (34 months)	0.001 (0.001)	0.000 (0.001)
Post-average 34 months	-0.033*** (0.01)	-0.03*** (0.01)
Sample size	8292	5325

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month. These estimates are conditional on whether an individual remains employed. Column (1) is the estimate for the full population, column (2) individuals who had positive wages and salary (employment) in at least one month in the 12 months before the health shock *** p<0.01, ** p<0.05, * p<0.1. Standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.
Source: own calculations use Statistics New Zealand's IDI

Table 30: The effect of a health shock on benefit recipiency

	Full Population (1)	Mild Health Shock (2)	Medium Health Shock (3)	Severe Health Shock (4)
Pre-average (34 months)	0.000 (0.000)	0.000 (0.000)	0.000 (0.001)	0.002 (0.003)
Post-average 34 months	0.004 (0.004)	0.00 (0.004)	0.008 (0.008)	0.053*** (0.017)
Sample size	8292	6042	1827	423

Notes: The above estimates are obtained by estimating the staggered difference-in-differences methodology developed by Callaway & Sant'Anna (2021). The analysis uses ACC data from 2005 to 2015 and IR data for 36 months before the accident month and 36 months after the accident month. Column (1) is the estimate for the full population, column (2) those who had a total cost of less than $700 (NZD), column (3) individuals who had a total cost of ≥700 and <9000, column (4) individuals whose total cost was ≥9000 *** p<0.01, ** p<0.05, * p<0.1.