Facial Masculinity and Academic Research Performance

Parker Wan

A thesis submitted to

Auckland University of Technology

in partial fulfilment of the requirements for the degree of

Master of Business (MBus)

2021

School of Business

Abstract

Folk wisdom believes "Don't judge a book by its cover". However, many studies find that an individual's appearance is a strong predictor to the person's behaviour. This paper aims to validate if one can actually "judge" a finance scholar's academic research performance by the person's facial masculinity, which is gauged by the facial width-height-ratio. By examining the finance scholars from the top 50 US institutions on the Times Higher Education's World University Rankings 2021, I find that there are evident differences in research productivity between the scholars with the high and low facial width-height-ratio, measured in both quantity and quality. Particularly, I document that the more facial-masculine scholars outperform the less facial-masculine scholars by 13% and 14% for the productivity quantity and quality, respectively. However, the differences seemingly exist only in male but not female scholars. My findings are robust to fitting estimation with different models as well as testing the relationship among different sub-samples.

Table of Contents

List of Figures

List of Tables

Attestation of Authorship

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

SIGNED: __________

DATE: July 2021

Acknowledgements

First and foremost, I am grateful for my supervisors' great support and genuine advice. Particularly, a big thanks to:

> my main supervisor Professor Nhut (Nick) Nguyen Hoang for his kind guidance, advice and patience;
>
> my second supervisor Dr Shushu Liao for her great technical support, advice and feedbacks;
>
> and my third supervisor Professor Aaron Gilbert for his genuine support, advice and feedbacks.

Also, I would like to thank the Finance Department of the Business School for the scholarship, awards, teaching assistant opportunities as well as all the supports during my master study in AUT.

I have really enjoyed doing this research. It has not only provided me an opportunity for learning by doing but also given me a taste of how the journal publication works in academia.

Chapter 1. Introduction

The well-known English idiom, “Don’t judge a book by its cover”, has been popular for centuries. It metaphorises that “outward appearance cannot be an indicator of someone or something’s value or worth” or “you cannot know what someone or something is like just by looking the person or thing’s appearance”.[1] On the contrary, a vast literature has documented that an individual’s appearance is a strong predictor to his/her behaviour. In particular, facial characteristics provide important information about a person’s emotion, personality, and behavioural dispositions (Carré & McCormick, 2008) and “facial structure is a reliable cue of aggressive behaviour” (Carré et al., 2009).

More specifically, research shows that facial width-height-ratio (fWHR) correlates with a series of masculine behavioural traits. For instance, fWHR is positively associated with aggressive behaviour (e.g., Carré & McCormick, 2008; Carré et al., 2009; McCormick et al., 2010; Geniole et al., 2012; Altschul et al., 2019), dominance trait (e.g., Geniole et al., 2014; Mileva et al., 2014; Anderl et al., 2016; Wen & Zheng, 2020), achievement-driving and status-striving (e.g., Lewis et al., 2012; Tsujimura & Banissy, 2013; He et al., 2019), competitiveness (Stirrat & Perrett, 2012) and assertiveness (Kausel et al., 2018). Additionally, fWHR also links with trustworthiness (Stirrat & Perrett, 2010; Ormiston et al., 2017), risk-seeking (e.g., Jia et al., 2014; Lu & Teo, 2018; Kamiya et al., 2019), unethical behaviours (e.g., Haselhuhn & Wong, 2012; Geniole et al., 2014; Jia et al., 2014) and egocentricity (Anderl et al., 2016).

In conjunction with Bertrand and Schoar’s (2003) finding that managers’ masculine behaviours define managing styles and thus, affect managerial decision-making, and eventually transmit into different corporate behaviours and outcomes, many accounting and finance studies examine the relationship between fWHR and managerial behaviours along with outcomes in a corporation context. For example, Jia et al. (2014) document that firms run by more masculine-faced male CEOs (larger fWHR) are more likely to engage in financial misreporting, insider trading and option backdating. Hahn et al. (2017) note that firms having leaders with higher-than-normal fWHR tend to donate more to charities and have higher environmental awareness. Additionally, the authors show that fWHR has a positive association with social rank in both profit and non-profit organisations. Moreover, Kamiya et al. (2019) demonstrate positive correlations between a male CEO’s fWHR and his firm’s stock return volatility, financial leverage and frequency of acquisitions. However, to the best of my knowledge, little research is conducted to examine the relationship between facial masculinity and academic performance on scholars from high-

[1] The source is from https://www.theidioms.com/cant-judge-a-book-by-its-cover/.

educational institutions. Academic research productivity offers a useful and valuable setting to evaluate the link between facial features and performance since faculty members' pictures are publicly attainable and their research performance are largely measurable.

My central testable hypothesis is that scholars with higher fWHR perform better in their research output. I gather information for finance scholars from the top 50 institutions based on the Times Higher Education's (THE) World University Rankings 2021 of accounting and finance subjects in the US region. Particularly, I hand collect the scholars' pictures and personal information from each institution's website and measure their fWHR using the Python program. Meanwhile, I manually collect the publications of each scholar from the Business Source Complete (EBSCO) database. I offer a close examination of both publication quantity and quality. I count the number of publication records and use the Journal Impact Factors (JIF) from the Clarivate's Journal Citation Reports (JCR) to infer publication quality. By fitting a multivariate linear regression model, I find that more masculine-faced scholars publish more and higher-quality papers. However, this positive relationship seemingly only appears in the male but not the female scholars. Finally, I show that my results are robust to estimation with different models, the use of alternative performance measures and restricting my sample to scholars with high-quality pictures, male scholars, and with two-word names.

A natural question that arises is how fWHR is associated with higher performance? Individuals with higher fWHR are shown to exhibit more masculine behaviours such as aggression, risk-taking, dominance and status-maintenance, perhaps through the mechanism of hormone testosterone influencing both facial structure and behaviours (e.g., Carré & McCormick, 2008; Stirrat & Perrett, 2012; Jia et al., 2014; MacDonell et al., 2018). In fact, aggressive and dominant behaviour are often associated with psychological sense of power (Keltner et al., 2003; Haselhuhn & Wong, 2012); while high-power people tend to be optimistic, provident and more goal-orientated (Overbeck & Park, 2001; Wong et al., 2011). Given these qualities, powerful people tend to be more successful, particularly in executive leadership roles (Wong et al., 2011), and thus have better performance. In addition, masculine-faced people tend to display some motivational behavioural traits, such as achievement-driving (e.g., Lewis et al., 2012; Tsujimura & Banissy, 2013; He et al., 2019), winning mentality (Tsujimura & Banissy, 2013) and assertiveness (Kausel et al., 2018). These positive traits potentially motivate them to achieve high performance.[2] Furthermore, academic publication environment is highly competitive (e.g., Card & DellaVigna, 2013; Schwert, 2021). Under such a competitive environment, high-fWHR individuals tend to exhibit more prosocial and altruistic behaviours rather than antisocial and egocentric tendencies, such as being more cooperative and generous, due to the status-seeking and achievement-striving

[2] "Positive" and "negative" behavioural and character traits are loosely defined in this paper. The traits that promote high performance are defined as "positive" traits; whereas those which lead to poor performance are defined as "negative" traits.

urges (van Honk et al., 2011; Stirrat & Perrett, 2012; He et al., 2019). As a result, high performance is expected. Finally, another possible mechanism for the relation between fWHR and performance is hormone testosterone. Literature from biology and neuroscience conjectures that testosterone affects masculine facial features (e.g., Dabbs & Dabbs, 2000; Lefevre et al., 2013). Testosterone can also directly influence performance. For instance, testosterone is known to serve as a performance-enhancing drug among athletes because it can create an advantage in sports (Hoberman, 2005; Smurawa & Congeni, 2007; Wood & Stanton, 2012). Some studies show that testosterone can stimulate brain to increase aggression and motivation for competition tendencies (Hermans et al., 2006; Gleason et al., 2009; Wood & Stanton, 2012).

This paper contributes to several strands of literature. First, this research provides the supportive evidence to the existing literature examining the relationship between facial structure and masculine behaviours. The prior studies document that there is a significant association between facial structure and masculine behaviours by measuring fWHR. Specifically, larger fWHR predicts more masculine behaviours (e.g., Carré & McCormick, 2008; Carré et al., 2009; Jia et al., 2014; He et al., 2019; Kamiya et al., 2019) . Additionally, this prediction is only valid in males but not females because fWHR is sexually dimorphic that men have higher facial ratios than women in general (e.g., Carré & McCormick, 2008; Carré et al., 2009), while others believe that fWHR is not a sexually dimorphic trait since the association is valid for women as well (e.g., Geniole et al., 2012; Arnocky et al., 2018; MacDonell et al., 2018; Wen & Zheng, 2020). Meanwhile, my results show that fWHR does predict masculine behaviours, and these behaviour tendencies are reflected on that scholar with large fWHR tends to publish more and higher quality papers. However, I do not find the sexual dimorphism in fWHR and provide limited evidence that these masculine behaviours do not exist in females due to the small female sample size.

Second, my results complement the literature regarding the association between fWHR and various outcomes in different contexts. Wong et al. (2011) examine the relationship fWHR and firm outcome in a corporate context and note that firm run by more facial-masculine CEOs achieve superior financial performance. Tsujimura and Banissy (2013) study the correlation between fWHR and athletes' performance in a sport context and demonstrate that baseball players with large fWHR tend to have better home run performance; while Kausel et al. (2018) test the linkage between fWHR and students' grades in a university context and argue that students with high fWHR tend to have better grades in non-quantitative courses. This paper complements the literature by examining the relationship between fWHR and research performance in a higher educational academia context and show that scholars with higher fWHR have better performance.

Third, I contribute to the research productivity literature. The importance of academic research has been documented in many extant studies. For instance, Jaffe (1989) argues that academic research has an indirect effect on industrial innovation. McLean and Pontiff (2016) assert that finance research helps promote financial market efficiencies. Jaffe et al. (2020) further highlight the importance by arguing that there are associations between research productivity and a country's intellectual and economic wealth. As a result, there is a large number of research conducted on examining scholars' research performance. For example, Chan et al. (2007) note that faculty size, catalyst effect and per capita budget have positive impact on scholars' research performance. I show that fWHR is a significant predictor for research output after controlling the institution effects, and also validate the research output hierarchy following the academic titles. Furthermore, Chan et al. (2009) and Chan et al. (2011) demonstrate that the US-trained scholars have higher research productivity than those non-US trained scholars. However, I do not find evidence that there is a difference between the productivity of the US- and non-US- trained scholars, albeit this finding is only suggestive due to the small non-US-trained scholars sample size.

The rest of the paper is organised as follows: Chapter 2 is the relevant literature review and hypotheses development. I describe the data collection and the methodology in details in Chapter 3. The empirical results are presented in Chapter 4. I conduct the robustness tests in Chapter 5. Chapter 6 discusses the limitations and concludes the paper.

Chapter 2. Literature review and hypothesis development

2.1. fWHR, behaviours and performance

A growing body of research documents that fWHR is tied to aggressiveness and dominance. Carré and McCormick (2008), Carré et al. (2009) and McCormick et al. (2010) demonstrate that fWHR is a reliable predictor of aggressiveness for men in a laboratory context. This association has been validated in an out-of-laboratory context. For example, by examining 343 Japanese professional soccer players, Fujii et al. (2016) find that the players with higher fWHR received larger number of penalty cards (yellow and red) in the forward offensive position. Additionally, Anderl et al. (2016) report that fWHR is significantly positively correlated with fearless dominance, while Wen and Zheng (2020) find that fWHR is significantly associated with dominance in women, but predicts physical violence in men in China.

Meanwhile, Keltner et al. (2003) and Haselhuhn and Wong (2012) argue that aggressiveness and dominance is associated with a psychological sense of power. Powerful people are better at spotting opportunities and tend to be big-picture thinkers, optimistic, provident, more task-focused and goal-orientated (Overbeck & Park, 2001; Wong et al., 2011), all of which are positive psychological foundations for the emergence of individual and organisational success. For example, Wong et al. (2011) assert that an individual's aggressiveness may be compensated by organisational success by showing that firms led by male CEOs with high fWHR tend to outperform those run by CEOs with low fWHR. In addition, after examining the face photos of mixed martial arts (MMA) fighters, Třebický et al. (2013) conclude that fWHR is positively associated with their actual fighting success due to perceived aggressiveness. Furthermore, Lebuda and Karwowski (2016) note that the stronger positive relationship between the total number of Nobel prize nominations and the chance of being granted the prize exists among higher fWHR literature writers due to possible higher impulsivity and dominance characteristics.

Some researchers argue that fWHR has an association with achievement-driving, winning mentality, and assertiveness behavioural and character traits. For example, by examining the 29 former US presidents in the history and utilising psychometric analysis measuring their character traits, Lewis et al. (2012) support that fHWR is also correlated with achievement-driving and status-striving along with aggression and dominance character traits. He et al. (2019) contend that fWHR is positively correlated with achievement drive among financial analysts. Tsujimura and Banissy (2013) suggest that fWHR is positively associated with achievement drive, winning mentality and aggression. Kausel et al. (2018) believe fWHR is positively related to assertiveness.

With the support of these positive behavioural and character traits, one would naturally expect that high fWHR predicts high performance. In line with this expectation, He et al. (2019) find that financial analysts with larger fWHR exert more effort by conducting more corporate site visits, which eventually lead to better performance due to their higher achievement drive. Tsujimura and Banissy (2013) document that there is positive correlation between fWHR and home run performance for professional Japanese baseball players because of possible achievement drive, winning mentality and aggression.[3] Furthermore, by examining 231 university students in a school of business and economics, Kausel et al. (2018) show that fWHR predicts the students' grades in non-quantitative courses due to possible assertiveness. These findings lead to my null hypotheses:

H1a: Scholars with larger fWHR have higher research performance in terms of quantity.

H1b: Scholars with larger fWHR have higher research performance in terms of quality.

However, this aggressiveness, dominance, assertiveness or achievement-driving and status-striving behavioural characteristics may also play a "Macbeth Tragedy" on an individual's performance (Chan et al., 2020). Chan et al. (2020) raise a conflicting argument with He et al. (2019) that masculine-faced analysts' superior performances might be an illusion due to their opportunistic forecasting behaviour. The authors propose that these positive traits may eventually evolve into unethical and opportunistic behaviours due to the obsession of achievement and ambition. They find that analysts with larger fWHR tend to provide more optimistic stock recommendations and, in fact, display lower stock picking ability.

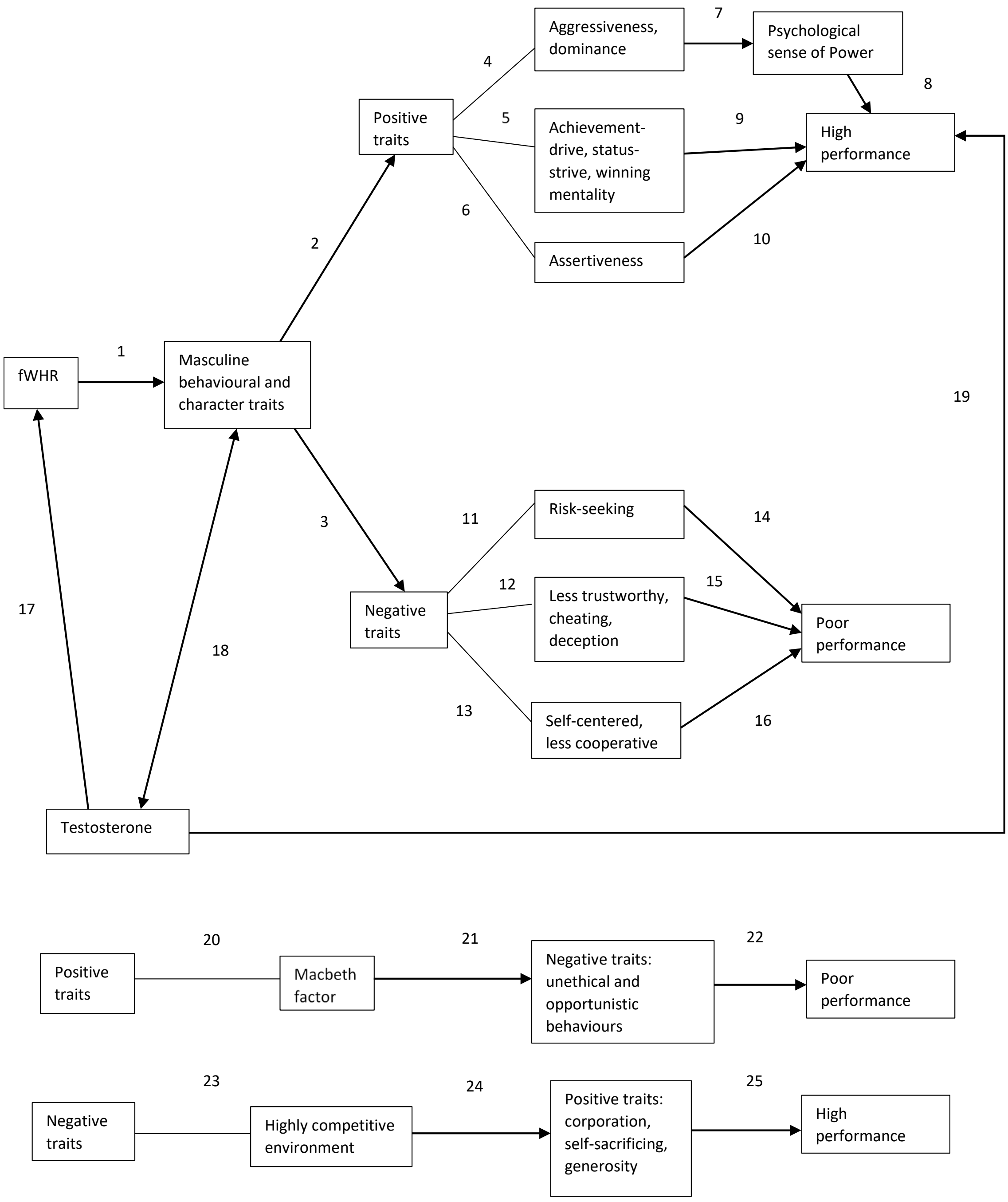
In addition, prior literature documents that high-fWHR is also correlated with unsociable behaviour tendencies, such as being less trustworthy (Stirrat & Perrett, 2010; Ormiston et al., 2017), risk-seeking (e.g., Jia et al., 2014; Lu & Teo, 2018; Kamiya et al., 2019), more willing to cheat (Haselhuhn & Wong, 2012; Geniole et al., 2014), egocentric (Anderl et al., 2016) and less cooperative (Haselhuhn et al., 2014). These negative character traits make individuals socially undesirable and thus jeopardise their performance, which eventually leads to poor outcomes. For instance, Lu and Teo (2018) show that more facially masculine fund managers underperform the less masculine counterparts and are exposed to greater operational risk. Haselhuhn et al. (2014) note that men with larger fWHR exhibit lower negotiation performance due to being less cooperative. These findings lead to my alternative hypotheses:

H2a: Scholars with larger fWHR have lower research performance in terms of quantity.

H2b: Scholars with larger fWHR have lower research performance in terms of quality.

[3] In a later paper, Mayew (2013) criticises Tsujimura and Banissy's (2013) finding by arguing that body mass but not fWHR predicts Japanese professional baseball players' superior home run performance. However, the author does not exclude the possibility that fWHR might play a more important role than body mass. Furthermore, the author proposes that testosterone possibly contributes most behind all these.

Notwithstanding, some researchers assert that the antisocial traits implied by high-fWHR are only situational when out-group competition is low. They further argue that fWHR people tend to exhibit prosocial tendencies, such as increased corporation, self-sacrificing and generous to in-group members under highly competitive environments (e.g., van Honk et al., 2011; Stirrat & Perrett, 2012; He et al., 2019), and thus have better performance (He et al., 2019). Given that academic publication environment is highly competitive (e.g., Card & DellaVigna, 2013; Schwert, 2021), these findings infer that high fWHR implies high performance. Again, I merge these arguments into my null hypotheses. The overall literature flow-chart is shown in *Figure 1* below.



This flowchart shows the relationship between fWHR, testosterone, masculine behaviours, and performance.

- The relation between fWHR and masculine behaviours [1] (e.g., Carré & McCormick, 2008; Carré et al., 2009; McCormick et al., 2010; Stirrat & Perrett, 2010; Haselhuhn & Wong, 2012; Haselhuhn et al., 2014; Jia et al., 2014; He et al., 2019; Kamiya et al., 2019; Chan et al., 2020).
- The relation between masculine behaviours and positive traits [2, 4, 5, 6] (e.g., Wong et al., 2011; Lewis et al., 2012; Stirrat & Perrett, 2012; Třebický et al., 2013; Tsujimura & Banissy, 2013; Lebuda & Karwowski, 2016; Kausel et al., 2018; He et al., 2019).
- The relation between aggressiveness and dominance, sense of power and high performance [7, 8] (e.g., Overbeck & Park, 2001; Keltner et al., 2003; Wong et al., 2011; Haselhuhn & Wong, 2012; Třebický et al., 2013; Lebuda & Karwowski, 2016).
- The relation between achievement-drive, status-striving, and winning mentality and high performance [9] (Tsujimura & Banissy, 2013; He et al., 2019).
- The relation between assertiveness and high performance [10] (Kausel et al., 2018).
- The relation between masculine behaviour and negative traits [3, 11, 12, 13] (e.g., Stirrat & Perrett, 2010; Haselhuhn & Wong, 2012; Geniole et al., 2014; Jia et al., 2014; Anderl et al., 2016; Ormiston et al., 2017; Kamiya et al., 2019).
- The relation between negative traits and poor performance [14, 15, 16] (e.g., Haselhuhn et al., 2014; Lu & Teo, 2018; Chan et al., 2020).
- The relation between fWHR and testosterone [17] (e.g., Verdonck et al., 1999; Dabbs & Dabbs, 2000; Vanderschueren et al., 2004).
- The relation between testosterone and masculine behaviours [18] (e.g., Worthman & Konner, 1987; Mazur & Booth, 1998; Schultheiss et al., 1999; Dabbs & Dabbs, 2000; Archer, 2006; Josephs et al., 2006; Apicella et al., 2008; Coates et al., 2009; Gleason et al., 2009; Wood & Stanton, 2012; Boksem et al., 2013; Fisk et al., 2017; Nadler et al., 2018).
- The relation between testosterone and high performance [19] (e.g., Dabbs & Dabbs, 2000; Hoberman, 2005; Smurawa & Congeni, 2007; Wood & Stanton, 2012).
- The mechanism of positive traits converting into negative traits via Macbeth factor, which eventually leads to poor performance [20, 21, 22] (Chan et al., 2020).
- The mechanism of negative traits converting into positive traits within highly competitive environments, which eventually leads to high performance [23, 24, 25] (e.g., van Honk et al., 2011; Stirrat & Perrett, 2012; He et al., 2019)

Figure 1. The relationship diagram between fWHR, behaviours, and performance

2.2. Testosterone, fWHR and performance

While a plethora of research documents the positive relationship between fWHR and a series of behavioural characteristics, the key factor contributing to this association remains a debatable subject. The prominent conjecture from the neuroscience is that androgens, such as testosterone, play an important role in this relationship (e.g., Carré & McCormick, 2008; Carré et al., 2009; Lefevre et al., 2013). In fact, testosterone has an impact on adolescent cranial growth, which directly influences facial structure development (e.g., Verdonck et al., 1999; Vanderschueren et al., 2004). On the other hand, testosterone is associated with many masculine behavioural and character traits, such as aggressiveness (e.g., Worthman & Konner, 1987; Archer, 2006), dominance (e.g., Mazur & Booth, 1998; Schultheiss et al., 1999; Josephs et al., 2006), risk-taking (e.g., Apicella et al., 2008; Coates et al., 2009; Fisk et al., 2017; Nadler et al., 2018) and competitiveness (e.g., Gleason et al., 2009; Wood & Stanton, 2012; Boksem et al., 2013). Moreover, by conducting a comprehensive research on examining the testosterone levels of 8,000 men, women and children on the foundation of the extant relevant literature, Dabbs and Dabbs (2000) conclude that testosterone virtually influences every aspect of human mind and destiny, which include physical growth, psychological development, cognition, what types of jobs people choose, what kind of people they become and so on. The authors further identify the significantly positive connections between testosterone and various aggressive behaviours, such as criminal behaviour and altruism. Given these connections, unsurprisingly, the positive association between fWHR and the behavioural traits can be explained by testosterone.

Particularly, many researchers consider testosterone as a "social hormone" since it is closely related to human social behaviours, such as maintenance of social status, status-striving, reputable-status seeking and/or achievement-driving (e.g., Eisenegger et al., 2010; Eisenegger et al., 2011; van Honk et al., 2011; Smeets-Janssen et al., 2015; van Honk et al., 2016; Fisk et al., 2017). As discussed in the previous section, these qualities are the vital ingredients for pursuing superior performance. Moreover, testosterone is reported to promote humans to grow bigger muscles, think quicker and run faster (e.g., Dabbs & Dabbs, 2000; Hoberman, 2005). As a result, many athletes consider testosterone doping as a performance-enhancing instrument (e.g., Hoberman, 2005; Smurawa & Congeni, 2007; Wood & Stanton, 2012).

Conversely, testosterone is also linked to some antisocial behaviours, such as abuse of trust and betrayal (Boksem et al., 2013), egoistic and materialistic (Wright et al., 2012; van Honk et al., 2016) and less collaborative (Wright et al., 2012). These undesirable behavioural traits could potentially deteriorate one's performance.

Given that both prosocial and antisocial behaviours are associated with testosterone, many studies turn to examine the relationship in different contexts. For example, Boksem et al. (2013) argue that testosterone tends to promote prosocial behaviours when social challenges and threats are absent. Josephs et al. (2003) show that high-testosterone males and females perform reversely on a math test under different stereotype-relevant situations. Mehta et al. (2008) demonstrate that testosterone promotes competitiveness in a winning situation.

While the true mechanisms of fWHR and behavioural characteristics are still remaining to be discovered, it is beyond the scope of this research, and I will leave it to the future studies. Nonetheless, the foundation of this paper has been validated by a vast literature, that is, fWHR is an "honest signal" of masculine behaviours (e.g., Carré & McCormick, 2008; Carré et al., 2009; Stirrat & Perrett, 2012; Jia et al., 2014).

2.3. Literature of academic research productivity

Academic research and its associated impact on industrial performance has garnered considerable attention. For example, Jaffe (1989) finds that university research significantly produces commercial spill-over effects on corporate patents and has an indirect effect on local innovation. Mansfield (1991) suggests that academic research greatly promotes the development and commercialisation of new products in many industries. Particularly, Grossman et al. (2001) examine the contributions of academic research to the industrial performance in five industries and find that all the contributions are substantial, albeit they vary from one industry to another.

In addition, Salter and Martin (2001) review a troupe of literature on the economic benefits of publicly funded basic research and conclude that the economic benefits from basic research are undoubtedly significant but in a variety of forms. Furthermore, by examining the research productivity across 27 disciplines over 62 countries, Jaffe et al. (2020) contend that research productivity is related to a country's intellectual and economic wealth.

Meanwhile, economic and social benefits of finance research is nothing short of remarkable. For instance, McLean and Pontiff (2016) demonstrate that academic research in finance promotes financial market efficiency by informing investors mispricing opportunities. Chan et al. (2004) list three institutional-level benefits of research success: i) attracts endowments, ii) provides direct information to administrators for recruitment, promotion, and salary adjustment decisions, and iii) assists graduate student enrolments. Interestingly, Swidler and Goldreyer (1998) show that the very first publication of the top 4 finance journals, namely the *Journal of Finance*, the *Journal of Financial Economics*, the *Review of Financial Studies* and the *Journal of Financial and Quantitative Analysis*, is worth between $19,493 and $33,754 to a scholar's salary.

A growing body of literature examines what determines finance research performance. For example, Fishe (1998) notes that full professors from the top 20 departments have 1 out of 3 papers on average published in the top three finance journals: namely the *Journal of Finance*, the *Journal of Financial Economics* and the *Review of Financial Studies*. Chan et al. (2001), Chan et al. (2005) and Chan et al. (2011) report that the universities in the Asia-Pacific region have significantly improved their research performance through the years from 1990 to 1998. By examining the global outputs in finance between 1990 and 2004, Chan et al. (2007) state that the most productive countries are US, UK, Canada, Hongkong and Australia, while US dominates the top 100 institution productivity ranking list occupying a total of 78 places. The authors also find that there are several institutional-level factors influencing research outputs, such as the faculty size, budget constraint and catalyst effect.

Another line of research focuses on gender inequality in research productivity. In fact, studies show that female scholars have low representation in the academic finance and economics profession (e.g., Boustan & Langan, 2019; Lundberg & Stearns, 2019; Sherman & Tookes, 2019; Hengel & Moon, 2020), and publish fewer papers than males (e.g., Long, 1992; Bentley, 2012; Lundberg & Stearns, 2019; Sherman & Tookes, 2019; Ghosh & Liu, 2020; Hengel & Moon, 2020); but their papers receive more citations (e.g., Long, 1992; Sherman & Tookes, 2019; Card et al., 2020; Hengel & Moon, 2020); though the gender gap is shrinking slowly in recent years (e.g., McDowell et al., 2006; Bentley, 2012; Lundberg & Stearns, 2019; Sherman & Tookes, 2019). Particularly, by examining the female and male academics' research productivity globally over 25 countries during the recent COVID-19 outbreak, Cui et al. (2021) show that female scholars' productivity decreased by 13.9% in relation to males', while the total research productivity increased by 35% in the 10 weeks after the lockdown in the US. Furthermore, the authors note that this gender inequality is more severe among the top-ranked universities. To date, little empirical evidence is provided on the association between scholars' research productivity and a physical trait that correlates with their achievement. This paper intends to fill this gap.

Chapter 3. Methodology and data

I start with identifying the institutions of my interest from the THE World University Rankings 2021 of accounting and finance subjects in the US region. Next, I hand-collect each finance scholar's information from the selected institutions' websites, which include the scholar's publication name, academic title, resume and photo. Then I apply my Python codes to calculate the fWHR. Last, the publications of each scholar are collected from the EBSCO database. I manually collect all available publication information for each scholar as well as the number of co-authors and the publishing year for each publication. [4]

Next, I identity the journals from the latest Financial Times' Top 50 journal list (FT50). There are several reasons for choosing this journal list. First, it is because of the relevance and international academic recognition. The list is selected from the survey polls of over 200 top business schools all over the world. The latest version was created in 2016, which expanded to include 50 journals from the previous 45 journals. [5] Second, it includes the most influential academic Business and Economics journals and "covers a good balance of subfields and also leaves room for a practitioner journal and openness for innovation" (Fassin, 2021). The diversity characteristic of the FT50 list accommodates the interdisciplinary nature of the finance discipline.

I measure a scholar's productivity in two aspects: quantity and quality. The quantity measure is to count the total number of publications of the scholar from the FT50 journal list. The larger counts mean the higher productivity. The publication quality, on the other hand, is measured by applying the journal impact factor (JIF) from the Clarivate's JCRs. Furthermore, by following Schwert's (2021) suggestion, I divide the quantity and quality measures separately into two more variations: with and without considering co-authorship. For instance, if a paper is written by professors A, B and C, under co-authorship adjustment, each professor will get 1/3 score for this paper, whereas they each will get a full score of one for this paper without adjustment for co-authorship. Then, the final quality adjusted score of a particular article for the author is the co-authorship (un)adjusted quality score multiplied by the corresponding JIF score from the Clarivate's JCRs for the year of the publication. The total academic performance quality score of a scholar is the total sum of each quality-adjusted points of his/her publications. [6]

Finally, I apply a multivariate linear regression to examine the relationship between fWHR and research productivity. The full specification of my model is:

[4] The detailed data collection processes are described in *Appendix D*.
[5] The source is from https://www.ft.com/content/3405a512-5cbb-11e1-8f1f-00144feabdc0.
[6] Please see *Appendix C* for the productivity quality measure selection based on the existing literature.

$$\log(PRODUCITVITY\_SCORE_i) = \alpha_i + \beta_1 * FWHR\_HIGH_i + \beta_2 * MALE_i + \beta_3 * US_i + \beta_{4-8} * TITLE_i + \beta_9 * 1st\_YEAR_i + \beta_{10} * EXP_i + \beta_{11} * (EXP_i)^2 + \beta_{12} * THE\_RANK_i + \varepsilon_i \quad [1]$$

where the $\log(PRODUCITVITY\_SCORE_i)$ is the natural log transformation of the productivity scores for scholar i, and $FWHR\_HIGH_i$ is the dummy variable which is transformed from the continuous variable FWHR by following Jia et al.'s (2014) and Kamiya et al.'s (2019) suggestions, and is equal to one if scholar i's fWHR is greater than the median and zero otherwise; and β_1 is the corresponding coefficient of $FWHR\_HIGH_i$. A statistically and economically significant β_1 with a positive sign will indicate that higher fWHR is associated with better research productivity.

$MALE_i$ is the gender control dummy variable equal to one for male scholars and zero otherwise, and β_2 is the corresponding coefficient. Since Sherman and Tookes (2019) document that female scholars publish less papers than males, and Card et al. (2020) find that females publish higher-quality paper than males, I expect a statistically significant β_2 with a positive sign for the productivity quantity scores. However, the significance and sign of β_2 for the productivity quality scores are unknown. This is because the quality scores are calculated as the sum of the total number of JIF-adjusted publications. Although females publish less papers, their papers receive more citations. Thus, the ex-ante combined effects cannot be determined.

The US_i variable is equal to one if scholar i attained the highest qualification in the US and zero otherwise, and β_3 is the corresponding coefficient. Based on Chan et al.'s (2009) and Chan et al.'s (2013) findings that scholars with the highest qualifications earned in the US are more productive than those with theirs obtained in a foreign country, I expect a statistically significant β_3 with a positive sign. The categorical variable $TITLE_i$ includes $PROF_i$, $ASSO\_PROF_i$, $ASSIS\_PROF_i$, SEN_LEC_i and $OTHER_i$, which represent the five levels with the baseline $PRAC_i$. The six levels of the academic titles are professor, associate professor, assistant professor, senior lecturer, other, and practice. β_4, β_5, β_6, β_7 and β_8 are the corresponding coefficients. In addition, a professor is perceived to be more productive than an associate professor, who is likely to be more productive than an assistant professor. According to this natural productivity hierarchy of the titles, I expect these coefficients are statistically significant and have positive signs, and their magnitudes are in a descending order.

The $1st\_YEAR_i$ variable represents the year of each scholar's very first publication with the coefficient β_9. Card and DellaVigna (2013) document that the number of published articles in the top journals is declining as the acceptance rate is falling from 15% to 6% between 1990 and 2012. In addition, Schwert (2021) finds the rejection rate of the *Journal of Financial Economics* increase through the years. Given these findings, presumably, a paper would get published more easily in the earlier years than later, which implies that a researcher who started publishing earlier tends to have higher research performance. Therefore, I expect the coefficient β_9 to be significant with a negative sign to accommodate the inverse relationship. EXP_i and $(EXP_i)^2$ are the publication experience control variables with the corresponding coefficients β_{10} and β_{11}. Evidently, more experience indicates higher performance. However, this relationship might not be linear since a scholar most likely receives a diminishing return of benefits along with the increase of the experience. For example, a scholar at 70 years old with 50 years' publication experience might not be as creative as a scholar at 50 years old with 30 years' experience. As a result, I would expect both the β_{10} and β_{11} to be statistically significant with a positive and negative sign individually to accommodate the quadratic association.

Finally, the THE_RANK_i variable is the THE ranking of the institution where scholar i works, which controls the possible affiliation differences. β_{12} is the corresponding coefficient. Chen and Huang (2007) argue that the authors affiliated with top-ranked institutions tend to have higher performance. In addition, Chan et al. (2007) suggest that faculty size, per capita budget and catalyst effect are also the factors that affect research outputs. Therefore, I expect β_{12} to be statistically significant with a negative sign given that the THE ranking is in an ascending order.

Chapter 4. Empirical results

4.1. Explanatory data analyses

I start with examining the three datasets individually, which I have collected above.

4.1.1. fWHR data analysis

There is a total of 1,916 finance scholars' information collected from the top 50 US institutions. Among these scholars, 1,636 of them have pictures displayed on the faculty websites and 1,327 of the pictures are classified as high-quality pictures based on the selection criteria. *Panel A* in *Table 1* shows that the mean of fWHR (1.864) is roughly equal to the median (1.863), which implies that the fWHR distribution is symmetric. This feature yields an even split when I transform the fWHR into the dummy variable FWHR_HIGH by following Jia et al.'s (2014) and Kamiya et al.'s (2019) suggestions. Additionally, the number of males dominates the sample with the percentages of 82.5% and 84.2% for all and high-quality pictures respectively.[7] These leave female scholars with less than 20% for both samples. Unsurprisingly, with such low female percentages, the fWHR distributions do not change much after removing the females from the sample.[8]

In addition, the mean differences between the fWHR of males and females are insignificant for both all and high-quality picture samples. These results suggest that fWHR is not sexually dimorphic, which contradicts the findings from some literature (e.g., Weston et al., 2007; Carré et al., 2009; Short et al., 2012) but coincides with the findings from other studies (e.g., Lefevre et al., 2012; Özener, 2012; Gómez-Valdés et al., 2013). However, I believe these results are only suggestive due to the small female population in the sample and thus, further investigation is required.

Table 1. The data and descriptive statistics

Panel A. Descriptive statistics summary of fWHR

Panel A reports the descriptive statistics of fWHR for the scholars with pictures and high-quality pictures, respectively. The top half of the table reports the descriptive statistics of fWHR for the scholars with pictures, whereas the bottom half of the table reports the descriptive statistics of fWHR for the scholars with high-quality pictures. Particularly, N represents the number of data points. 25th represents the 25th percentile and 75th represents the 75th percentile. The mean difference t-tests are testing the mean fWHR difference between males and females.

Gender	N	Mean	SD	25th	Median	75th	Min	Max
				All pictures				
All	1,636	1.864	0.142	1.762	1.863	1.958	1.445	2.410

[7] 82.5% = 1350/1636 and 84.2% = 1117/1327.
[8] The distribution plots are shown in *Appendix B*.

Female	286	1.868	0.133	1.774	1.86	1.957	1.513	2.294
Male	1,350	1.864	0.144	1.760	1.863	1.958	1.445	2.410
Male-Female Difference	t=-0.480	P-value=0.632						
			High quality pictures					
All	1,327	1.867	0.143	1.765	1.865	1.96	1.445	2.41
Female	210	1.874	0.136	1.777	1.862	1.966	1.513	2.294
Male	1,117	1.865	0.144	1.763	1.866	1.959	1.445	2.410
Male-Female Difference	t=-0.799	P-value=0.425						

Panel B. The data summary for each institution [9]

This table reports the data summary for each institution. THE_RANK represents the THE 2021 US regional institution rankings of accounting and finance subjects. N denotes the number of scholars. Mean_fWHR and SD_fWHR represent the mean and standard deviation of fWHR of the corresponding institution. Last, Female% represents the percentage of the female finance scholars over all the scholars. Foreign% represents the percentage of the scholars with their highest qualification obtained in non-US countries.

THE_RANK	Institution	N	Mean_fWHR	SD_fWHR	Female%	Foreign%
1	Stanford University	42	1.902	0.135	0.143	0.075
2	Harvard University	53	1.867	0.150	0.151	0.080
3	Massachusetts Institute of Technology	45	1.831	0.141	0.089	0.100
4	Yale University	21	1.862	0.113	0.095	0
5	The University of Chicago	41	1.871	0.127	0.122	0.125
6	Johns Hopkins University	41	1.862	0.157	0.122	0.200
7	University of Pennsylvania	76	1.892	0.151	0.125	0.085
8	University of California, Los Angeles	36	1.845	0.146	0.083	0.056
9	Columbia University	124	1.863	0.132	0.205	0.127
10	Duke University	19	1.841	0.139	0.158	0.053
11	University of Michigan-Ann Arbor	34	1.910	0.136	0.200	0.067
12	Northwestern University	36	1.861	0.150	0.167	0.028
13	New York University	113	1.858	0.132	0.113	0.081
14	Carnegie Mellon University	20	1.919	0.163	0.211	0.250
15	University of Washington	33	1.843	0.145	0.207	0.061
16	University of California, San Diego	14	1.885	0.190	0.143	0.071
17	Georgia Institute of Technology	15	1.879	0.108	0.200	0.214
18	University of Texas at Austin	99	1.849	0.149	0.221	0.081
19	University of Illinois at Urbana-Champaign	72	1.860	0.145	0.200	0.031
20	University of Wisconsin-Madison	30	1.869	0.130	0.148	0.045
21	Washington University in St Louis	32	1.868	0.165	0.103	0.161

[9] The *University of California, Santa Barbara* at the rank 26 is left out. This is because it does not have a finance related department but has an economics department. Instead, *Northeastern University* at the rank 51 is included.

22	University of Southern California	82	1.882	0.133	0.203	0.055
23	Boston University	48	1.800	0.128	0.098	0.219
24	University of North Carolina at Chapel Hill	30	1.839	0.142	0.241	0.100
25	University of California, Davis	7	1.817	0.146	0.143	0
27	Ohio State University	49	1.898	0.133	0.286	0.093
28	Emory University	22	1.904	0.156	0.091	0.091
29	University of Minnesota	34	1.877	0.129	0.241	0.040
30	University of Maryland, College Park	29	1.834	0.112	0.276	0.069
31	Purdue University West Lafayette	16	1.875	0.154	0.188	0.167
32	University of California, Irvine	21	1.880	0.146	0.286	0.143
33	Dartmouth College	16	1.851	0.148	0.133	0.067
34	Michigan State University	34	1.867	0.201	0.176	0.091
35	Vanderbilt University	11	1.983	0.117	0.182	0.091
36	Penn State University	34	1.829	0.123	0.129	0.030
37	University of Virginia	14	1.957	0.117	0.143	0.214
38	Georgetown University	30	1.880	0.150	0.167	0.034
39	Case Western Reserve University	14	1.843	0.181	0.231	0.154
40	University of Arizona	23	1.879	0.101	0.261	0.043
41	Rice University	29	1.925	0.138	0.185	0.038
42	University of Colorado Boulder	37	1.852	0.161	0.147	0.118
43	University of Pittsburgh	9	1.850	0.077	0.167	0
44	Indiana University	61	1.799	0.134	0.163	0.114
45	University of Rochester	22	1.845	0.144	0.136	0.182
46	University of Florida	37	1.868	0.111	0.088	0.062
47	Tufts University	12	1.852	0.141	0.083	0
48	Rutgers, the State University of New Jersey	45	1.797	0.121	0.211	0.100
49	University of Alabama	34	1.872	0.097	0.409	0.043
50	University of Notre Dame	57	1.897	0.137	0.173	0.158
51	Northeastern University	63	1.859	0.148	0.258	0.063

Two evident results can be seen from *Panel B* in *Table 1*. First, the female ratio is significantly low for all the 50 institutions.[10] The average percentage of female scholars across all the 50

[10] Female ratio is the percentage of female scholars in the sample. It is calculated as the number of female scholars divided by the total number of male and female finance scholars for each institution. It is worth mentioning that the sum of male and female finance scholars is not equal, but very close to the total number of scholars (N). This is because I have been conservative when collecting the data. If a scholar's personal bio neither implies a gender nor display any photos, I leave the gender of this scholar as a missing value.

institutions is around 17%, including the *University of Alabama* with the highest female ratio at 41% and the *University of California, Los Angeles*, and *Tufts University* with the equally lowest female ratio both at 8%. The result coincides with Sherman and Tookes' (2019) findings in two ways. First, they find women represent around 16% of finance academics in the top-100 US business schools between 2009 and 2017. The average female percentage in my sample is around 17%, which is very close to 16%. Second, they demonstrate a very slow increase in the female percentage of the finance faculty from 14.9% in 2009 to 16.8% in 2017; whereas my result shows that the female ratio is 17% in 2021, which is indeed a very small increase from 16.8% in 2017.

Panel B also shows that the foreign ratio is significantly low for all the 50 institutions. The average percentage of scholars with the highest qualification granted from a non-US institution is around 9% across all the institutions in the sample. *Carnegie Mellon University* has the highest foreign ratio at 25%; whereas *Yale University*, the *University of California, Davis*, the *University of Pittsburgh*, and *Tufts University* have a tie for the last place at 0%. It is worth noting that *Stanford University*, *Harvard University*, *Massachusetts Institute of Technology*, and *Yale University* have only an average of 6% foreign ratio. This result suggests that the top US institutions tend to hire academic staff with domestic educational backgrounds. There are two possible reasons. First, a US-trained and/or US-experienced scholar tends to have higher research productivity than those from the non-US background (Chan et al., 2009; Chan et al., 2011). Second, the non-US scholars tend to focus more on the non-US finance research because of the "home bias" (Karolyi, 2016). Meanwhile, the top journals, which are mostly US-based, tend to publish the US related research over the non-US related studies because of the "foreign bias" (Karolyi, 2016). As a result, the non-US scholars tend to be disadvantaged over the US scholars in terms of the research productivity measures.

4.1.2. JCR data analysis

In order to cover the possible name changes or early termination of any journals, I collect all the journal lists across all the available years of all the four categories: finance, economics, business and management from the JCRs. At the end, I have a total of 125 journals from the finance category; 427 journals from the economics category; 173 from the business category; and 253 journals from the management category. Overall, I have 809 journals in total.[11] This total number is less than the actual sum of the numbers of the four categories because JCRs recognise the interdisciplinary feature of some journals and thus include them into multi-categories. For

[11] In fact, the total number of all the journals from the four JCR categories is 808. The journal *Production and Operations Management* from the FT50 journal list is classified into the engineering category by the JCRs. This is also one of the reasons why I do not follow the JCR journal list to measure the finance scholars' productivity in this research. Apparently, the FT50 journal list is more relevant to the publications from the finance discipline.

example, the *Journal of Finance* has been included into both the finance and economics category. Furthermore, this final JCR journal list is used to determine the scholars' publication experience.

Table 2. The descriptive statistics of JIFs for each journal from the FT50 journal list

This table presents the descriptive statistics of JIF, JIF_adj and JIF_five for each journal from the FT50 journal list. Specifically, JIF represents the impact factors of a particular journal. JIF_adj represents a journal's impact factors adjusted for self-citations, and JIF_five represents the five-year average impact factors of a journal. Finally, Occurrence represents the total number of publications for each journal. The journals are ranked based on the descending order of the total numbers of occurrences.

		JIF		JIF_adj		JIF_five	
Journal	*Occurrence*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Journal of Finance	2418	4.06	1.395	3.615	1.524	7.425	1.542
Journal of Financial Economics	2031	3.34	1.068	2.944	1.083	6.159	1.301
Review of Financial Studies	1451	2.802	1.268	2.518	1.163	5.308	1.214
Journal of Financial and Quantitative Analysis	748	1.39	0.551	1.309	0.541	2.588	0.523
American Economic Review	744	2.752	1.08	2.603	1.053	5.026	1.379
Management Science	401	2.095	0.907	1.87	0.881	4.045	0.811
Journal of Political Economy	352	3.423	1.147	3.34	1.146	6.081	0.73
Quarterly Journal of Economics	263	5.562	2.232	5.339	2.238	9.926	2.644
Harvard Business Review	211	2.268	1.302	2.264	1.302	3.294	1.703
Econometrica	196	3.026	0.836	2.815	0.799	5.405	0.649
Review of Finance	182	1.951	0.393	1.893	0.399	2.77	0.295
Review of Economic Studies	177	2.701	1.197	2.614	1.195	4.804	0.993
Journal of Accounting and Economics	174	2.54	1.171	2.117	1.145	5.119	1.15
Operations Research	119	1.92	0.361	1.722	0.36	2.838	0.405
Accounting Review	99	1.972	0.914	1.582	0.901	3.822	0.943
Journal of Accounting Research	84	2.219	1.117	1.872	1.071	4.219	1.142
Journal of Business Ethics	64	1.258	1.089	0.863	1.081	2.56	1.464
Journal of Applied Psychology	62	3.402	1.234	2.962	1.183	6.864	0.803
Journal of International Business Studies	61	3.118	2.287	2.4	1.755	6.368	1.957
Contemporary Accounting Research	46	1.473	0.524	1.293	0.505	2.445	0.772
Strategic Management Journal	38	3.261	1.174	2.832	1.107	6.655	0.87
Review of Accounting Studies	34	1.801	0.437	1.614	0.43	2.506	0.547
Academy of Management Journal	29	4.68	1.917	4.261	1.783	10.016	1.734
Journal of Management	27	3.964	2.847	3.701	2.764	8.573	3.117
Sloan Management Review	24	1.64	0.657	1.546	0.667	2.676	1.127
Journal of Management Studies	23	2.616	1.695	2.246	1.607	5.631	1.697
Journal of Consumer Research	21	2.802	1.084	2.27	1.06	4.927	1.206
Production and Operations Management	21	1.514	0.731	1.003	0.575	2.661	0.523

Human Resource Management	19	1.623	0.586	1.47	0.552	2.545	0.916
Journal of Marketing Research	19	2.58	0.862	2.319	0.867	4.483	1.105
Organization Science	19	2.615	0.999	2.3	0.896	5.683	0.363
Journal of Marketing	18	3.948	1.499	3.452	1.4	7.737	1.296
Academy of Management Review	17	5.975	2.276	5.494	2.192	11.171	2.222
Human Relations	17	1.615	0.923	1.411	0.864	3.11	1.014
Organizational Behaviour and Human Decision Processes	15	2.025	0.696	1.839	0.689	3.723	0.477
Accounting Organizations and Society	14	1.57	0.993	1.209	0.921	3.487	0.897
Journal of Business Venturing	14	2.617	2.107	2.308	1.94	5.96	2.808
Manufacturing and Service Operations Management	14	1.992	0.819	1.764	0.779	2.852	0.572
Administrative Science Quarterly	13	3.98	1.592	3.694	1.564	7.559	1.311
Research Policy	13	2.484	1.386	2.086	1.304	5.052	1.553
Information Systems Research	11	2.123	0.872	1.907	0.849	4.926	0.928
Journal of the Academy of Marketing Science	10	3.169	2.588	2.829	2.396	5.998	3.411
Entrepreneurship Theory and Practice	9	3.486	2.465	3.066	2.559	5.7	2.712
Journal of Management Information Systems	9	1.991	0.859	1.509	0.709	3.713	0.738
Journal of Consumer Psychology	8	2.156	1.234	1.51	0.849	3.405	1.057
Journal of Operations Management	8	3.98	1.624	3.371	1.46	6.994	2.142
Strategic Entrepreneurship Journal	6	2.601	1.422	2.313	1.393	3.746	1.173
MIS Quarterly	5	4.115	1.657	3.713	1.591	9.522	1.458
Marketing Science	3	2.344	0.869	1.587	0.463	3.455	0.509
Organization Studies	3	2.021	0.908	1.684	0.808	3.814	1.006

From *Table 2*, the FT50 journal list covers a diversity of journals from all the disciplines of a typical business school. Particularly, the well-recognisable top three finance journals, the *Journal of Finance,* the *Journal of Financial Economics*, and the *Review of Financial Studies* (e.g., Currie & Pandher, 2011; Sherman & Tookes, 2019; Bajo et al., 2020; Currie & Pandher, 2020) are included in the list. These confirm that the FT50 journal list includes the most influential academic journals and offers a good balance of academic fields (Fassin, 2021). In addition, the majority of the journals (total 38 journals) have the full 23-year JCR records from 1997 to 2019. This result coincides with Fassin's (2021) finding that "the FT50 list confirms the stability of the top-tier journals". The remaining 12 journals have a partial JCR record ranging from a minimum of 10 to a maximum of 21 years. Among these, the *Strategic Entrepreneurship Journal* and the *Review of Finance* have the minimum of 10-year JCR records and the *Journal of the Academy of Marketing Science* has the maximum of 21-year JCR records. These results match Fassin's (2021) claim that the FT50 list "leaves room for a practitioner journal and openness for innovation".

Furthermore, the *Academy of Management Review* has the highest JIF mean at 5.975; whereas the *Journal of Business ethics* has the lowest JIF mean at 1.258. The top and bottom places are the same for the other two alternative JIFs: the JIF adjusted for self-citations (JIF_adj) and the five-year average JIF (JIF_five). Interestingly, the JIF and JIF_adj of the *Journal of Management* have the largest variation through the years with standard deviations of 2.847 and 2.764, respectively. In contrast, the *Operations Research* journal has the smallest variation in JIF and JIF_adj with a standard deviation of 0.361 and 0.36, respectively. On the other hand, for the JIF_five, the *Journal of the Academy of Marketing Science* has the largest variation with a standard deviation of 3.411, while the *Review of Finance* has the lowest standard deviation of 0.295.

4.1.3. Publication data analysis

I find a total number of 1,435 scholars having publication records in the EBSCO database under the elimination line of a maximum 2,000 records. Despite the filtering mechanisms, errors still exist. [12] For example, *John Graham*, from *Duke University*, has research experience of 130 years, from the earliest publication in the year 1892 to the latest in the year 2021. This happens because *John Graham* is a common name, and I have over-collected publications from several different John-Grahams. Therefore, I need to trim off these insensible outliers. The histograms of the estimated publication experience from the publication records before matching to any journal lists, after matching to the JCR journal list and the FT50 journal list, suggest a clear turning point at the year 60 across. [13] Moreover, the 60-year research experience is a reasonable cut-off point. Some productive scholars have a life-long research experience. For example, Professor *Eugene F. Fama* published his first paper *The Behaviour of Stock Market Prices* in the *Journal of Business* in 1965. Until today, he is still an active scholar with nearly 60 years' publication experience.

Moreover, I count the occurrences of each journal from the FT50 journal list in the sample data. Then, I rank the journals based on the total numbers of their occurrences in a descending order. From *Table 2*, the total number of occurrences across all the journals are 10,364. Interestingly, the top three finance journals: the *Journal of Finance,* the *Journal of Financial Economics* and the *Review of Financial Studies* comfortably reside in the top three places on the occurrence list. The combined publications occurred in these three journals are over 50% of the total occurrences from all the 50 journals. [14] This result provides two implications. First, it shows the evidence to support the validity of my sample data, that is, finance scholars publishing their papers in finance journals. Second, it indicates the popularity of the three journals among finance scholars. In fact, Karolyi (2016) and Schwert (2021) document that the top three finance journals are still the most

[12] Please see *Appendix D* for the detailed description of the publication data collection.
[13] Please see *Appendix B* for the histograms.
[14] The percentage of the total occurrences of the top three finance journal is at 56.9% = (2,418+2,031+1,451)/10,364.

attractive journals for finance scholars given their premier quality and well-known reputation, despite that the rejection rates of these three journals have increased dramatically over the years.

4.2. Model fitting and main results

4.2.1. Descriptive statistics

After the explanatory data analyses, I merge the three datasets and calculate the productivity scores for each scholar within the sample. More specifically, I calculate four different groups of productivity scores by applying the quality-unweighted, JIF, JIF_adj and JIF_five. Additionally, by following Schwert's (2021) suggestion, I further divide each productivity score into two measures with and without taking account of co-authorship. The detailed descriptions of each variable are shown in *Glossary*.

From *Panel A* of *Table 3*, given the same group of productivity scores, the scores with co-authorship adjustment are always lower than the counterparts without the adjustment. The scores with the co-authorship adjustment consider only partial contributions to the scholar when a paper has multiple authors, whereas the counterparts without the adjustment consider full contributions for the same paper. Moreover, all the productivity scores are positive, which suggests a natural log transformation when fitting in a model.

Furthermore, the correlation matrix in *Panel B* of *Table 3* shows that there are significantly high correlations between all the productivity scores. The minimum correlation is 0.853 between JIF_adj_auth and N_coauth while the maximum correlation is 0.998 between JIF_auth and JIF_adj_auth. Particularly, the quality-adjusted productivity scores of JIF, JIF_adj and JIF_five are all highly correlated with each other at a minimum of 0.923. These indicate that the JIF of the journals is not sensitive to possible self-citations as well as short-term fluctuations. This provides suggestive evidence to support Fassin's (2021) conclusion that the FT50 list contains the stable top-tier journals. Interestingly, the quantity measures of the productivity scores, N_coauth and N_auth, are also highly correlated with the quality measures. There are two possible reasons. First, more publications mean more citations since JIF is citation-based quality measure. The second reason might be that scholars with high-quality productivity are more likely to publish papers in high-quality journals. In fact, Schwert (2021) points out that scholars with the past successful publication experience in a top-tier journal are most likely to get their papers published again in the same journal. Overall, given the high correlations between each productivity scores shown in the table, it is unsurprising that one productivity score's statistical significance will most likely imply the same to the others.

Table 3. The descriptive statistics and the correlation matrix of the productivity scores

Panel A shows the descriptive statistics of each productivity score and Panel B shows the correlation matrix between each productivity score. 25th_per represents the 25th percentile and 75th_per represents the 75th percentile. N_coauth and N_auth represent the number of publications of a scholar with and without co-authorship adjustment respectively; JIF_coauth and JIF_auth represent the scores weighted by JIF and with and without co-authorship adjustment respectively. Similarly, JIF_adj_coauth and JIF_adj_auth represent the scores weighted by JIF_adj and with and without co-authorship adjustment respectively; and JIF_five_coauth and JIF_five_auth represent the scores weighted by JIF_five and with and without co-authorship adjustment respectively.

Panel A. The descriptive statistics of each productivity score

	Mean	**SD**	**25th_per**	**Median**	**75th_per**	**Min**	**Max**
N_coauth	4.824	5.505	1.333	3.167	5.917	0.05	50.476
N_auth	9.925	10.879	3	7	13	1	86
JIF_coauth	13.792	14.553	4.142	9.787	17.962	0.107	122.31
JIF_auth	29.827	31.192	8.246	20.074	40.853	0.321	295.903
JIF_adj_coauth	12.199	12.824	3.56	8.686	15.843	0.088	116.835
JIF_adj_auth	26.581	27.868	7.326	17.79	35.954	0.131	282.583
JIF_five_coauth	22.908	24.981	6.929	15.774	29.148	0.092	218.774
JIF_five_auth	48.431	51.279	13.939	33.123	64.214	0.899	458.454

Panel B. The correlation matrix between each productivity score

	N_coauth	**N_auth**	**JIF_coauth**	**JIF_auth**	**JIF_adj_coauth**	**JIF_adj_auth**	**JIF_five_coauth**	**JIF_five_auth**
N_coauth	1							
N_auth	0.966	1						
JIF_coauth	0.937	0.937	1					
JIF_auth	0.876	0.939	0.965	1				
JIF_adj_coauth	0.916	0.922	0.997	0.967	1			
JIF_adj_auth	0.853	0.921	0.958	0.998	0.965	1		
JIF_five_coauth	0.972	0.955	0.986	0.937	0.975	0.923	1	
JIF_five_auth	0.920	0.970	0.968	0.989	0.963	0.980	0.964	1

4.2.2. Simple regression model fitting

Next, I start fitting a simple linear model. The model specification is:

$$\log(PRODUCTIVITY\_SCORE_i) = \alpha + \beta_1 * FWHR\_HIGH_i + \varepsilon_i \qquad [2]$$

where $\log(PRODUCTIVITY\_SCORE_i)$ is the natural log transformation of the productivity scores for scholar i, and $FWHR\_HIGH_i$ is a dummy variable equal to one if scholar i's fWHR is greater than the median and zero otherwise.

In the columns (1), (4), (7) and (10) in *Table 4*, the results for this regression that FWHR_HIGH coefficient is positive and statistically significant, which indicates a statistically significant positive relationship between FWHR_HIGH and log(PRODUCTIVITY_SCORE). The slope coefficients in the models (1) and (4) suggest an approximate 14.7% increase in productivity score in terms of quantity between the high fWHR scholars and the low fWHR scholars on average;[15] whereas the slope coefficients in the models (7) and (10) suggest an approximate 13.1% increase in productivity score in terms of quality.[16] However, the adjusted-R^2 value shows that all four models have an extremely low model goodness of fit with only around 0.3% variation of the dependent variables explained by the models. These lead me to fit models with more control variables.

4.2.3. Reduced linear regression model fitting

Before I add more control variables into the previous simple linear regression models, I de-mean the variable 1st_YEAR and EXP in order for having a meaningful intercept. More specifically, I deduct the year of the first publication of each scholar by 1998 and the publication experience of each scholar by 18 years.[17] The reduced regression specification is:

$$\log(PRODUCTIVITY\_SCORE_i) = \alpha + \beta_1 * FWHR\_HIGH_i + \beta_2 * MALE_i + \beta_3 * US_i + fixed\_effects(TITLE_i) + \beta_9 * 1st\_YEAR_i + \beta_{10} * EXP_i + \beta_{11} * (EXP_i)^2 + \varepsilon_i \qquad [3]$$

where $\log(PRODUCTIVITY\_SCORE_i)$ is the natural log transformation of the productivity scores for scholar i and $FWHR\_HIGH_i$ is a dummy variable equal to one if scholar i's fWHR is greater than the median and zero otherwise; $MALE_i$ is the gender dummy equal to one if scholar i is male and zero otherwise; US_i is equal to one for scholar i with the highest qualification attained in the US and zero otherwise; $fixed\_effects(TITLE_i)$ are the five academic title categorical variables: PROF, ASSO_PROF, ASSIS_PROF, SEN_LEC and OTHER, with the baseline variable PRAC; $1st\_YEAR_i$ represents the year of the very first publication for scholar

[15] 14.7% = (0.145+0.149)/2.
[16] 13.1% = (0.128+0.134)/2.
[17] 1998 and 18 are the means of the variables 1st_YEAR and EXP separately.

i; EXP_i is the publication experience of scholar i and $(EXP_i)^2$ represents the second order of the variable EXP_i.

The regression results in the columns (2), (5), (8) and (11) of *Table 4* show that the FWHR_HIGH dummy remains positive and statistically significant. Additionally, the PROF and ASSO_PROF variables have positive significant coefficients across all the models. This is institutive since I expect that the research productivity of a professor or associate professor is higher than that of an industry professional. The variable EXP has the highly significant positive coefficients for all the models. This is reasonable because a scholar with longer publication experience would be expected to have higher research productivity. The coefficients of the variable EXP^2 are in negative signs, which indicates a diminishing return of the publication experience on research productivity as expected. Furthermore, the adjusted-R^2 is improved remarkedly to around 50% for all the four models. It is noteworthy that the variable 1st_YEAR is highly significant but with positive signs as unexpected.

Table 4. The regression summaries of the simple, reduced, and full linear models

This table shows the regression results for the simple, reduced, and full linear models. Specifically, the models (1), (4), (7) and (10) are the simple linear regression models. The models (2), (5), (8) and (11) are the reduced linear regression models without the THE_RANK control variable, while the model (3), (6), (9) and (12) are the full linear regression models. Further, the dependent variable of the models (1), (2) and (3) is log(N_coauth); the dependent variable of the models (4), (5) and (6) is log(N_auth); the dependent variable of the model (8), (9) and (10) is log(JIF_coauth); and the dependent variable of the models (10), (11) and (12) is log(JIF_auth). Last, FWHR_HIGH is the fWHR dummy; MALE is the gender dummy; US is the US domestic highest qualification dummy; the TITLE variables are the academic title dummies; 1st_YEAR represents the year of the very first publication of a scholar; EXP is the publication experience of a scholar; EXP^2 represents the second order of the variable EXP; and THE_RANK represents the THE ranking for the corresponding institutions. The standard errors are HAC robust standard errors. *, ** and *** indicate the significances at 10%, 5% and 1% levels individually.

	log(N_coauth)			**log(N_auth)**			**log(JIF_coauth)**			**log(JIF_auth)**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
FWHR_HIGH	0.145**	0.125**	0.123**	0.149**	0.133***	0.131***	0.128*	0.136**	0.133**	0.134*	0.146***	0.144***
	(0.071)	(0.050)	(0.048)	(0.068)	(0.046)	(0.045)	(0.072)	(0.056)	(0.054)	(0.071)	(0.053)	(0.052)
MALE		0.020	-0.006		0.053	0.032		0.076	0.041		0.101	0.073
		(0.068)	(0.068)		(0.060)	(0.061)		(0.076)	(0.076)		(0.069)	(0.071)
US		0.141	0.137		0.056	0.053		0.200*	0.195*		0.121	0.117
		(0.096)	(0.094)		(0.084)	(0.083)		(0.105)	(0.101)		(0.093)	(0.090)
PROF		1.010***	0.998***		0.958***	0.949***		1.274***	1.259***		1.224***	1.212***
		(0.138)	(0.134)		(0.118)	(0.115)		(0.159)	(0.155)		(0.145)	(0.142)
ASSO_PROF		0.693***	0.753***		0.567***	0.613***		1.016***	1.093***		0.888***	0.951***
		(0.143)	(0.141)		(0.123)	(0.122)		(0.166)	(0.164)		(0.152)	(0.151)
ASSIS_PROF		0.074	0.151		-0.028	0.031		0.447***	0.546***		0.337**	0.418***
		(0.151)	(0.150)		(0.131)	(0.131)		(0.172)	(0.170)		(0.158)	(0.158)
SEN_LEC		-0.171	-0.306		-0.121	-0.224		-0.041	-0.214		-0.004	-0.146
		(0.252)	(0.240)		(0.202)	(0.201)		(0.279)	(0.265)		(0.268)	(0.266)
OTHER		-0.137	-0.174		-0.103	-0.132		-0.095	-0.142		-0.079	-0.118
		(0.244)	(0.254)		(0.208)	(0.218)		(0.276)	(0.291)		(0.239)	(0.252)
1st_YEAR		0.019***	0.016***		0.030***	0.027***		0.049***	0.045***		0.060***	0.057***
		(0.004)	(0.004)		(0.004)	(0.004)		(0.006)	(0.006)		(0.005)	(0.005)
EXP		21.961***	20.517***		24.082***	22.971***		28.069***	26.212***		30.650***	29.130***
		(1.978)	(1.984)		(1.729)	(1.761)		(2.470)	(2.466)		(2.288)	(2.310)
EXP^2		-1.811*	-2.053**		-2.998***	-3.184***		-0.753	-1.065		-2.064*	-2.319**
		(0.994)	(0.983)		(0.971)	(0.966)		(1.071)	(1.049)		(1.064)	(1.052)
THE_RANK			-0.012***			-0.009***			-0.015***			-0.012***
			(0.002)			(0.002)			(0.002)			(0.002)
Constant	0.975***	0.172	0.176	1.729***	1.046***	1.049***	2.070***	0.901***	0.906***	2.842***	1.794***	1.798***
	(0.049)	(0.169)	(0.164)	(0.047)	(0.142)	(0.139)	(0.049)	(0.191)	(0.186)	(0.049)	(0.167)	(0.165)
N	964	935	935	964	935	935	964	935	935	964	935	935
Adjusted R^2	0.003	0.529	0.554	0.004	0.547	0.564	0.002	0.429	0.470	0.003	0.458	0.486

4.2.4. Full linear regression model fitting

Finally, I fit a full model by adding an additional variable for controlling the institutional differences. Similarly, I de-mean the rank variable by deducting the mean rank 23. The model is specified in the equation [1] above.

The regression results in the columns (3), (6), (9) and (12) of *Table 4* confirm the significance of the fWHR_HIGH dummy. However, comparing to the results from the simple and reduced regressions, the magnitudes reduce slightly to around 12.7% for the quantity productivity measures: N_coauth and N_auth; whereas the magnitudes increase slightly to around 13.9% for the quality measures: JIF_coauth and JIF_auth.[18] These provide evidence that facial masculinity is positively associated with research productivity in terms of quality as well as quantity. This finding is consistent with the conclusions from many existing studies. For example, Carré and McCormick (2008) and Carré et al. (2009) document the positive relationship between fWHR and aggressive behaviour. Wong et al. (2011) find that firms led by more masculine-faced CEOs tend to have higher financial performance. Stirrat and Perrett (2012) argue that men with wider faces are more competitive than those with narrower faces. Moreover, many studies try to explain possible reasons for this association. For instance, both Campbell et al. (2010) and Jia et al. (2014) support that hormone testosterone influence is the main driver for males to have such masculine behaviours, which include impulsivity, aggressiveness, achievement-drive and maintenance of social status.

The MALE dummy is insignificant across all the models. This suggests that there is no statistically significant difference of the productivity in either quality or quantity between men and women scholars. In fact, the gender gap in research performance has been well-documented in many studies (e.g., McPherson et al., 2013; Sherman & Tookes, 2019; Ghosh & Liu, 2020). More specifically, Sherman and Tookes (2019) note that female finance scholars publish approximately 17.3% fewer papers than males in the US institutions. On the other hand, by using citations as a proxy for quality, Card et al. (2020) document that the papers from female authors receive 25% more citations than the similar papers written by male authors. Meanwhile, since the proportion of females in my sample is relatively small at around 17%, my MALE results are only suggestive and further investigation is required.

In addition, the US dummy is insignificant in all the models. This result indicates an indifference of the productivity between the scholars who obtained their highest qualifications in the US and the ones who earned theirs in a foreign country. This result contradicts Chan et al.'s (2009) and Chan et al.'s (2013) findings. Again, I recognise the small proportion of the scholars with the

[18] 12.7% = (0.123+0.131)/2 and 13.9% = (0.133+0.144)/2.

highest qualification granted outside the US in the sample. Thus, the result is only suggestive and further investigation is recommended.

The PROF and ASSO_PROF categorical variables remain highly significant across the full models. Interestingly, the ASSIS_PROF variable is only significant in the quality-adjusted productivity scores, which suggests that the scholars with an assistant professor title have only superior research quality but not quantity to the industry professionals. The coefficients of SEN_LEC and OTHER are insignificant across all the models, which suggests there is no differences in research productivity between the scholars with these two titles and the industry professionals. It is noteworthy that the descending order of the coefficients of the first three title levels shows the hierarchy of productivity in general: professors > associate professors > assistant professors > industry professionals.

The EXP variable is still highly positively significant, and the coefficient of the second order of EXP shows negative, which indicates a diminishing research productivity of academic experience. The new added variable THE_RANK is highly statistically significant with negative signs. Since the institution ranking is in an ascending order, the negative signs suggest an institution with a small ranking number (higher ranking) has higher productivity as originally expected.

Interestingly, 1st_YEAR is still highly significant with positive signs. This contradicts my conjecture that scholars who started publishing earlier would have higher performance due to the lower rejection rates of journals in earlier years. In fact, there are two possible reasons for this discrepancy. First, lower rejection rates do not necessarily mean that a paper would be more easily published. Although the rejection rates increase over time, the application volumes increase drastically as well (Card & DellaVigna, 2013; Schwert, 2021). This disproportional volume increase gives an illusion of the earlier papers getting published more easily. Second, the variable 1st_YEAR captures some of the model variation from the variables EXP and EXP^2. This is reasonable since the earlier the first publication is, the longer the experience is for an active scholar. Unsurprisingly, after I remove the 1st_YEAR variable from the models, the EXP^2 variable become highly significant in all the models while the other results remain consistent.[19] Finally, although the 1st_YEAR variable is statistically significant, it is economically insignificant with negligible magnitudes. Therefore, it does not raise my concerns.

[19] The results are not reported but available upon request.

4.3. Alternative model fitting[20]

4.3.1. Linear fixed effect model

Given that each institution has a different level of research productivity, which is reflected in its ranking measured by the ranking system, presumably scholars from different institutions will have different levels of productivity. In fact, Chen and Huang (2007) find that scholars affiliated with top ranked institutions tend to have higher performance scores. Moreover, different faculty size, per capita budget and catalyst effect of each institution have impact on the scholars within each institution (Chan et al., 2007). Naturally, one may consider using the INSITUTION categorical variable with 49 levels to control the institution effects. As a result, I fit a full model with the institution fixed effects. The specification is:

$$\log(PRODUCTIVITY\_SCORE_i) = \alpha + \beta_1 * FWHR\_HIGH_i + Control\ variables + fixed\_effects(INSITUTION_i) + \varepsilon_i \qquad [4]$$

where $\log(PRODUCTIVITY\_SCORE_i)$ is the natural log transformation of the productivity scores for scholar *i,* and $FWHR\_HIGH_i$ is a dummy variable equal to one if scholar i's fWHR is greater than the median and zero otherwise; the control variables are the aforementioned variables, which include $MALE_i$, US_i, the title categorical variables $TITLE_i$, $1st\_YEAR_i$, EXP_i and $(EXP_i)^2$; and $INSITUTION_i$ are the 49 institution categorical variables with *Boston University* as the baseline.

The results are presented in *Panel A* of *Table 5*. First of all, the FWHR_HIGH coefficients are still significant across all four models but slightly lower in magnitudes compared to the previous models with 10.8% for productivity quantity and 11.6% for quality measure, respectively. [21] The other variables PROF, ASSO_PROF, 1st_YEAR and EXP are still highly significant with consistent signs. EXP^2 is in negative signs which are consistent with previous results. The insignificant results for MALE and US dummies persist. Furthermore, the INSTITUTION dummies show only a minority of significant effects at the cost of 49 additional degrees of freedom being used. [22]

[20] Two additional model fitting: the OLS of the publication rate on fWHR and the linear mixed effect model are presented in *Appendix E*.
[21] 10.8% = (0.1+0.115)/2 and 11.6% = (0.108+0.124)/2.
[22] The results are not reported but available upon request.

Table 5. The regression summaries for the alternative models

This table shows the regression results of the alternative models. In particular, FWHR_HIGH is the fWHR dummy; MALE is the gender dummy; US is the US domestic highest qualification dummy; the TITLE variables are the academic title dummies; 1st_YEAR represents the year of the very first publication of a scholar; EXP is the publication experience of a scholar; EXP^2 represents the second order of the variable EXP; THE_RANK represents the THE ranking for the corresponding institutions; and INSTITUTION represents the institution categorical dummies. Moreover, Panel A shows the regression results of the ordinary least squares regression (OLS) models with the institution fixed effects. Panel B shows the regression results of the logistic regression with the THE ranking as the institution control variable. The standard errors are the HAC robust standard errors. *, ** and *** indicate the significances at 10%, 5% and 1% levels individually.

Panel A. Linear fixed effect model

	log(N_coauth) (1)	log(N_auth) (2)	log(JIF_coauth) (3)	log(JIF_auth) (4)
FWHR_HIGH	0.100**	0.115**	0.108**	0.124**
	(0.049)	(0.045)	(0.054)	(0.051)
MALE	-0.017	0.028	0.034	0.071
	(0.068)	(0.062)	(0.076)	(0.071)
US	0.095	0.015	0.135	0.061
	(0.090)	(0.080)	(0.100)	(0.090)
PROF	0.990***	0.935***	1.254***	1.202***
	(0.136)	(0.119)	(0.154)	(0.143)
ASSO_PROF	0.790***	0.634***	1.146***	0.988***
	(0.147)	(0.127)	(0.166)	(0.153)
ASSIS_PROF	0.162	0.029	0.569***	0.429***
	(0.159)	(0.138)	(0.178)	(0.164)
SEN_LEC	-0.374	-0.304	-0.255	-0.199
	(0.238)	(0.209)	(0.261)	(0.269)
OTHER	-0.232	-0.190	-0.190	-0.159
	(0.275)	(0.233)	(0.306)	(0.261)
1ST_YEAR	0.013***	0.023***	0.041***	0.053***
	(0.005)	(0.004)	(0.006)	(0.006)
EXP	19.586***	21.842***	25.106***	27.847***
	(2.133)	(1.878)	(2.621)	(2.415)
EXP^2	-2.265**	-3.427***	-1.364	-2.664**
	(0.971)	(0.958)	(1.049)	(1.049)
Constant	0.016	0.818***	0.524*	1.324***
	(0.224)	(0.228)	(0.287)	(0.315)
INSTITUTION (effects)	Fixed	Fixed	Fixed	Fixed
N	935	935	935	935
Adjusted R^2	0.573	0.584	0.496	0.515

Panel B. Logistic regression

	N_coauth_high (1)	N_auth_high (2)	JIF_coauth_high (3)	JIF_auth_high (4)
FWHR_HIGH	0.426**	0.510***	0.514***	0.539***
	(0.177)	(0.186)	(0.165)	(0.170)
MALE	-0.109	-0.128	0.013	0.128
	(0.252)	(0.257)	(0.233)	(0.231)
US	0.634**	0.512	0.724**	0.447
	(0.277)	(0.320)	(0.283)	(0.304)
PROF	2.375***	2.690***	2.374***	2.402***
	(0.393)	(0.388)	(0.413)	(0.416)
ASSO_PROF	1.307***	1.495***	2.021***	1.992***
	(0.436)	(0.427)	(0.449)	(0.447)
ASSIS_PROF	-0.254	-0.354	0.357	-0.004
	(0.523)	(0.525)	(0.504)	(0.512)
SEN_LEC	-0.530	-1.023	-0.893	-0.762
	(0.869)	(1.127)	(1.148)	(1.184)
OTHER	-0.212	-0.019	0.144	-0.358
	(0.787)	(0.834)	(0.794)	(0.891)
1ST_YEAR	0.072***	0.115***	0.145***	0.206***
	(0.017)	(0.024)	(0.027)	(0.032)
EXP	52.395***	71.915***	73.044***	95.129***

	(7.995)	(10.980)	(11.748)	(13.915)
EXP²	-8.948***	-15.481***	-5.291*	-9.565***
	(3.235)	(3.585)	(3.061)	(3.246)
THE_RANK	-0.032***	-0.029***	-0.035***	-0.031***
	(0.006)	(0.006)	(0.006)	(0.006)
Constant	-2.202***	-2.301***	-2.668***	-2.522***
	(0.511)	(0.525)	(0.499)	(0.506)
N	935	935	935	935
Akaike Inf. Crit.	855.033	778.758	940.628	895.557

4.3.2. Logistic regression

Often, researchers transform a continuous variable into a categorical variable before fitting in a model when doing quantitative analyses. This method is sometimes called "dichotomisation". There are several reasons for this transformation. First, a variable may have high variation and some of which is unwanted, such as, the variation from measurement errors. By dichotomising the variable, one could potentially reduce the measure error in the interest of the variable (Greene, 2000). Second, dichotomisation may help reveal the economic significance of a variable within a model (Jia et al., 2014). For these reasons, Jia et al. (2014) and Kamiya et al. (2019) transform the continuous variable FWHR into a categorical dummy with two levels. Likewise, I am tempted to fit a logistic regression by dichotomising the dependent variable – PRODUCTIVITY_SCORE. The specification is:

$$\text{logit}(PRODUCTIVITY\_SCORE\_HIGH_i) = \alpha + \beta_1 * FWHR\_HIGH_i + Control\ variables + \beta_{12} * THE\_RANK_i + \varepsilon_i \quad [5]$$

where $PRODUCTIVITY\_SCORE\_HIGH_i$ is the dummy variable transformed from the productivity scores, which is equal to one if scholar i's productivity score is greater than the median and zero otherwise; $FWHR\_HIGH_i$ is a dummy variable equal to one if scholar i's fWHR is greater than the median and zero otherwise; the control variables are the aforementioned variables, which include $MALE_i$, US_i, the title categorical variables $TITLE_i$, $1st\_YEAR_i$, EXP_i and $(EXP_i)^2$; and THE_RANK_i is the control variable for controlling the institutional differences.

The regression results are listed in *Panel B* of *Table 5.* First of all, the FWHR_HIGH dummy is still significant with positive signs in all the models. In the logistic regression context, the coefficients are interpreted as that the log-odds in high productivity quantity of a scholar with high fWHR is 0.47 higher than that of a scholar with low fWHR, and the log-odds in high productivity quality of a scholar with high fWHR is 0.53 higher than that of a scholar with low fWHR.[23] The statistical significance of the other variables is consistent with the results from the previous models.

[23] 0.47 = (0.426+0.51)/2 and 0.53 = (0.514+0.539)/2.

To sum up, I have fitted five different series of models. The results consistently show that there are positive associations between fWHR and productivity scores. Further, I find that the associations exist between fWHR and productivity, in terms of not only quantity, but also quality.

Chapter 5. Robustness tests

5.1. Journal impact factor alternatives

The journal impact factor was first mentioned by Dr *Eugene Garfield* in 1955 and then formally developed and published to help select additional source journals in the 1960s (Garfield, 1999, 2006). For a particular journal, the JIF is calculated as the total citations in the current year to the items published in the previous two years divided by the total number of citable items in the previous two years (Garfield, 2006). Although JIF is easy to calculate and provides an immediate and straightforward metric for evaluating the quality of journals and publications, many authors have raised several limitations of JIF. For example, Kurmis (2003) summarises several biases of JIF, and two of them are self-citations and the time lag of publication. The self-citation bias is quite intuitive. Since JIF is measured from the citations, many self-citations could potentially tilt the corresponding JIF of a paper upwards. Additionally, as JIF is a two-year average citation measure, it favours the papers that can be cited immediately or in near future. In this case, some publications in slow-moving fields may be disadvantaged. For instance, research in clinical fields often takes a long time to complete publication due to long experiment process, ethics approval and so on (Kurmis, 2003). As a result, a long publication time lag could potentially tilt the corresponding JIF of a paper downwards.

To address the two biases mentioned above, a JIF adjusted for self-citations (JIF_adj), and a five-year average JIF (JIF_five) are invented. As the names suggest, JIF_adj is calculated by deducting self-citations from the numerator of JIF and then dividing by the same denominator of JIF, whereas JIF_five is calculated as the total citations in the current year to the items published in the previous five years divided by the total number of citable items in the previous five years.[24] To test if my results are sensitive to the possible biases of JIF, I substitute JIF with JIF_adj and JIF_five successively to re-calculate the productivity scores for each scholar and then I re-fit the full models specified in [1] and [5].[25] The results are shown in *Panel A* of *Table 6*.

Evidently, the fWHR_HIGH coefficients are statistically significant with an economic magitude of around 14%, similar to the previous results from *Table 4*. In addition, the results of the other variables are consistent with the previous models. As a result, I conclude that my results are not sensitive to the different adjustments of JIF.

[24] For the detailed calculation formulas, please see http://jcr.help.clarivate.com/Content/home.htm.
[25] The logistic regression results are displayed in *Appendix B*.

Table 6. The OLS regression summaries on the different sub-samples

This table shows the summaries of the full OLS regressions on the different sub-samples. In particular, FWHR_HIGH is the fWHR dummy; MALE is the gender dummy; US is the US domestic highest qualification dummy; the TITLE variables are the academic title dummies; 1st_YEAR represents the year of the very first publication of a scholar; EXP is the publication experience of a scholar; EXP^2 represents the second order of the variable EXP; THE_RANK represents the THE ranking for the corresponding institutions. The standard errors are the HAC robust standard errors. *, ** and *** indicate the significances at 10%, 5% and 1% levels individually.

Panel A. OLS regressions for self-citation adjusted JIF and 5-year average JIF

The models (1) and (2) have the dependent variable PRODUCTIVITY_SCORE calculated from the JIF_adj, whereas the models (3) and (4) have the dependent variable PRODUCTIVITY_SCORE calculated from the JIF_five.

	log(JIF_adj_coauth)	**log(JIF_adj_auth)**	**log(JIF_five_coauth)**	**log(JIF_five_auth)**
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.136**	0.147***	0.124**	0.135***
	(0.055)	(0.053)	(0.052)	(0.050)
MALE	0.043	0.073	0.029	0.065
	(0.076)	(0.071)	(0.073)	(0.068)
US	0.185*	0.109	0.184*	0.105
	(0.101)	(0.091)	(0.100)	(0.089)
PROF	1.275***	1.229***	1.195***	1.144***
	(0.159)	(0.146)	(0.140)	(0.129)
ASSO_PROF	1.114***	0.972***	1.027***	0.885***
	(0.168)	(0.155)	(0.151)	(0.140)
ASSIS_PROF	0.574***	0.446***	0.487***	0.359**
	(0.174)	(0.162)	(0.160)	(0.149)
SEN_LEC	-0.250	-0.185	-0.476*	-0.397
	(0.309)	(0.330)	(0.273)	(0.264)
OTHER	-0.122	-0.100	-0.096	-0.067
	(0.296)	(0.256)	(0.296)	(0.253)
1ST_YEAR	0.049***	0.061***	0.033***	0.045***
	(0.006)	(0.006)	(0.005)	(0.005)
EXP	26.799***	29.777***	24.140***	26.837***
	(2.622)	(2.467)	(2.309)	(2.168)
EXP2	-0.804	-2.064*	-0.782	-2.050**
	(1.059)	(1.061)	(1.012)	(1.008)
THE_RANK	-0.016***	-0.013***	-0.015***	-0.012***
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.771***	1.666***	1.482***	2.366***
	(0.188)	(0.167)	(0.169)	(0.150)
N	935	935	935	935
Adjusted R^2	0.459	0.476	0.489	0.496

Panel B. OLS regressions for a sub-sample of male scholars only

It is noteworthy that the MALE dummies are left out of the models as this sub-sample has the males only.

	log(N_coauth)	**log(N_auth)**	**log(JIF_coauth)**	**log(JIF_auth)**
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.166***	0.172***	0.184***	0.189***
	(0.053)	(0.050)	(0.058)	(0.057)
US	0.140	0.057	0.181	0.108
	(0.109)	(0.096)	(0.117)	(0.106)
PROF	0.980***	0.965***	1.195***	1.176***
	(0.142)	(0.126)	(0.162)	(0.154)
ASSO_PROF	0.753***	0.639***	1.075***	0.949***
	(0.151)	(0.134)	(0.171)	(0.162)
ASSIS_PROF	0.193	0.081	0.588***	0.459***
	(0.164)	(0.146)	(0.179)	(0.170)
SEN_LEC	-0.337	-0.175	-0.262	-0.120
	(0.255)	(0.216)	(0.289)	(0.291)
OTHER	-0.064	0.038	-0.168	-0.088
	(0.278)	(0.223)	(0.333)	(0.287)
1ST_YEAR	0.014***	0.026***	0.041***	0.053***
	(0.005)	(0.004)	(0.006)	(0.005)
EXP	19.332***	21.279***	24.100***	26.477***
	(1.943)	(1.740)	(2.359)	(2.211)

EXP²	-2.263**	-3.077***	-1.407	-2.297**
	(0.990)	(0.973)	(1.042)	(1.046)
THE_RANK	-0.013***	-0.011***	-0.016***	-0.014***
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.194	1.083***	1.025***	1.932***
	(0.174)	(0.151)	(0.191)	(0.175)
N	793	793	793	793
Adjusted R²	0.559	0.562	0.470	0.479

Panel C. OLS regressions for a sub-sample of scholars with high-quality facial pictures

	log(N_coauth)	**log(N_auth)**	**log(JIF_coauth)**	**log(JIF_auth)**
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.108**	0.104**	0.119**	0.117**
	(0.054)	(0.051)	(0.059)	(0.058)
MALE	0.015	0.034	0.047	0.059
	(0.079)	(0.072)	(0.091)	(0.084)
US	0.284**	0.165*	0.331***	0.216**
	(0.111)	(0.098)	(0.118)	(0.106)
PROF	0.920***	0.887***	1.208***	1.178***
	(0.163)	(0.141)	(0.187)	(0.178)
ASSO_PROF	0.680***	0.555***	1.051***	0.922***
	(0.171)	(0.149)	(0.200)	(0.191)
ASSIS_PROF	0.084	-0.019	0.510**	0.398**
	(0.178)	(0.158)	(0.207)	(0.200)
SEN_LEC	-0.469*	-0.321	-0.341	-0.209
	(0.279)	(0.221)	(0.317)	(0.307)
OTHER	-0.421*	-0.347*	-0.389	-0.336
	(0.228)	(0.195)	(0.264)	(0.233)
1ST_YEAR	0.014***	0.024***	0.043***	0.053***
	(0.005)	(0.004)	(0.006)	(0.006)
EXP	17.892***	19.851***	22.680***	24.896***
	(1.980)	(1.736)	(2.475)	(2.257)
EXP²	-0.996	-1.982**	-0.097	-1.199
	(0.889)	(0.874)	(0.990)	(1.002)
THE_RANK	-0.011***	-0.008***	-0.014***	-0.011***
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.103	1.013***	0.825***	1.756***
	(0.197)	(0.167)	(0.224)	(0.204)
N	744	744	744	744
Adjusted R²	0.557	0.562	0.467	0.475

Panel D. OLS regressions for a sub-sample of authors with publications in the top three finance journals

The top three finance journals are JF, JFE and RFS. It is worth mention that all the productivity scores have been re-calculated based on the top three finance journals.

	log(N_coauth_top3)	**log(N_auth_top3)**	**log(JIF_coauth_top3)**	**log(JIF_auth_top3)**
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.154***	0.151***	0.161***	0.160***
	(0.057)	(0.053)	(0.060)	(0.057)
MALE	0.073	0.117*	0.106	0.146*
	(0.073)	(0.067)	(0.083)	(0.077)
US	0.117	0.058	0.133	0.075
	(0.096)	(0.087)	(0.102)	(0.094)
PROF	0.980***	0.894***	1.097***	1.024***
	(0.151)	(0.139)	(0.163)	(0.150)
ASSO_PROF	0.587***	0.440***	0.764***	0.629***
	(0.159)	(0.146)	(0.172)	(0.157)
ASSIS_PROF	-0.042	-0.198	0.197	0.040
	(0.167)	(0.153)	(0.180)	(0.162)
SEN_LEC	-0.188	-0.145	-0.138	-0.098
	(0.288)	(0.293)	(0.298)	(0.292)
OTHER	0.082	-0.144	0.002	-0.213
	(0.713)	(0.694)	(0.775)	(0.768)
1ST_YEAR	0.031***	0.040***	0.057***	0.067***
	(0.005)	(0.005)	(0.006)	(0.006)

EXP	17.996***	19.961***	22.692***	25.003***
	(2.166)	(2.160)	(2.358)	(2.346)
EXP²	-0.555	-1.717*	0.310	-0.911
	(1.055)	(1.033)	(1.084)	(1.082)
THE_RANK	-0.005***	-0.003	-0.007***	-0.004*
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	-0.182	0.718***	0.877***	1.780***
	(0.172)	(0.158)	(0.188)	(0.170)
N	781	781	781	781
Adjusted R²	0.378	0.406	0.295	0.331

Panel E. OLS regressions for a sub-sample of scholars with two-word names

	log(N_coauth)	log(N_auth)	log(JIF_coauth)	log(JIF_auth)
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.101*	0.109**	0.106*	0.111*
	(0.057)	(0.053)	(0.062)	(0.060)
MALE	-0.010	0.020	0.030	0.055
	(0.074)	(0.066)	(0.084)	(0.078)
US	0.217**	0.115	0.292***	0.198**
	(0.102)	(0.090)	(0.111)	(0.099)
PROF	0.996***	0.970***	1.307***	1.284***
	(0.162)	(0.137)	(0.184)	(0.168)
ASSO_PROF	0.681***	0.586***	1.083***	0.984***
	(0.166)	(0.142)	(0.193)	(0.177)
ASSIS_PROF	0.117	0.048	0.568***	0.489***
	(0.175)	(0.151)	(0.199)	(0.184)
SEN_LEC	-0.712**	-0.286	-0.391	0.006
	(0.328)	(0.231)	(0.369)	(0.312)
OTHER	-0.302	-0.198	-0.290	-0.208
	(0.258)	(0.224)	(0.295)	(0.259)
1ST_YEAR	0.018***	0.026***	0.047***	0.056***
	(0.006)	(0.005)	(0.008)	(0.007)
EXP	17.923***	19.296***	22.545***	24.231***
	(2.276)	(1.925)	(2.831)	(2.568)
EXP²	-1.958**	-3.302***	-1.418	-2.850**
	(0.984)	(0.976)	(1.088)	(1.112)
THE_RANK	-0.010***	-0.008***	-0.014***	-0.011***
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.075	0.933***	0.730***	1.607***
	(0.184)	(0.153)	(0.210)	(0.184)
N	702	702	702	702
Adjusted R²	0.561	0.574	0.485	0.499

Panel F. OLS regressions for a sub-sample of scholars with more-than-two-word names

	log(N_coauth)	log(N_auth)	log(JIF_coauth)	log(JIF_auth)
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.219**	0.241**	0.256**	0.295***
	(0.099)	(0.096)	(0.106)	(0.105)
MALE	0.069	0.137	0.137	0.188
	(0.152)	(0.148)	(0.168)	(0.155)
US	-0.139	-0.194	-0.243	-0.294
	(0.226)	(0.206)	(0.230)	(0.211)
PROF	1.078***	1.007***	1.149***	1.090***
	(0.248)	(0.237)	(0.277)	(0.258)
ASSO_PROF	1.078***	0.773***	1.030***	0.738***
	(0.285)	(0.260)	(0.314)	(0.282)
ASSIS_PROF	0.210	-0.018	0.420	0.206
	(0.332)	(0.305)	(0.349)	(0.317)
SEN_LEC	0.225	-0.113	-0.107	-0.426
	(0.290)	(0.373)	(0.362)	(0.448)
OTHER	0.402	0.236	0.520	0.360
	(0.743)	(0.674)	(0.757)	(0.694)
1ST_YEAR	0.008	0.025***	0.034***	0.053***
	(0.006)	(0.006)	(0.008)	(0.008)
EXP	8.930***	10.972***	11.580***	13.965***

	(1.289)	(1.404)	(1.535)	(1.536)
EXP²	-0.568	-0.722	0.250	0.017
	(0.958)	(0.950)	(0.950)	(0.920)
THE_RANK	-0.015***	-0.013***	-0.019***	-0.016***
	(0.003)	(0.003)	(0.004)	(0.004)
Constant	0.345	1.265***	1.418***	2.359***
	(0.342)	(0.331)	(0.363)	(0.344)
N	233	233	233	233
Adjusted R²	0.507	0.505	0.431	0.457

5.2. Males versus females

Since Olds and Shaver (1980) document that the association between masculinity and competitiveness only exists in males, many researchers afterwards find similar results. For instance, Carré and McCormick (2008) argue that the fWHR is a "reliable cue" of aggressive behaviour, and however, such implication is only valid for men but not women. They propose that fWHR is sexually dimorphic that men have higher ratios than women in general, and thus testosterone can explain the linkage between fWHR and aggressive behaviour. Many studies support this argument (e.g., Weston et al., 2007; Carré et al., 2009; Short et al., 2012), whereas others assert that fWHR is not a sexually dimorphic trait. For example, Lefevre et al. (2012) find no differences of fWHR between males and females after studying four large samples of adults. Özener (2012) believes that the fWHR's sexually dimorphic characteristic does not exist at least in the Turkish population. Furthermore, Gómez-Valdés et al. (2013) extend the sample to global population genetics and find no evidence supporting the fWHR sexual dimorphism. Given the agreements and disagreements, I have discussed earlier that facial masculinity is still a reliable predictor for masculine behaviours regardless of whether fWHR is sexually dimorphic or not. Unlike Jia et al. (2014) and Kamiya et al. (2019) examining the male CEOs only, I decide to incorporate the disagreements on fWHR sexually dimorphic characteristic by examining both male and female scholars in the sample.

As presented previously, the existence of females does not affect the distribution of the entire sample due to their low representation. Furthermore, I have shown in *Table 1* that fWHR is not different between male and female scholars. To ensure that my results are not sensitive to the possible gender differences, I re-fit the models for a sample with male scholars only. The specification is the same as the equation [1] above except for the MALE dummy.

The regression results are listed in *Panel B* of *Table 6*, which show that the FWHR_HIGH coefficients are highly significant at the 1% level in all the four models. Interestingly, the economic significance increases to 16.9% and 18.6% for productivity quantity and quality, respectively.[26] Similar to the previous results, the variables PROF and ASSO_PROF are still highly significant and the variables 1st_YEAR and EXP are also highly significant with positive

[26] 16.9% = (0.166+0.172)/2 and 18.6% = (0.184+0.189)/2.

signs, while the coefficients of the EXP^2 variable are still negative. Again, the THE_RANK variable is highly significant with negative signs. Similarly, I test the female scholars in the same fashion and find the FWHR_HIGH variable is not significant in all the models.[27] This seemingly suggests that the relationship between fWHR and research productivity is only found in males but not females, which coincides with the findings from prior studies (e.g., Carré & McCormick, 2008; Carré et al., 2009). However, this result is only suggestive due to the small sample size of female scholars and further investigation is required. Furthermore, I also test the logistic regressions [5] on the male sub-sample. The results are consistent and listed in *Appendix B*.

5.3. High-quality pictures

When Carré and McCormick (2008), Carré et al. (2009), and Jia et al. (2014) measure fWHR in their research, they use a Java-based image processing software called *ImageJ*, which is developed by the US National Institutes of Health. Additionally, they create three criteria to classify the photos into two categories: high- and low-quality pictures, which are based on picture resolution, unnatural expression, and face tilting. However, I decide to follow Kamiya et al.'s (2019) suggestion to use the Python artificial-intelligence (AI) program to measure fWHR from the scholar's photos; particularly, I use the *Face Recognition* module in Python 3. There are several reasons for my choice here. First, Kamiya et al. (2019) find that the Python AI measured fWHR produce an almost identical result to that from the *ImageJ* in terms of both statistical and economic significance (0.13 vs. 0.129 with both t-stats = 2.50). Second, the *Face Recognition* package can better address the possible measurement errors when encountering low resolution and face tilted picture. The algorithm will first turn a picture into a histogram of oriented gradients and identify the human faces within. Next, it rotates, scales, and shears the image so that the eyes and mouth are as centred as possible and afterwards a two-dimensions of face landmark estimation with a total of 68 points will be mapped.[28] Last, the fWHR of the person will be calculated based on these 68 landmarks.[29] As for the possible measurement error from unnatural expressions, Jia et al. (2014) explain that genuine smiles or laughing will move the face muscles and thus lower the facial height. However, I do not find any pictures having such expressions given that all the photos are directly from the institutions' websites. Presumably, they need to meet certain professional standards before the institutions allowing them to be published on the websites. Third, there are benefits from economies of scale for me using Python as well. I use Python for not only calculating fWHR but also web-scaping for scholars' information and JCR

[27] The regression results are listed in *Appendix B*.
[28] The mechanism is described in detail in one of the module creator *Adam Geitgey*'s blogs: https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78.
[29] Please see *Appendix A* for the detailed demonstration.

data from the websites. By using one programming language, it greatly reduces the learning cost, smooths the learning curve, and saves time.

Nevertheless, I distinguish the high-quality pictures from the low-quality pictures by applying the similar criteria suggested by Jia et al. (2014). To ensure that my results are not affected by the picture qualities, I fit a full model [1] on a sub-sample of the high-quality pictures only after excluding 191 low-quality picture data points. [30] The results are listed in *Panel C* of *Table 6*. Most importantly, the FWHR_HIGH dummy is still statistically significant across all four models but with slightly lower magnitudes at 10.6% and 11.8% for productivity quantity and quality separately; while the other variables: PROF, ASSO_PROF, 1st_YEAR, EXP and THE_RANK remain significant with the same corresponding signs as the results from *Table 4*. Interestingly, the US dummy becomes positive and statistically significant, which implies that the scholars with a US educational background have higher productivity than those with a foreign educational background; this result seems to support Chan et al.'s (2009) and Chan et al.'s (2013) findings. However, the results from the logistic models shown in *Appendix B* show that the US dummy is insignificant as previously reported in *Table 4*. Therefore, further investigation for the possible effect of US education is recommended. Moreover, the other results from the logistic regressions [5] are consistent and displayed in *Appendix B*. As a result, I conclude that my results are not sensitive to the picture quality differences.

5.4. Top three finance journals

The journal ranking approaches are commonly summarised into two board categories: a survey-based approach and a citation-based approach (Currie & Pandher, 2011; Chan et al., 2013). The FT50 journal list is created under the survey-based approach. However, the perceptions of journal quality from surveys are quite subjective, which are influenced by the respondents' personal factors, such as geographic origin, research interests and experience, and possible journal affiliations (Oltheten et al., 2005). On the other hand, the two most well-known journal ranking systems under citation-based approach are the Clarivate's journal ranking (JCR) powered by the Web of Science and the Scimago Journal & Country Rank (SJR) journal ranking collaborated with the Scopus. Moreover, many researchers develop their own journal ranking lists. For example, Currie and Pandher (2011) conduct a web-based survey among active finance scholars from 37 countries over all the world to rank 83 finance journals. Chen and Huang (2007) rank 41 finance journals by constructing an indicator called *Author Affiliation Index*. Unsurprisingly, the journal ranking from different sources can vary substantially. For instance, Currie and Pandher (2020) document that there are significant differences between the Clarivate's JCR rankings and

[30] 191 = 935 – 744.

Scimago's SJR rankings. However, despite all the possible variations, there are three finance journals stably remaining the top three positions in almost all the finance journal ranking lists: namely the *Journal of Finance*, the *Journal of Financial Economics* and the *Review of Financial Studies.*(e.g., Oltheten et al., 2005; Chen & Huang, 2007; Currie & Pandher, 2011; Chan et al., 2013)

In addition, as shown in *Table 2*, these three journals occupy the top three positions with more than 50% of the total publications in the sample. Naturally, by using these three journals as the journal sub-list, I test if the results are sensitive to the journal list selection. Consequently, I re-calculate the productivity scores for each scholar based on these three journals and re-fit the full model [1]. The results are shown in *Panel D* of *Table 6*. It can be seen that the FWHR_HIGH dummy is highly significant across all four models with slightly increased economic significance, and also the results for other variables remain consistent with the previous models. Moreover, I test the logistic regressions [5] on this sub-sample and find consistent results listed in *Appendix B*. As a result, I conclude that my results are not sensitive to the journal list selections.

5.5. Name length sensitivity

I have deployed three mechanisms to help minimise the possible measurement errors when collecting the publications from the EBSCO database. The three mechanisms are automation, publication names confirmation and journal matching.[31] Despite that, there might be still some extent of measurement errors existing in the sample. Particularly, some scholars with more than two-word names may not include their middle name(s) in their published papers, which results in wrong identification when searching by names. For instance, when searching for the name *An Li,* a total of 78,804 records shows up in EBSCO, whereas the result will reduce drastically to 177 records when searching for the name *An Yao Li* instead.

Thus, following this idea, I re-fit the models [1] on a sub-sample of scholars with two-word names scholars and a sub-sample of those with more-than-two-word names separately to see if the results are sensitive to the possible measurement errors. The regression results for these sub-samples are displayed in *Panels E* and *F* of *Table 6*, respectively. First, I notice that the entire sample has been split into two parts: 702 data points for the two-words-name scholars and 233 for the rest. In addition, the FWHR_HIGH dummy is significant in all eight models. However, the magnitudes of the FWHR_HIGH coefficients in the two-word-name scholar sample are less than half of those in the sample of scholars with longer names. Particularly, the two-word-name sample has the FWHR_HIGH coefficients of 10.5% for the productivity quantity at and 10.9% for the productivity quality, whereas the longer name sample has larger corresponding coefficients of

[31] Please see *Appendix D* for the detailed description.

23% and 27.6%. Combining these two sub-samples will even out the coefficients to the average value of 13% and 14% for quantity and quality individually. The significances of the other variables are consistent with the previous results. Furthermore, I test the logistic regressions [5] on the sub-samples. The results are consistent and listed in *Appendix B*.

Chapter 6. Discussion and conclusion

Considering the prior studies that facial structure is a reliable predictor for masculine behaviours (e.g., Carré & McCormick, 2008; Carré et al., 2009; Stirrat & Perrett, 2012; Jia et al., 2014), I examine the relationship between facial masculinity and academic research performance of the finance scholars from the top 50 US institutions on the THE World University Ranking 2021. By using fWHR as the proxy of masculine behaviours, I find scholars with high fWHR indeed exhibit masculine behaviour tendencies, and such behaviours promote them to publish more and higher-quality papers. In particular, the scholars with high fWHR outperform the ones with low fWHR by 13% and 14% in terms of the research productivity quantity and quality separately. I test different models, and the results are consistent. Additionally, my results are robust and survive a battery of sub-sample tests.

Moreover, I find that the association between facial masculinity and productivity exists only in men but not women. This result is consistent with the findings from the existing literature (e.g., Olds & Shaver, 1980; Carré & McCormick, 2008; Carré et al., 2009). However, I do not find evidence of a difference between the research productivity of male and female scholars, or any evidence that the scholars with the highest qualifications granted in the US have higher productivity than those with their highest qualifications obtained in non-US countries. Furthermore, unlike some literature which believes that fWHR is sexual dimorphic (e.g., Carré & McCormick, 2008; Carré et al., 2009), I do not find a statistical difference in fWHR between male and female scholars in my sample. In these cases, my results are subject to a caveat that I have relatively small samples for both female and foreign scholars. Further investigations are recommended.

Overall, my results lend support that fWHR is a reliable cue of masculine behaviours in not only laboratory, sport, corporate, and university but also high-educational academic contexts. However, my research does not aim to specifically identify underlying factors that drive high-fWHR scholars to publish more and higher-quality papers. He et al. (2019) argue that achievement drive is the main factor contributing to high performance, which is the most pronounced masculine behavioural trait implied by high-fWHR; while some researchers believe that aggressiveness, dominance, winning mentality and assertiveness masculine behavioural traits have an important impact on superior performance (Wong et al., 2011; Tsujimura & Banissy, 2013; Kausel et al., 2018), and others assert that highly competitive environment does also play a role (van Honk et al., 2011; Stirrat & Perrett, 2012). However, I take comfort in that no evidence shows that one behavioural trait dominates the others for promoting high performance, since fWHR is documented to have associations with a series of masculine behaviours, and these behaviours interconnect with each other. Here, I only speculate that all the positive traits implied by high-fWHR and the highly competitive publication environment jointly promote scholars to have superior performance compared to low-fWHR counterparts. In addition, the prominent biology and neuroscience literature conjectures that testosterone underlies the relationship between fWHR and masculine behaviours. In a laboratory environment, one could draw saliva and/or serum samples from participants for examining testosterone levels. However, I cannot test this conjecture since it is infeasible to take testosterone test samples on the scholars from the top 50 US institutions.

Finally, I cannot rule out the possible endogeneity and omitted variable issues in the models. Particularly, a scholar who publishes more and higher-quality papers more likely comes from a higher-ranked institution. Although the possible co-authorship with scholars from other institutions and/or a scholar's possible multi-institutional affiliations could alleviate this issue, I interpret my results with caveats. Meanwhile, it is very unlikely that fWHR is endogenous in the models. First, an individual's face shape is a biogenetic measure, which typically stops developing after puberty (e.g., Jia et al., 2014; Kamiya et al., 2019). Second, it is unlikely that either a person's academic research performance or publication level factors have an influence on his/her facial growth. Furthermore, prior studies document that journal affiliation may affect a scholar's performance, that is, a scholar who serves as an editor of a journal is more likely to get papers published in the journal (Schwert, 2021). This finding infers that an EDITOR dummy may be required for controlling such factor. However, I am unable to obtain this information in the sample. [32] In my view, these limitations open avenues for possible future studies.

[32] Although some of scholars' resumes do provide some information, the information is limited and scattered.

References

Altschul, D., Robinson, L., Coleman, K., Capitanio, J., & Wilson, V. (2019). An exploration of the relationships among facial dimensions, age, sex, dominance status, and personality in rhesus macaques (Macaca mulatta). *International journal of primatology, 40*(4), 532-552. https://doi.org/10.1007/s10764-019-00104-y

Anderl, C., Hahn, T., Schmidt, A.-K., Moldenhauer, H., Notebaert, K., Clément, C. C., & Windmann, S. (2016). Facial width-to-height ratio predicts psychopathic traits in males. *Personality and Individual Differences, 88*, 99-101. https://doi.org/10.1016/j.paid.2015.08.057

Apicella, C. L., Dreber, A., Campbell, B., Gray, P. B., Hoffman, M., & Little, A. C. (2008). Testosterone and financial risk preferences. *Evolution and human behavior, 29*(6), 384-390. https://doi.org/10.1016/j.evolhumbehav.2008.07.001

Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neuroscience & Biobehavioral Reviews, 30*(3), 319-345. https://doi.org/10.1016/j.neubiorev.2004.12.007

Arnocky, S., Carré, J. M., Bird, B. M., Moreau, B. J., Vaillancourt, T., Ortiz, T., & Marley, N. (2018). The facial width-to-height ratio predicts sex drive, sociosexuality, and intended infidelity. *Archives of Sexual Behavior, 47*(5), 1375-1385. https://doi.org/10.1007/s10508-017-1070-x

Arnold, T., Butler, A. W., Crack, T. F., & Altintig, A. (2003). Impact: what influences finance research? *The Journal of Business, 76*(2), 343-361. http://dx.doi.org/10.1086/367753

Bajo, E., Barbi, M., & Hillier, D. (2020). Where should I publish to get promoted? A finance journal ranking based on business school promotions. *Journal of Banking & Finance, 114*, 105780. https://doi.org/10.1016/j.jbankfin.2020.105780

Bentley, P. (2012). Gender differences and factors affecting publication productivity among Australian university academics. *Journal of Sociology, 48*(1), 85-103. https://doi.org/10.1177/1440783311411958

Bertrand, M., & Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *The Quarterly journal of economics, 118*(4), 1169-1208. https://doi.org/10.1162/003355303322552775

Besancenot, D., Huynh, K., & Serranito, F. (2017). Co-authorship and research productivity in economics: Assessing the assortative matching hypothesis. *Economic Modelling, 66*, 61-80. https://doi.org/10.1016/j.econmod.2017.05.018

Boksem, M. A., Mehta, P. H., Van den Bergh, B., van Son, V., Trautmann, S. T., Roelofs, K., Smidts, A., & Sanfey, A. G. (2013). Testosterone inhibits trust but promotes reciprocity. *Psychological science, 24*(11), 2306-2314. https://doi.org/10.1177/0956797613495063

Boustan, L., & Langan, A. (2019). Variation in Women's Success across PhD Programs in Economics. *Journal of Economic Perspectives, 33*(1), 23-42. https://doi.org/10.1257/jep.33.1.23

Campbell, B. C., Dreber, A., Apicella, C. L., Eisenberg, D. T., Gray, P. B., Little, A. C., Garcia, J. R., Zamore, R. S., & Lum, J. K. (2010). Testosterone exposure, dopaminergic reward, and sensation-seeking in young men. *Physiology & behavior, 99*(4), 451-456. https://doi.org/10.1016/j.physbeh.2009.12.011

Card, D., & DellaVigna, S. (2013). Nine facts about top journals in economics. *Journal of Economic Literature, 51*(1), 144-161. http://dx.doi.org/10.1257/jel.51.1.144

Card, D., DellaVigna, S., Funk, P., & Iriberri, N. (2020). Are referees and editors in economics gender neutral? *The Quarterly journal of economics, 135*(1), 269-327. https://doi.org/10.1093/qje/qjz035

Carré, J. M., & McCormick, C. M. (2008). In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B: Biological Sciences, 275*(1651), 2651-2656. https://doi.org/10.1098/rspb.2008.0873

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological science, 20*(10), 1194-1198. https://doi.org/10.1111/j.1467-9280.2009.02423.x
Chan, K. C., Chang, C.-H., & Chang, Y. (2013). Ranking of finance journals: Some Google Scholar citation perspectives. *Journal of Empirical Finance, 21*, 241-250. https://doi.org/10.1016/j.jempfin.2013.02.001
Chan, K. C., Chen, C. R., & Fung, H. G. (2009). Pedigree or placement? An analysis of research productivity in finance. *Financial Review, 44*(1), 87-111. https://doi.org/10.1111/j.1540-6288.2008.00211.x
Chan, K. C., Chen, C. R., & Lee, T. C. (2011). A long-term assessment of finance research performance among Asia–Pacific academic institutions (1990–2008). *Pacific-Basin Finance Journal, 19*(1), 157-171. https://doi.org/10.1016/j.pacfin.2010.09.006
Chan, K. C., Chen, C. R., & Lung, P. P. (2005). Ranking of finance programs in the Asia-Pacific region: An update. *Pacific-Basin Finance Journal, 13*(5), 584-600. https://doi.org/10.1016/j.pacfin.2005.05.001
Chan, K. C., Chen, C. R., & Lung, P. P. (2007). One-and-a-half decades of global research output in finance: 1990–2004. *Review of Quantitative Finance and Accounting, 28*(4), 417-439. http://dx.doi.org/10.1007/s11156-007-0018-y
Chan, K. C., Chen, C. R., & Steiner, T. L. (2001). Research productivity of the finance profession in the Asia-Pacific region. *Pacific-Basin Finance Journal, 9*(3), 265-280. https://doi.org/10.1016/S0927-538X(01)00011-7
Chan, K. C., Chen, C. R., & Steiner, T. L. (2002). Production in the finance literature, institutional reputation, and labor mobility in academia: A global perspective. *Financial Management*, 131-156. https://doi.org/10.2307/3666177
Chan, K. C., Chen, C. R., & Steiner, T. L. (2004). Who is publishing? An analysis of finance research productivity in the European region. *Journal of Business Finance & Accounting, 31*(3-4), 401-437. https://doi.org/10.1111/j.0306-686X.2004.00545.x
Chan, K. C., & Liano, K. (2009). Threshold citation analysis of influential articles, journals, institutions and researchers in accounting. *Accounting & Finance, 49*(1), 59-74. https://doi.org/10.1111/j.1467-629X.2007.00254.x
Chan, K. H., Wang, R. R., & Wang, R. (2020). The Macbeth Factor: The Dark Side of Achievement-driving Analysts. *Abacus*. https://doi.org/10.1111/abac.12187
Chen, C. R., & Huang, Y. (2007). Author Affiliation Index, finance journal ranking, and the pattern of authorship. *Journal of Corporate Finance, 13*(5), 1008-1026. https://doi.org/10.1016/j.jcorpfin.2007.04.011
Coates, J. M., Gurnell, M., & Rustichini, A. (2009). Second-to-fourth digit ratio predicts success among high-frequency financial traders. *Proceedings of the national academy of sciences, 106*(2), 623-628. https://doi.org/10.1073/pnas.0810907106
Cui, R., Ding, H., & Zhu, F. (2021). Gender inequality in research productivity during the COVID-19 pandemic. *Manufacturing & Service Operations Management*. http://www.ragcyt.org.ar/descargas/500_doc.pdf
Currie, R. R., & Pandher, G. S. (2011). Finance journal rankings and tiers: An active scholar assessment methodology. *Journal of Banking & Finance, 35*(1), 7-20. https://doi.org/10.1016/j.jbankfin.2010.07.034
Currie, R. R., & Pandher, G. S. (2020). Finance journal rankings: Active scholar assessment revisited. *Journal of Banking & Finance, 111*, 105717. https://doi.org/10.1016/j.jbankfin.2019.105717
Dabbs, J. M., & Dabbs, M. G. (2000). *Heroes, rogues, and lovers: Testosterone and behavior*. McGraw-Hill.
Eisenegger, C., Haushofer, J., & Fehr, E. (2011). The role of testosterone in social interaction. *Trends in cognitive sciences, 15*(6), 263-271. https://doi.org/10.1016/j.tics.2011.04.008
Eisenegger, C., Naef, M., Snozzi, R., Heinrichs, M., & Fehr, E. (2010). Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature, 463*(7279), 356-359. https://doi.org/10.1038/nature08711

Fassin, Y. (2021). Does the Financial Times FT50 journal list select the best management and economics journals? *Scientometrics*, 1-33. https://doi.org/10.1007/s11192-021-03988-x
Fishe, R. P. (1998). What are the research standards for full professor of finance? *The Journal of Finance, 53*(3), 1053-1079. https://doi.org/10.1111/0022-1082.00043
Fisk, S. R., Miller, B. J., & Overton, J. (2017). Why social status matters for understanding the interrelationships between testosterone, economic risk-taking, and gender. *Sociology Compass, 11*(3), e12452. https://doi.org/10.1111/soc4.12452
Fujii, T., Goto, A., & Takagishi, H. (2016). Does facial width-to-height ratio predict Japanese professional football players' athletic performance? *Letters on Evolutionary Behavioral Science, 7*(1), 37-40. https://doi.org/10.5178/lebs.2016.49
Garfield, E. (1999). Journal impact factor: a brief review. *Cmaj, 161*(8), 979-980. https://doi.org/10.1001/jama.295.1.90
Garfield, E. (2006). The history and meaning of the journal impact factor. *jama, 295*(1), 90-93. http://doi.org/10.29024/sar.31
Gelman, A. (2005). Analysis of variance—why it is more important than ever. *Annals of statistics, 33*(1), 1-53. https://doi.org/10.1214/009053604000001048
Geniole, S. N., Keyes, A. E., Carré, J. M., & McCormick, C. M. (2014). Fearless dominance mediates the relationship between the facial width-to-height ratio and willingness to cheat. *Personality and Individual Differences, 57*, 59-64. https://doi.org/10.1016/j.paid.2013.09.023
Geniole, S. N., Keyes, A. E., Mondloch, C. J., Carré, J. M., & McCormick, C. M. (2012). Facing aggression: Cues differ for female versus male faces. *Plos one, 7*(1), e30366. https://doi.org/10.1371/journal.pone.0030366
Ghosh, P., & Liu, Z. (2020). Coauthorship and the gender gap in top economics journal publications. *Applied Economics Letters, 27*(7), 580-590. https://doi.org/10.1080/13504851.2019.1644420
Gleason, E. D., Fuxjager, M. J., Oyegbile, T. O., & Marler, C. A. (2009). Testosterone release and social context: when it occurs and why. *Frontiers in neuroendocrinology, 30*(4), 460-469. https://doi.org/10.1016/j.yfrne.2009.04.009
Gómez-Valdés, J., Hünemeier, T., Quinto-Sánchez, M., Paschetta, C., de Azevedo, S., González, M. F., Martínez-Abadías, N., Esparza, M., Pucciarelli, H. M., & Salzano, F. M. (2013). Lack of support for the association between facial shape and aggression: A reappraisal based on a worldwide population genetics perspective. *Plos one, 8*(1), e52317. https://doi.org/10.1371/journal.pone.0052317
Greene, W. H. (2000). *Econometric analysis 4th edition*. Upper Saddle River, NJ: Prentice Hall.
Grossman, J. H., Reid, P. P., & Morgan, R. P. (2001). Contributions of academic research to industrial performance in five industry sectors. *The Journal of Technology Transfer, 26*(1), 143-152. https://doi.org/10.1023/A:1007848631448
Hahn, T., Winter, N. R., Anderl, C., Notebaert, K., Wuttke, A. M., Clément, C. C., & Windmann, S. (2017). Facial width-to-height ratio differs by social rank across organizations, countries, and value systems. *Plos one, 12*(11), e0187957. https://doi.org/10.1371/journal.pone.0187957
Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: facial structure predicts unethical behaviour. *Proceedings of the Royal Society B: Biological Sciences, 279*(1728), 571-576. https://doi.org/10.1098/rspb.2011.1193
Haselhuhn, M. P., Wong, E. M., Ormiston, M. E., Inesi, M. E., & Galinsky, A. D. (2014). Negotiating face-to-face: Men's facial structure predicts negotiation performance. *The Leadership Quarterly, 25*(5), 835-845. https://doi.org/10.1016/j.leaqua.2013.12.003
He, X., Yin, H., Zeng, Y., Zhang, H., & Zhao, H. (2019). Facial structure and achievement drive: Evidence from financial analysts. *Journal of Accounting Research, 57*(4), 1013-1057. https://doi.org/10.1111/1475-679X.12259
Hengel, E., & Moon, E. (2020). Gender and equality at top economics journals. Retrieved from https://erinhengel.github.io/Gender-Quality/quality.pdf
Hermans, E. J., Putman, P., Baas, J. M., Koppeschaar, H. P., & van Honk, J. (2006). A single administration of testosterone reduces fear-potentiated startle in humans. *Biological psychiatry, 59*(9), 872-874. https://doi.org/10.1016/j.biopsych.2005.11.015

Hoberman, J. (2005). *Testosterone dreams*. University of California Press.
Jaffe, A. B. (1989). Real effects of academic research. *The American economic review*, 957-970. Retrieved June 23, 2021, from http://www.jstor.org/stable/1831431
Jaffe, K., Ter Horst, E., Gunn, L. H., Zambrano, J. D., & Molina, G. (2020). A network analysis of research productivity by country, discipline, and wealth. *Plos one, 15*(5), e0232458. https://doi.org/10.1371/journal.pone.0232458
Jarnecic, E., Segara, R., Segara, L., & Westerholm, J. P. (2008). The scholarly output of universities and academics in the Asia-Pacific region who publish in major finance journals: 2000-2007. *Australasian Accounting, Business and Finance Journal, 2*(3), 26-49. http://dx.doi.org/10.14453/aabfj.v2i3.2
Jia, Y., LENT, L. V., & Zeng, Y. (2014). Masculinity, testosterone, and financial misreporting. *Journal of Accounting Research, 52*(5), 1195-1246. https://doi.org/10.1111/1475-679X.12065
Josephs, R. A., Newman, M. L., Brown, R. P., & Beer, J. M. (2003). Status, testosterone, and human intellectual performance: Stereotype threat as status concern. *Psychological science, 14*(2), 158-163. https://doi.org/10.1111/1467-9280.t01-1-01435
Josephs, R. A., Sellers, J. G., Newman, M. L., & Mehta, P. H. (2006). The mismatch effect: when testosterone and status are at odds. *Journal of personality and social psychology, 90*(6), 999. https://doi.org/10.1037/0022-3514.90.6.999
Kamiya, S., Kim, Y. H., & Park, S. (2019). The face of risk: CEO facial masculinity and firm risk. *European Financial Management, 25*(2), 239-270. https://doi.org/10.1111/eufm.12175
Karolyi, G. A. (2016). Home bias, an academic puzzle. *Review of Finance, 20*(6), 2049-2078. https://doi.org/10.1093/rof/rfw007
Kausel, E. E., Ventura, S., Datawheel, L., Díaz, D., & Vicencio, F. (2018). Does facial structure predict academic performance? *Personal Individ Differ, 129*, 1-5. https://doi.org/10.1016/j.paid.2018.02.041
Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological review, 110*(2), 265. https://doi.org/10.1037/0033-295X.110.2.265
Kerl, A., Miersch, E., & Walter, A. (2018). Evaluation of academic finance conferences. *Journal of Banking & Finance, 89*, 26-38. https://doi.org/10.1016/j.jbankfin.2018.01.014
Kurmis, A. P. (2003). Understanding the limitations of the journal impact factor. *JBJS, 85*(12), 2449-2454. https://doi.org/10.2106/00004623-200312000-00028
Lebuda, I., & Karwowski, M. (2016). Written on the Writer's Face: Facial Width-to-Height Ratio among Nominees and Laureates of the Nobel Prize in Literature. *Creativity Research Journal, 28*(2), 207-211. https://doi.org/10.1080/10400419.2016.1162572
Lefevre, C. E., Lewis, G. J., Bates, T. C., Dzhelyova, M., Coetzee, V., Deary, I. J., & Perrett, D. I. (2012). No evidence for sexual dimorphism of facial width-to-height ratio in four large adult samples. *Evolution and human behavior, 33*(6), 623-627. https://doi.org/10.1016/j.evolhumbehav.2012.03.002
Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. (2013). Telling facial metrics: facial width is associated with testosterone levels in men. *Evolution and human behavior, 34*(4), 273-279. https://doi.org/10.1016/j.evolhumbehav.2013.03.005
Lewis, G. J., Lefevre, C. E., & Bates, T. C. (2012). Facial width-to-height ratio predicts achievement drive in US presidents. *Personality and Individual Differences, 52*(7), 855-857. https://doi.org/10.1016/j.paid.2011.12.030
Long, J. S. (1992). Measures of sex differences in scientific productivity. *Social Forces, 71*(1), 159-178. https://doi.org/10.1093/sf/71.1.159
Lu, Y., & Teo, M. (2018). Do Alpha males deliver Alpha? Facial structure and hedge funds. *Research Collection Lee Kong Chian School Of Business.*, 1-61. https://doi.org/10.2139/ssrn.3100645
Lundberg, S., & Stearns, J. (2019). Women in economics: Stalled progress. *Journal of Economic Perspectives, 33*(1), 3-22. http://dx.doi.org/10.1257/jep.33.1.3
MacDonell, E. T., Geniole, S. N., & McCormick, C. M. (2018). Force versus fury: Sex differences in the relationships among physical and psychological threat potential, the

facial width-to-height ratio, and judgements of aggressiveness. *Aggressive behavior, 44*(5), 512-523. https://doi.org/10.1002/ab.21771
Mansfield, E. (1991). Academic research and industrial innovation. *Research policy, 20*(1), 1-12. https://doi.org/10.1016/0048-7333(91)90080-A
Mayew, W. J. (2013). Reassessing the association between facial structure and baseball performance: a comment on Tsujimura & Banissy (2013). *Biology letters, 9*(5), 20130538. https://doi.org/10.1098/rsbl.2013.0538
Mazur, A., & Booth, A. (1998). Testosterone and dominance in men. *Behavioral and brain sciences, 21*(3), 353 - 363. https://doi.org/10.1017/S0140525X98001228
McCormick, C. M., Mondloch, C. J., Carré, J. M., & Short, L. (2010). The facial width-to-height ratio as a basis for estimating aggression from emotionally neutral faces. *Journal of Vision, 10*(7), 599-599. https://doi.org/10.1167/10.7.599
McDowell, J. M., Singell Jr, L. D., & Stater, M. (2006). Two to tango? Gender differences in the decisions to publish and coauthor. *Economic inquiry, 44*(1), 153-168. https://doi.org/10.1093/ei/cbi065
McLean, R. D., & Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance, 71*(1), 5-32. https://doi.org/10.1111/jofi.12365
McPherson, M. A., Kim, M., Dorman, M., & Perera, N. (2013). Research output at US economics departments. *Applied Economics Letters, 20*(9), 889-892. https://doi.org/10.1080/13504851.2012.761331
Mehta, P. H., Jones, A. C., & Josephs, R. A. (2008). The social endocrinology of dominance: basal testosterone predicts cortisol changes and behavior following victory and defeat. *Journal of personality and social psychology, 94*(6), 1078. https://doi.org/10.1037/0022-3514.94.6.1078
Mileva, V. R., Cowan, M. L., Cobey, K. D., Knowles, K., & Little, A. (2014). In the face of dominance: Self-perceived and other-perceived dominance are positively associated with facial-width-to-height ratio in men. *Personality and Individual Differences, 69*, 115-118. https://doi.org/10.1016/j.paid.2014.05.019
Moosa, I. (2011). The demise of the ARC journal ranking scheme: an ex post analysis of the accounting and finance journals. *Accounting & Finance, 51*(3), 809-836. https://doi.org/10.1111/j.1467-629X.2011.00432.x
Nadler, A., Jiao, P., Johnson, C. J., Alexander, V., & Zak, P. J. (2018). The bull of wall street: Experimental analysis of testosterone and asset trading. *Management Science, 64*(9), 4032-4051. https://doi.org/10.1287/mnsc.2017.2836
Olds, D. E., & Shaver, P. (1980). Masculinity, femininity, academic performance, and health: Further evidence concerning the androgyny controversy 1. *Journal of Personality, 48*(3), 323-341. https://doi.org/10.1111/j.1467-6494.1980.tb00837.x
Oltheten, E., Theoharakis, V., & Travlos, N. G. (2005). Faculty perceptions and readership patterns of finance journals: A global view. *Journal of Financial and Quantitative Analysis, 40*(1), 223-239. https://doi.org/10.1017/S0022109000001800
Ormiston, M. E., Wong, E. M., & Haselhuhn, M. P. (2017). Facial-width-to-height ratio predicts perceptions of integrity in males. *Personality and Individual Differences, 105*, 40-42. https://doi.org/10.1016/j.paid.2016.09.017
Overbeck, J. R., & Park, B. (2001). When power does not corrupt: superior individuation processes among powerful perceivers. *Journal of personality and social psychology, 81*(4), 549. https://doi.org/10.1037/0022-3514.81.4.549
Özener, B. (2012). Facial width-to-height ratio in a Turkish population is not sexually dimorphic and is unrelated to aggressive behavior. *Evolution and human behavior, 33*(3), 169-173. https://doi.org/10.1016/j.evolhumbehav.2011.08.001
Pavel, A.-P. (2015). Global university rankings-a comparative analysis. *Procedia economics and finance, 26*, 54-63. https://doi.org/10.1016/S2212-5671(15)00838-2
Salter, A. J., & Martin, B. R. (2001). The economic benefits of publicly funded basic research: a critical review. *Research policy, 30*(3), 509-532. https://doi.org/10.1016/S0048-7333(00)00091-3

Schultheiss, O. C., Campbell, K. L., & McClelland, D. C. (1999). Implicit power motivation moderates men's testosterone responses to imagined and real dominance success. *Hormones and behavior, 36*(3), 234-241. https://doi.org/10.1006/hbeh.1999.1542

Schwert, G. W. (2021, 2021/06/01/). The remarkable growth in financial economics, 1974–2020. *Journal of Financial Economics, 140*(3), 1008-1046. https://doi.org/10.1016/j.jfineco.2021.03.010

Seibert, S. E., Kacmar, K. M., Kraimer, M. L., Downes, P. E., & Noble, D. (2017). The role of research strategies and professional networks in management scholars' productivity. *Journal of Management, 43*(4), 1103-1130. https://doi.org/10.1177/0149206314546196

Shamos, M. I. (2011). *Handbook of academic titles* http://euro.ecom.cmu.edu/people/faculty/mshamos/Handbook.pdf

Sherman, M. G., & Tookes, H. (2019). Female representation in the academic finance profession. *Journal of Finance. Forthcoming 2021*. http://dx.doi.org/10.2139/ssrn.3438653

Short, L. A., Mondloch, C. J., McCormick, C. M., Carré, J. M., Ma, R., Fu, G., & Lee, K. (2012). Detection of propensity for aggression based on facial structure irrespective of face race. *Evolution and human behavior, 33*(2), 121-129. https://doi.org/10.1016/j.evolhumbehav.2011.07.002

Smeets-Janssen, M. M., Roelofs, K., Van Pelt, J., Spinhoven, P., Zitman, F. G., Penninx, B. W., & Giltay, E. J. (2015). Salivary testosterone is consistently and positively associated with extraversion: results from The Netherlands study of depression and anxiety. *Neuropsychobiology, 71*(2), 76-84. ttps://doi.org/10.1159/000369024

Smurawa, T. M., & Congeni, J. A. (2007). Testosterone precursors: use and abuse in pediatric athletes. *Pediatric Clinics of North America, 54*(4), 787-796. https://doi.org/10.1016/j.pcl.2007.05.002

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological science, 21*(3), 349-354. https://doi.org/10.1177/0956797610362647

Stirrat, M., & Perrett, D. I. (2012). Face structure predicts cooperation: Men with wider faces are more generous to their in-group when out-group competition is salient. *Psychological science, 23*(7), 718-722. https://doi.org/10.1177/0956797611435133

Swidler, S., & Goldreyer, E. (1998). The value of a finance journal publication. *The Journal of Finance, 53*(1), 351-363. https://doi.org/10.1111/0022-1082.135230

Třebický, V., Havlíček, J., Roberts, S. C., Little, A. C., & Kleisner, K. (2013). Perceived aggressiveness predicts fighting performance in mixed-martial-arts fighters. *Psychological science, 24*(9), 1664-1672. https://doi.org/10.1177/0956797613477117

Tsujimura, H., & Banissy, M. J. (2013). Human face structure correlates with professional baseball performance: insights from professional Japanese baseball players. *Biology letters, 9*(3), 20130140. https://doi.org/10.1098/rsbl.2013.0140

van Honk, J., Terburg, D., & Bos, P. A. (2011). Further notes on testosterone as a social hormone. *Trends in cognitive sciences, 15*(7), 291-292. https://doi.org/10.1016/j.tics.2011.05.003

van Honk, J., Will, G.-J., Terburg, D., Raub, W., Eisenegger, C., & Buskens, V. (2016). Effects of testosterone administration on strategic gambling in poker play. *Scientific reports, 6*(1), 1-10. https://doi.org/10.1038/srep18096

Vanderschueren, D., Vandenput, L., Boonen, S., Lindberg, M. K., Bouillon, R., & Ohlsson, C. (2004). Androgens and bone. *Endocrine reviews, 25*(3), 389-425. https://doi.org/10.1210/er.2003-0003

Verdonck, A., Gaethofs, M., Carels, C., & de Zegher, F. (1999). Effect of low-dose testosterone treatment on craniofacial growth in boys with delayed puberty. *The European Journal of Orthodontics, 21*(2), 137-143. https://doi.org/10.1093/ejo/21.2.137

Wen, G., & Zheng, L. (2020). Facial width to height ratio predicts physical aggression in committed relationships in men and dominance in women in China. *Personality and Individual Differences, 157*, 109832. https://doi.org/10.1016/j.paid.2020.109832

Weston, E. M., Friday, A. E., & Liò, P. (2007). Biometric evidence that sexual selection has shaped the hominin face. *Plos one, 2*(8), e710. https://doi.org/10.1371/journal.pone.0000710

Wong, E. M., Ormiston, M. E., & Haselhuhn, M. P. (2011). A face only an investor could love: CEOs' facial structure predicts their firms' financial performance. *Psychological science, 22*(12), 1478-1483. https://doi.org/10.1177/0956797611418838

Wood, R. I., & Stanton, S. J. (2012). Testosterone and sport: current perspectives. *Hormones and behavior, 61*(1), 147-155. https://doi.org/10.1016/j.yhbeh.2011.09.010

Worthman, C. M., & Konner, M. J. (1987). Testosterone levels change with subsistence hunting effort in! Kung San men. *Psychoneuroendocrinology, 12*(6), 449-458. https://doi.org/10.1016/0306-4530(87)90079-5

Wright, N. D., Bahrami, B., Johnson, E., Di Malta, G., Rees, G., Frith, C. D., & Dolan, R. J. (2012). Testosterone disrupts human collaboration by increasing egocentric choices. *Proceedings of the Royal Society B: Biological Sciences, 279*(1736), 2275-2280. https://doi.org/10.1098/rspb.2011.2523

Glossary

Table 7. The variable definitions

This table shows the detailed calculation and definition of each variable in the models.

Variable Name	Formula	Description
N_coauth	$N\_coauth_j = \sum \frac{1}{N\_coauth_i}$	N_coauth is the total sum of one divided by the number of authors of a publication for scholar j. The publications are based on the FT50 journals.
N_coauth_high	$N\_coauth\_high_j = \begin{cases} 1, if\ N\_coauth_j \geq median \\ 0, \quad otherwise \end{cases}$	N_coauth_high is the N_coauth dummy equal to one if greater than the median and zero otherwise for scholar j. The publications are based on the FT50 journals.
N_auth	$N\_auth_j = \sum 1_i$	N_auth is the total counts of the publications of scholar j. The publications are based on the FT50 journals.
N_auth_high	$N\_auth\_high_j = \begin{cases} 1, if\ N\_auth_j \geq median \\ 0, \quad otherwise \end{cases}$	N_auth_high is the N_auth dummy equal to one if greater than the median and zero otherwise for scholar j. The publications are based on the FT50 journals.
JIF_coauth	$JIF\_coauth_j = \sum \frac{JIF_i}{N\_coauth_i}$	JIF_coauth is the total sum of the corresponding JIF of a journal at a given year divided by the number of authors of that publication for scholar j. The publications are based on the FT50 journals.
JIF_coauth _high	$JIF\_coauth\_high_j = \begin{cases} 1, if\ JIF\_coauth_j \geq median \\ 0, \quad otherwise \end{cases}$	JIF_coauth _high is the JIF_coauth dummy equal to one if greater than the median and zero otherwise for scholar j. The publications are based on the FT50 journals.
JIF_auth	$JIF\_auth_j = \sum JIF_i$	JIF_auth is the total sum of the corresponding JIF of a journal at a given year when a paper published for scholar j. The publications are based on the FT50 journals.
JIF_auth _high	$JIF\_auth\_high_j = \begin{cases} 1, if\ JIF\_auth_j \geq median \\ 0, \quad otherwise \end{cases}$	JIF_auth _high is the JIF_auth dummy equal to one if greater than the median and zero otherwise for scholar j. The publications are based on the FT50 journals.
JIF_adj_coauth	$\text{JIF\_adj\_coauth}_j = \sum \frac{JIF\_adj_i}{N\_coauth_i}$	JIF_adj_coauth is the total sum of the corresponding JIF_adj of a journal at a given year divided by the number of authors of a publication for scholar j. The publications are based on the FT50 journals.
JIF_adj_coauth _high	$JIF\_adj\_coauth\ \_high_j = \begin{cases} 1, if\ JIF\_adj\_coauth_j \geq median \\ 0, \quad otherwise \end{cases}$	JIF_adj_coauth _high is the JIF_adj_coauth dummy equal to one if greater than the median and zero otherwise for scholar j. The publications are based on the FT50 journals.
JIF_adj_auth	$JIF\_adj\_auth_j = \sum JIF\_adj_i$	JIF_adj_auth is the total sum of the corresponding JIF_adj of a journal at a given year when a paper published for scholar j. The publications are based on the FT50 journals.

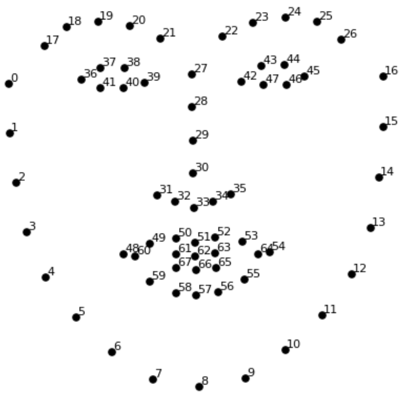
JIF_adj_auth _high	$JIF\_adj\_auth\ \_high_j = \begin{cases} 1, if\ JIF\_adj\_auth_j \geq median \\ 0, \quad otherwise \end{cases}$	JIF_adj_auth _high is the JIF_adj_auth dummy equal to one if greater than the median and zero otherwise for scholar j. The publications are based on the FT50 journals.
JIF_five_coauth	$JIF\_five\_coauth_j = \sum \frac{JIF\_five_i}{N\_coauth_i}$	JIF_five_coauth is the total sum of the corresponding JIF_five of a journal at a given year divided by the number of authors of a publication for scholar j. The publications are based on the FT50 journals.
JIF_five_coauth _high	$JIF\_five\_coauth\ \_high_j = \begin{cases} 1, if\ JIF\_five\_coauth_j \geq median \\ 0, \quad otherwise \end{cases}$	JIF_five_coauth _high is the JIF_five_coauth dummy equal to one if greater than the median and zero otherwise for scholar j. The publications are based on the FT50 journals.
JIF_five_auth	$JIF\_five\_auth_j = \sum JIF\_five_i$	JIF_five_auth is the total sum of the corresponding JIF_five of a journal at a given year when a paper published for scholar j. The publications are based on the FT50 journals.
JIF_five_auth _high	$JIF\_five\_auth\ \_high_j = \begin{cases} 1, if\ JIF\_five\_auth_j \geq median \\ 0, \quad otherwise \end{cases}$	JIF_five_auth _high is the JIF_five_auth dummy equal to one if greater than the median and zero otherwise for scholar j. The publications are based on the FT50 journals.
N_coauth_top3	$N\_coauth\_top3_j = \sum \frac{1}{N\_coauth_i}$	N_coauth_top3 is the total sum of one divided by the number of authors of a publication for scholar j. The publications are based on the top three finance journals.
N_coauth_top3_high	$N\_coauth\_top3\_high_j = \begin{cases} 1, if\ N\_coauth\_top3_j \geq median \\ 0, \quad otherwise \end{cases}$	N_coauth_top3_high is the N_coauth_top3 dummy equal to one if greater than the median and zero otherwise for scholar j. The publications are based on the top three finance journals.
N_auth_top3	$N\_auth\_top3_j = \sum 1_i$	N_auth_top3 is the total counts of the publications of scholar j. The publications are based on the top three finance journals.
N_auth_top3_high	$N\_auth\_top3\_high_j = \begin{cases} 1, if\ N\_auth\_top3_j \geq median \\ 0, \quad otherwise \end{cases}$	N_auth_top3_high is the N_auth_top3 dummy equal to one if greater than the median and zero otherwise for scholar j. The publications are based on the top three finance journals.
JIF_coauth_top3	$JIF\_coauth\_top3_j = \sum \frac{JIF_i}{N\_coauth_i}$	JIF_coauth_top3 is the total sum of the corresponding JIF of a journal at a given year divided by the number of authors of that publication for scholar j. The publications are based on the top three finance journals.
JIF_coauth _top3_high	$JIF\_coauth\_top3\_high_j = \begin{cases} 1, if\ JIF\_coauth\_top3_j \geq median \\ 0, \quad otherwise \end{cases}$	JIF_coauth _top3_high is the JIF_coauth_top3 dummy equal to one if greater than the median and zero otherwise for a particular scholar j. The publications are based on the top three finance journals.
JIF_auth_top3	$JIF\_auth\_top3_j = \sum JIF\_top3_i$	JIF_auth_top3 is the total sum of the corresponding JIF of a journal at a given year when a paper published for scholar j. The publications are based on the top three finance journals.
JIF_auth _top3_high	$JIF\_auth\_top3\_high_j = \begin{cases} 1, if\ JIF\_auth\_top3_j \geq median \\ 0, \quad otherwise \end{cases}$	JIF_coauth _top3_high is the JIF_coauth_top3 dummy equal to one if greater than the median and zero otherwise for scholar j. The publications are based on the top three finance journals.
FWHR_HIGH	$FWHR\_HIGH_j = \begin{cases} 1, if\ FWHR_j \geq median \\ 0, \quad otherwise \end{cases}$	FWHR_HIGH is the fWHR dummy equal to one if scholar j's fWHR is greater than the median and zero otherwise for a particular scholar j.
MALE	$MALE_j = \begin{cases} 1, \quad if\ male \\ 0, otherwise \end{cases}$	MALE is the gender dummy equal to one if male and zero if female for scholar j.

US	$US_j = \begin{cases} 1, & if\ U.S. \\ 0, & otherwise \end{cases}$	US is the US domestic highest qualification dummy equal to one if the highest qualification attained in the US and zero otherwise for scholar j.
TITLE	$PROF_j = \begin{cases} 1, if\ professor \\ 0, if\ otherwise \end{cases}$ and $ASSO\_PROF_j = \begin{cases} 1, if\ associate\ professor \\ 0, \quad if\ otherwise \end{cases}$ and $ASSIS\_PROF_j = \begin{cases} 1, if\ assistant\ professor \\ 0, \quad if\ otherwise \end{cases}$ and $SEC\_LEC_j = \begin{cases} 1, if\ senior\ lecturer \\ 0, \quad if\ otherwise \end{cases}$ and $OTHER_j = \begin{cases} 1, \quad if\ other \\ 0, if\ otherwise \end{cases}$ and The baseline dummy is $PRAC_j$.	TITLE is the academic title categorical variable with six levels: PROF, ASSO_PROF, ASSIS_PROF, SEC_LEC, OTHER and the baseline dummy PRAC.
1st_YEAR		1st_YEAR is the continuous variable which represents the very first publication year of scholar j. It has been de-meaned by deducting the mean 1998 before fitting in the models.
EXP	$EXP_j = LATEST\_YEAR_j - 1st\_YEAR_j + 1$	EXP is the continuous variable which represents the publication experience of scholar j. Also, LATEST_YEAR represents the latest publication year of scholar j in the formula. The variable EXP has been de-meaned by deducting the mean 18 before fitting in the models.
EXP^2	${EXP_j}^2 = (EXP_j)^2$	EXP^2 is the quadratic term of the variable EXP, which intends to capture the non-linear relationship between productivity and experience.
THE_RANK		THE_RANK is the continuous variable which represents the institution ranking from the THE ranking 2021 where scholar j is affiliated with. It has been de-meaned by deducting the mean rank 23 before fitting in the models. *Boston University* is ranked at 23.

Appendix A. Facial width-to-height ratio calculation

I developed a Python program based on the *Face Recognition* module for calculating the fWHR of the scholars in the sample.[33] This package is one of the prevailing modules for any face recognition tasks. According to the author *Adam Geitgey*, it has an accuracy of 99.38% pair matching face verification from the *Labelled Faces in the Wild* benchmark.[34]

The mechanism of how the fWHR is calculated from a scholar's photo is described as follows. First, the program de-colours a picture into grey scales, as the colour data is redundant for the face detection. Next, it transforms the black and white image into a histogram of oriented gradients which is a method invented in 2005. Afterwards, it searches and identifies any faces within the image.[35] An error returns if no human faces are detected. Once a face has been found, the program will rotate, scale, and shear the face so that the eyes and mouth are as centred as possible followed by a two-dimension (2D) map with 68 specific landmarks projected. The projected landmarks of a typical face are illustrated in *Figure 2* below. Last, the fWHR of the face is calculated based on these landmarks.

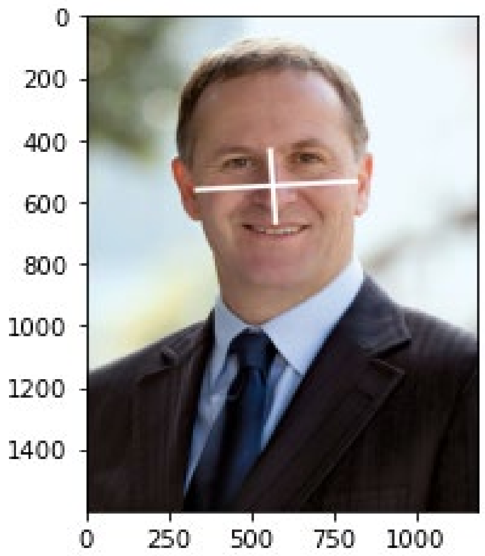


This plot shows the 68 specific landmarks mapped from a typical human face. The points 0 to 16 represent the chin; the points 17 to 21 represent the left eyebrow; the points 22 to 26 represent the right eyebrow; the points 27 to 30 represent the nose bridge; the points 31 to 35 represent the nose tip; the points 36 to 41 represent the left eye; the points 42 to 47 represent the right eye; the points 48 to 54 and 60 to 64 represent the upper lip; and the points 48 and 54 to 60 and 64 to 67 represent the lower lip.

Figure 2. The 2D landmark map of a typical face[36]

[33] This source code is from https://notebook.community/TiesdeKok/fWHR_calculator/FWHR_calculator. I make modifications in order to match Kamiya et al.'s (2019) fWHR measuring method.
[34] The source is from https://pypi.org/project/face-recognition/.
[35] The program allows identifying multiple human faces in a picture. However, this functionality has no use to me since almost all the photos in my sample contain only one face.
[36] The plot is from one of the author's blogs. https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78.

By following Kamiya et al.'s (2019) method, the facial width is measured as the distance between the highest points of the left and right zygomatic bone, which is the Euclidean distance between the points 1 and 15 in the graph above; whereas the facial height is measured as the distance between the midpoints of two inner ends of the eyebrows and the upper lip, that is the Euclidean distance between the middle point of the points 21 and 22 and the point 51. Consequently, the corresponding fWHR is calculated as the facial width divided by the height. A sample photo of the 38th Prime Minister of New Zealand, Sir *John Key*, illustrates the calculation method in *Figure 3* below.



This photo illustrates the fWHR calculation by following Kamiya et al.'s (2019) method. The horizontal line indicates the facial width, while the vertical line shows the facial height. As a result, the corresponding fWHR is calculated by using the width dividing by the height. At the end, Sir *John Key's* fWHR calculated from the photo is 2.002.

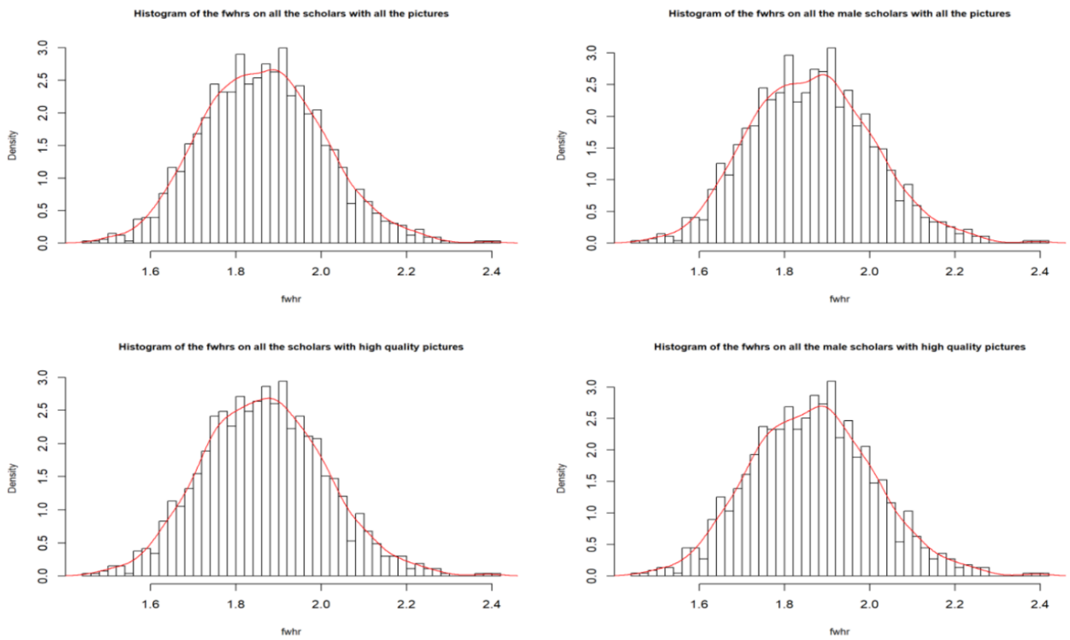
Figure 3. The fWHR calculation on a sample photo of Sir *John Key*

Moreover, I apply four criteria to distinguish the high-quality pictures from the low-quality pictures. The first criterion is to use the difference between the heights of the two eyes divided by the facial width. If this ratio is greater than a conventional number 5/100, the picture is classified as low-quality. The second criterion is to use the width between the top point and bottom point of the nose divided by the facial width. If this ratio is greater than a conventional number 3.5/100, then I categorise this picture as low-quality. The third criterion is to calculate the ratio of the distance between the left eye and the left face edge and the distance between the right eye and the right face edge. Again, if this ratio is greater than a conventional number 3/100, then the picture is defined as low-quality.[37] Finally, I set a criterion for the picture resolution. By following

[37] These conventional numbers are suggested by users on the Github. Please see https://notebook.community/TiesdeKok/fWHR_calculator/FWHR_calculator.

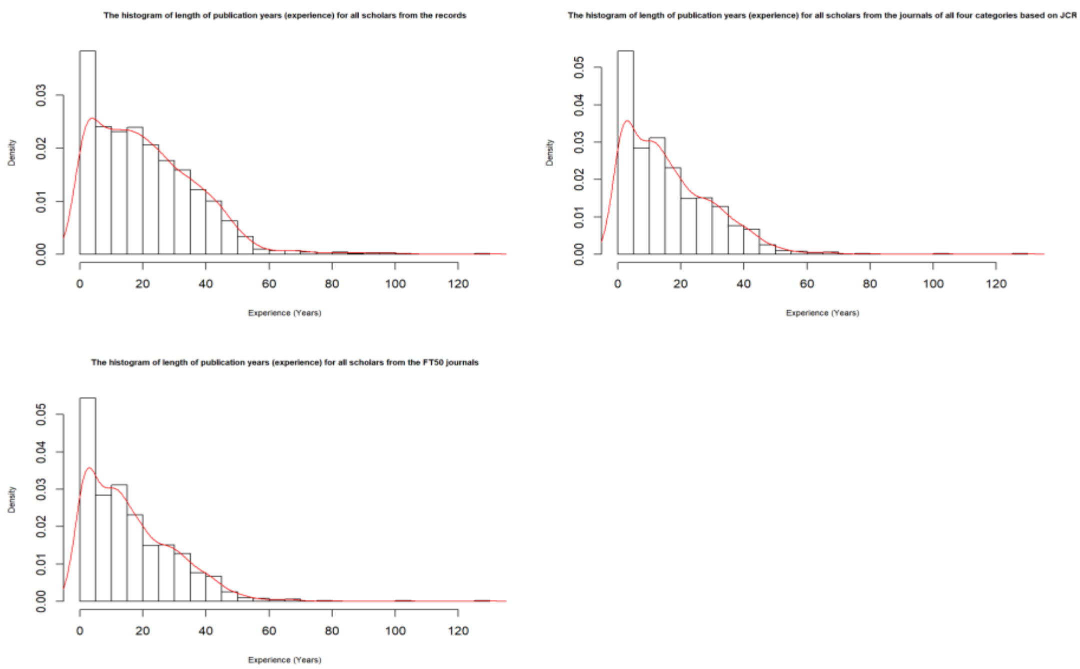
Kamiya et al.'s (2019) suggestion, if a picture has a resolution lower than 120×150 pixels, then the picture is low-quality.

Appendix B. Additional results



The top left is the histogram of fWHR of all the scholars with a facial picture. The bottom left is the histogram of fWHR of all the scholars with a high-quality facial picture; while the top right is the histogram of fWHR of all the male scholars with a facial picture and the bottom right is the histogram of fWHR of all the male scholars with a high-quality facial picture. From the side-by-side comparisons, evidently, the distributions are approximately the same between all scholars and male scholars only due to the small percentages of females.

Figure 4. The histograms of the fWHR on different quality picture samples



This figure shows the three histograms of the estimated publication experience in years from three different sample data. The top-left plot shows the histogram of the estimated publication experience from the publication records before matching any journal lists. The top-right plot shows the histogram of the estimated publication experience from the publication records after matching the JCR journal list, which is created from combining all the journal lists from the four categories: finance, business, economics, and management. The bottom plot shows the histogram of the estimated publication experience from the publication records after matching the FT50 journal list. Evidently, the point of 60 years is the turning points in common for all three histograms.

Figure 5. The histograms of publication experience on different samples

Table 8. The descriptive statistics of the fWHR for each institution

This table shows the descriptive statistics for each institution. The left-most column of the table is the THE 2021 US regional institution rankings of accounting and finance subjects. The middle part of the table is the descriptive statistics of fWHR of the scholars with a face picture for each institution. Particularly, THE_RANK represents the THE ranking. N represents the total number of scholars with an available fWHR for each institution. 25th represents the 25th percentile and 75th represents the 75th percentile.

THE_RANK	Institution	N	Mean	SD	25th	Median	75th	Min	Max
1	Stanford University	42	1.625	0.135	1.805	1.920	2.000	1.625	1.625
2	Harvard University	52	1.620	0.150	1.743	1.862	1.971	1.620	1.620
3	Massachusetts Institute of Technology	45	1.593	0.141	1.721	1.829	1.938	1.593	1.593
4	Yale University	21	1.600	0.113	1.804	1.877	1.913	1.600	1.600
5	The University of Chicago	40	1.564	0.127	1.792	1.881	1.946	1.564	1.564
6	Johns Hopkins University	30	1.599	0.157	1.773	1.845	1.939	1.599	1.599
7	University of Pennsylvania	60	1.608	0.151	1.776	1.899	1.964	1.608	1.608
8	University of California, Los Angeles	36	1.445	0.146	1.727	1.851	1.942	1.445	1.445
9	Columbia University	64	1.569	0.132	1.798	1.868	1.923	1.569	1.569
10	Duke University	17	1.479	0.139	1.759	1.887	1.950	1.479	1.479
11	University of Michigan-Ann Arbor	28	1.643	0.136	1.853	1.914	1.966	1.643	1.643
12	Northwestern University	36	1.516	0.150	1.752	1.844	1.953	1.516	1.516
13	New York University	98	1.624	0.132	1.747	1.861	1.941	1.624	1.624
14	Carnegie Mellon University	19	1.648	0.163	1.808	1.914	1.983	1.648	1.648
15	University of Washington	28	1.569	0.145	1.743	1.823	1.907	1.569	1.569
16	University of California, San Diego	14	1.579	0.190	1.773	1.848	2.052	1.579	1.579
17	Georgia Institute of Technology	15	1.719	0.108	1.812	1.842	1.970	1.719	1.719
18	University of Texas at Austin	86	1.565	0.149	1.755	1.826	1.939	1.565	1.565
19	University of Illinois at Urbana-Champaign	44	1.592	0.145	1.760	1.850	1.947	1.592	1.592
20	University of Wisconsin-Madison	27	1.678	0.130	1.779	1.855	1.907	1.678	1.678
21	Washington University in St Louis	28	1.596	0.165	1.748	1.839	1.985	1.596	1.596
22	University of Southern California	73	1.565	0.133	1.793	1.877	1.982	1.565	1.565
23	Boston University	38	1.516	0.128	1.717	1.795	1.884	1.516	1.516
24	University of North Carolina at Chapel Hill	29	1.591	0.142	1.723	1.836	1.917	1.591	1.591
25	University of California, Davis	7	1.628	0.146	1.712	1.776	1.949	1.628	1.628
27	Ohio State University	37	1.619	0.133	1.802	1.902	1.963	1.619	1.619
28	Emory University	22	1.564	0.156	1.869	1.940	2.003	1.564	1.564

29	University of Minnesota	29	1.657	0.129	1.786	1.863	1.931	1.657	1.657
30	University of Maryland, College Park	29	1.589	0.112	1.750	1.801	1.907	1.589	1.589
31	Purdue University West Lafayette	16	1.561	0.154	1.750	1.893	1.972	1.561	1.561
32	University of California, Irvine	19	1.629	0.146	1.774	1.870	1.979	1.629	1.629
33	Dartmouth College	15	1.591	0.148	1.758	1.836	1.916	1.591	1.591
34	Michigan State University	34	1.495	0.201	1.737	1.828	1.999	1.495	1.495
35	Vanderbilt University	11	1.795	0.117	1.894	2.023	2.068	1.795	1.795
36	Penn State University	29	1.520	0.123	1.761	1.844	1.925	1.520	1.520
37	University of Virginia	14	1.791	0.117	1.870	1.968	2.039	1.791	1.791
38	Georgetown University	24	1.508	0.150	1.787	1.876	1.961	1.508	1.508
39	Case Western Reserve University	13	1.530	0.181	1.769	1.791	1.911	1.530	1.530
40	University of Arizona	23	1.603	0.101	1.819	1.911	1.925	1.603	1.603
41	Rice University	26	1.708	0.138	1.847	1.900	2.009	1.708	1.708
42	University of Colorado Boulder	33	1.566	0.161	1.744	1.826	1.951	1.566	1.566
43	University of Pittsburgh	6	1.721	0.077	1.826	1.853	1.905	1.721	1.721
44	Indiana University	48	1.513	0.134	1.717	1.777	1.862	1.513	1.513
45	University of Rochester	22	1.505	0.144	1.729	1.866	1.987	1.505	1.505
46	University of Florida	33	1.614	0.111	1.808	1.872	1.958	1.614	1.614
47	Tufts University	12	1.572	0.141	1.744	1.891	1.955	1.572	1.572
48	Rutgers, the State University of New Jersey	31	1.491	0.121	1.709	1.816	1.864	1.491	1.491
49	University of Alabama	22	1.656	0.097	1.818	1.863	1.957	1.656	1.656
50	University of Notre Dame	52	1.653	0.137	1.796	1.902	1.989	1.653	1.653
51	Northeastern University	59	1.523	0.148	1.754	1.872	1.967	1.523	1.523

Table 9. The logistic regression summaries on the different sub-samples

This table shows the summaries of the full logistic regressions on the different sub-samples. FWHR_HIGH is the fWHR dummy; MALE is the gender dummy; US is the US domestic highest qualification dummy; the TITLE variables are the academic title dummies; 1ST_YEAR represents the year of the very first publication of a scholar; EXP is the publication experience of a scholar; EXP^2 represents the second order of the variable EXP; and THE_RANK represents the THE ranking for the corresponding institutions. The standard errors are the HAC robust standard errors. *, ** and *** indicate the significances at 10%, 5% and 1% levels individually.

Panel A. Logistic regressions for self-citation adjusted JIF and 5-year average JIF

The models (1) and (2) have the dependent variable PRODUCTIVITY_SCORE_HIGH which is transformed from the PRODUCTIVITY_SCORE calculated from the JIF_adj. Whereas the models (3) and (4) have the dependent variable PRODUCTIVITY_SCORE_HIGH which is transformed from the PRODUCTIVITY_SCORE calculated from the JIF_five. PRODUCTIVITY_SCORE_HIGH is equal to one if it is greater than the median and zero otherwise.

	JIF_adj_coauth_high	**JIF_adj_auth_high**	**JIF_five_coauth_high**	**JIF_five_auth_high**
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.553***	0.438**	0.533***	0.494***
	(0.166)	(0.173)	(0.171)	(0.172)
MALE	0.055	0.049	-0.055	0.107
	(0.230)	(0.230)	(0.249)	(0.248)
US	0.670**	0.457	0.608**	0.590**
	(0.282)	(0.304)	(0.287)	(0.297)
PROF	2.280***	2.538***	2.848***	2.531***
	(0.421)	(0.421)	(0.436)	(0.410)
ASSO_PROF	2.152***	2.127***	2.232***	1.805***
	(0.454)	(0.455)	(0.466)	(0.449)
ASSIS_PROF	0.301	0.150	0.685	-0.086
	(0.507)	(0.511)	(0.522)	(0.529)
SEN_LEC	-0.844	-0.535	-14.158***	-14.300***

	(1.176)	(1.231)	(0.522)	(0.509)
OTHER	0.112	-0.199	0.465	0.199
	(0.791)	(0.900)	(0.808)	(0.810)
1ST_YEAR	0.182***	0.254***	0.115***	0.155***
	(0.032)	(0.035)	(0.022)	(0.027)
EXP	87.013***	113.487***	63.608***	78.425***
	(13.797)	(15.043)	(9.841)	(11.930)
EXP2	-3.828	-9.605***	-5.156	-8.863***
	(3.128)	(3.233)	(3.180)	(3.354)
THE_RANK	-0.036***	-0.029***	-0.038***	-0.031***
	(0.005)	(0.006)	(0.006)	(0.006)
Constant	-2.655***	-2.596***	-2.903***	-2.639***
	(0.506)	(0.511)	(0.526)	(0.508)
N	935	935	935	935
Akaike Inf. Crit.	936.735	886.040	904.253	875.059

Panel B. Logistic regressions for a sub-sample of male scholars only

It is noteworthy that the MALE dummies are left out of the models as this sub-sample has the males only. PRODUCTIVITY_SCORE_HIGH is transformed from the PRODUCTIVITY_SCORE, which is equal to one if it is greater than the median and zero otherwise.

	N_coauth_high	**N_auth_high**	**JIF_coauth_high**	**JIF_auth_high**
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.532***	0.543***	0.584***	0.584***
	(0.195)	(0.201)	(0.181)	(0.185)
US	0.734**	0.672*	0.820**	0.545
	(0.297)	(0.349)	(0.325)	(0.351)
PROF	2.430***	2.913***	2.307***	2.306***
	(0.434)	(0.449)	(0.451)	(0.456)
ASSO_PROF	1.415***	1.842***	2.049***	1.928***
	(0.491)	(0.499)	(0.499)	(0.498)
ASSIS_PROF	-0.236	-0.073	0.303	0.054
	(0.598)	(0.608)	(0.567)	(0.567)
SEN_LEC	-0.366	-0.729	-0.840	-0.737
	(0.904)	(1.152)	(1.164)	(1.202)
OTHER	-0.046	0.393	0.262	-0.263
	(0.868)	(0.942)	(0.865)	(0.959)
1ST_YEAR	0.074***	0.106***	0.135***	0.193***
	(0.018)	(0.024)	(0.027)	(0.032)
EXP	50.168***	63.886***	65.804***	85.220***
	(7.874)	(10.249)	(10.992)	(13.288)
EXP^2	-7.718**	-13.792***	-4.570	-8.710***
	(3.399)	(3.644)	(3.149)	(3.278)
THE_RANK	-0.037***	-0.035***	-0.038***	-0.035***
	(0.006)	(0.007)	(0.006)	(0.006)
Constant	-2.392***	-2.684***	-2.568***	-2.224***
	(0.509)	(0.547)	(0.536)	(0.548)
N	793	793	793	793
Akaike Inf. Crit.	716.684	661.047	795.316	766.651

Panel C. Logistic regressions for a sub-sample of scholars with high-quality facial pictures

PRODUCTIVITY_SCORE_HIGH is transformed from the PRODUCTIVITY_SCORE, which is equal to one if it is greater than the median and zero otherwise.

	N_coauth_high	**N_auth_high**	**JIF_coauth_high**	**JIF_auth_high**
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.390*	0.389*	0.447**	0.375**
	(0.200)	(0.206)	(0.183)	(0.189)
MALE	0.071	0.017	-0.119	0.046
	(0.282)	(0.288)	(0.269)	(0.259)
US	0.777**	0.791**	0.738**	0.541
	(0.314)	(0.352)	(0.337)	(0.368)
PROF	2.114***	2.435***	2.209***	2.256***
	(0.453)	(0.422)	(0.477)	(0.478)

ASSO_PROF	1.088**	1.089**	1.854***	1.811***
	(0.503)	(0.468)	(0.515)	(0.512)
ASSIS_PROF	-0.576	-0.643	0.290	-0.038
	(0.611)	(0.580)	(0.575)	(0.578)
SEN_LEC	-0.611	-1.167	-0.891	-0.810
	(0.937)	(1.164)	(1.199)	(1.240)
OTHER	-1.145	-0.839	-0.621	-1.407
	(0.828)	(0.847)	(0.826)	(1.033)
1ST_YEAR	0.073***	0.103***	0.133***	0.193***
	(0.020)	(0.025)	(0.029)	(0.035)
EXP	49.177***	59.120***	62.339***	81.698***
	(8.565)	(10.536)	(11.498)	(13.849)
EXP^2	-4.358	-9.606***	-2.647	-7.461**
	(3.329)	(3.728)	(3.111)	(3.358)
THE_RANK	-0.031***	-0.026***	-0.033***	-0.030***
	(0.007)	(0.007)	(0.006)	(0.007)
Constant	-2.200***	-2.323***	-2.347***	-2.290***
	(0.556)	(0.546)	(0.578)	(0.594)
N	744	744	744	744
Akaike Inf. Crit.	676.037	630.136	765.932	724.914

Panel D. Logistic regressions for a sub-sample of authors with publications in the top three finance journals

The top three finance journals are JF, JFE and RFS. It is worth mention that all the productivity scores have been re-calculated based on the top three finance journals. PRODUCTIVITY_SCORE_HIGH is transformed from the PRODUCTIVITY_SCORE, which is equal to one if it is greater than the median and zero otherwise.

	N_coauth_top3_high	N_auth_top3_high	JIF_coauth_top3_high	JIF_auth_top3_high
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.429**	0.437**	0.451***	0.371**
	(0.175)	(0.180)	(0.168)	(0.170)
MALE	0.099	0.262	0.232	0.454**
	(0.253)	(0.251)	(0.250)	(0.229)
US	0.239	0.319	0.377	0.388
	(0.284)	(0.308)	(0.289)	(0.296)
PROF	2.117***	2.191***	2.547***	2.161***
	(0.502)	(0.503)	(0.634)	(0.591)
ASSO_PROF	1.221**	1.484***	1.982***	1.601***
	(0.523)	(0.532)	(0.646)	(0.603)
ASSIS_PROF	-0.825	-0.836	0.197	-0.446
	(0.626)	(0.655)	(0.691)	(0.661)
SEN_LEC	-13.936***	-13.792***	-13.234***	0.587
	(0.650)	(0.661)	(0.789)	(1.102)
OTHER	0.593	0.576	1.275	0.693
	(1.346)	(1.481)	(1.567)	(1.456)
1ST_YEAR	0.066***	0.088***	0.128***	0.166***
	(0.018)	(0.020)	(0.026)	(0.028)
EXP	34.194***	44.757***	47.973***	62.244***
	(7.159)	(8.183)	(10.164)	(11.196)
EXP^2	-4.982*	-11.412***	-3.093	-6.607**
	(2.770)	(3.041)	(2.697)	(2.873)
THE_RANK	-0.009	-0.013**	-0.019***	-0.004
	(0.006)	(0.006)	(0.006)	(0.006)
Constant	-1.912***	-2.228***	-2.693***	-2.444***
	(0.580)	(0.601)	(0.694)	(0.659)
N	781	781	781	781
Akaike Inf. Crit.	842.605	798.389	892.065	871.743

Panel E. Logistic regressions for a sub-sample of scholars with two-word names

PRODUCTIVITY_SCORE_HIGH is transformed from the PRODUCTIVITY_SCORE, which is equal to one if it is greater than the median and zero otherwise.

	N_coauth_high	N_auth_high	JIF_coauth_high	JIF_auth_high
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.528**	0.523**	0.462**	0.488**
	(0.208)	(0.218)	(0.193)	(0.201)
MALE	-0.014	0.052	0.072	0.079
	(0.269)	(0.279)	(0.252)	(0.251)
US	0.798**	0.572	0.771**	0.484

	(0.320)	(0.372)	(0.329)	(0.346)
PROF	2.228***	2.747***	2.502***	2.539***
	(0.450)	(0.456)	(0.499)	(0.491)
ASSO_PROF	0.972**	1.439***	2.051***	2.092***
	(0.491)	(0.489)	(0.529)	(0.518)
ASSIS_PROF	-0.484	-0.257	0.355	0.091
	(0.570)	(0.566)	(0.585)	(0.579)
SEN_LEC	-0.800	-14.432***	-14.302***	-14.182***
	(1.190)	(0.594)	(0.638)	(0.636)
OTHER	-0.585	-0.140	0.025	-0.836
	(0.828)	(0.958)	(0.888)	(1.046)
1ST_YEAR	0.088***	0.120***	0.184***	0.231***
	(0.027)	(0.036)	(0.041)	(0.042)
EXP	50.632***	63.157***	76.283***	89.667***
	(10.526)	(13.717)	(15.328)	(15.642)
EXP^2	-8.179**	-14.929***	-5.358*	-10.351***
	(3.262)	(3.753)	(3.083)	(3.372)
THE_RANK	-0.023***	-0.016**	-0.027***	-0.022***
	(0.007)	(0.007)	(0.007)	(0.007)
Constant	-2.451***	-2.762***	-3.167***	-3.017***
	(0.573)	(0.604)	(0.596)	(0.589)
N	702	702	702	702
Akaike Inf. Crit.	635.034	582.801	692.143	660.569

Panel F. Logistic regressions for a sub-sample of scholars with more-than-two-word names

PRODUCTIVITY_SCORE_HIGH is transformed from the PRODUCTIVITY_SCORE, which is equal to one if it is greater than the median and zero otherwise. It is worth mentioning that the FWHR_HIGH dummy is all statistically significant across all the models except for the models (1) and (2). However, the sample size is relatively small with only 233 data points, and they are the only two models showing insignificant results. Therefore, these do not raise my concerns.

	N_coauth_high	**N_auth_high**	**JIF_coauth_high**	**JIF_auth_high**
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.214	0.564	0.764**	0.888**
	(0.361)	(0.387)	(0.345)	(0.365)
MALE	-0.476	-0.815	-0.196	0.541
	(0.712)	(0.714)	(0.690)	(0.639)
US	0.270	0.318	0.412	0.074
	(0.600)	(0.644)	(0.618)	(0.796)
PROF	3.098***	2.928***	2.217***	2.408***
	(0.961)	(0.870)	(0.845)	(0.888)
ASSO_PROF	2.811**	1.403	1.862*	1.231
	(1.187)	(1.037)	(1.009)	(1.026)
ASSIS_PROF	0.155	-16.199***	0.613	-0.232
	(1.619)	(0.925)	(1.133)	(1.096)
SEN_LEC	0.041	-0.236	-0.381	0.044
	(1.443)	(1.421)	(1.379)	(1.475)
OTHER	0.973	0.546	0.460	0.532
	(2.334)	(2.403)	(2.235)	(2.115)
1ST_YEAR	0.040	0.103***	0.092***	0.167***
	(0.025)	(0.030)	(0.030)	(0.053)
EXP	19.637***	32.520***	26.042***	38.836***
	(5.426)	(7.001)	(5.949)	(9.986)
EXP^2	-4.190	-5.131	-0.729	-0.623
	(3.282)	(3.800)	(2.993)	(3.250)
THE_RANK	-0.054***	-0.059***	-0.052***	-0.053***
	(0.012)	(0.014)	(0.011)	(0.011)
Constant	-2.010*	-1.129	-1.823	-1.913
	(1.202)	(1.144)	(1.176)	(1.303)
N	233	233	233	233
Akaike Inf. Crit.	231.805	200.534	258.986	240.743

Table 10. OLS regression summaries for a sub-sample of female scholars only

This table shows the summaries of the full OLS regression on the female-only sample. FWHR_HIGH is the fWHR dummy; US is the US domestic highest qualification dummy; the TITLE variables are the academic title dummies; 1ST_YEAR represents the year of the very first publication of a scholar; EXP is the publication experience of a scholar; EXP^2 represents the second order of the variable EXP; THE_RANK represents the THE ranking for the corresponding institutions. It is noteworthy that the MALE dummies are left out of the models since this sub-sample has the females only. The standard errors are the HAC robust standard errors. *, ** and *** indicate the significances at 10%, 5% and 1% levels individually.

	log(N_coauth)	**log(N_auth)**	**log(JIF_coauth)**	**log(JIF_auth)**
	(1)	(2)	(3)	(4)
FWHR_HIGH	-0.141	-0.110	-0.138	-0.096
	(0.123)	(0.103)	(0.132)	(0.115)
US	0.117	0.018	0.149	0.043
	(0.205)	(0.175)	(0.199)	(0.176)
PROF	1.119***	0.934***	1.551***	1.393***
	(0.392)	(0.298)	(0.431)	(0.342)
ASSO_PROF	0.759**	0.511*	1.133***	0.922***
	(0.376)	(0.288)	(0.428)	(0.348)
ASSIS_PROF	-0.022	-0.180	0.301	0.170
	(0.398)	(0.316)	(0.458)	(0.388)
SEN_LEC	-0.058	-0.310	0.190	-0.047
	(0.674)	(0.568)	(0.669)	(0.647)
OTHER	-1.030***	-1.313***	-0.440	-0.716*
	(0.394)	(0.304)	(0.452)	(0.378)
1ST_YEAR	0.028	0.040***	0.088***	0.099***
	(0.018)	(0.013)	(0.023)	(0.019)
EXP	5.942***	7.668***	9.772***	11.690***
	(1.989)	(1.518)	(2.406)	(1.952)
EXP^2	-0.416	-0.719	-0.030	-0.374
	(0.891)	(0.883)	(1.001)	(0.977)
THE_RANK	-0.007	-0.001	-0.009*	-0.004
	(0.004)	(0.004)	(0.005)	(0.004)
Constant	-56.264	-78.119***	-174.783***	-195.467***
	(35.994)	(26.635)	(45.293)	(36.962)
N	142	142	142	142
Adjusted R^2	0.488	0.568	0.490	0.558

Appendix C. Productivity quality measure selection

In order to measure a scholar's research performance quality, I need to find a method to quantify the quality of the publications. In fact, Kerl et al. (2018) have classified the approaches into two categories. First, each paper's quality is measured directly based on its citations. Second, a paper's quality is proxied indirectly by the quality of the journal where it is published. The first approach assesses a paper's quality based on some commonly used citation indices, such as the *Social Science Citation Index*, *Web of Science*, *Scopus,* and *Google Scholar*. However, these citation indices have been criticised mostly for suffering from self-citation bias (Chan et al., 2002). Furthermore, they do not take into account that authors affiliated with top-ranked institutions are more likely to publish their papers in high quality journals (Chen & Huang, 2007). There are more complex methods developed based on citation approach. Bibliometrics is a good example. It uses statistical methods to analyse publications' quality in multidimensions. It is widely used in analysing science publications but receives less attention in the finance discipline (Karolyi, 2016).

The second category from Kerl et al.'s (2018) classification is to use the ranking of journals to infer the quality of articles published in them.[38] There are three prevailing methods for ranking journals: journal page-based method, peer assessment method and publication citation-based method (Currie & Pandher, 2011). The well-known page-based approach is *Journal of Finance-equivalent pages* (JF-Pages). The *Journal of Finance* is regarded as the highest-quality journal in the finance discipline. Chan et al. (2001) construct a JF-pages factor by first randomly selecting three full-text pages without any equations, footnotes nor graphs for each journal, then count and take average of the total words. Next, they use this representative word number of a journal to compare with that of the *Journal of Finance*. The final ratio will be the journal's JF-Pages factor. For example, if a representative page of the *Journal of Financial Economics* has an average of 500 words, and that of the *Journal of Finance* is 450 words, then the *Journal of Financial Economics*' JF-Pages factor is 450/500 = 0.9. As a result, the *Journal of Financial Economics*' final JF-Pages score is the total page count multiply the corresponding JF-Pages factor. Based on the description, it is noticeable that the JP-Pages are coarse approximations of ratings and suffer from a downward bias. The authors themselves point out in a later paper that some high quality theoretical and mathematical articles can be very short in pages (Chan et al., 2011).

On the other hand, the peer assessment method is also called the survey approach, which is to conduct a survey to rank journals based on the opinions of experts. For example, Oltheten et al. (2005) conduct a global survey in order to minimise a potential regional bias and they find that the remarkable consistency in the top four finance journals: the *Journal of Finance*, the *Journal*

[38] By doing this, it implies that the articles published on the same journals should have the same quality scores before any adjustments.

of Financial Economics, the *Review of Financial Studies*, and the *Journal of Financial and Quantitative Analysis*. However, the perception of quality for the remaining journals varies along with the respondents' location, interests, seniority and journal affiliation (Oltheten et al., 2005). Additionally, Currie and Pandher (2011) and Currie and Pandher (2020) narrow the survey targets down to the active finance scholars by arguing that "active scholars are more aware and current in their knowledge of journal quality and awareness in the field". However, it has been pointed out that the peer assessment method generally suffers from non-response biases, respondents' perceptions biases and sampling biases (Chan & Liano, 2009; Moosa, 2011).

Finally, the publication citation-based method is the most commonly used. Many academic institutes and associations publish their own journal rankings. The three popular journal rankings are the Australian Business Deans Council (ABDC) journal classification list, the Association of Business Schools (ABS) Academic Journal Quality Guide and the Journal Impact Factor (JIF) published by the Clarivate (formerly Institute of Scientific Information, ISI). The former two only give ordinal rankings of journals. For instance, ABS Academic Journal Quality Guide classifies journals into five categories from grades 4* to 1. The 4* ranking represents the top-quality internationally recognised journals. Whereas the ABDC journal classification list classifies journals into four categories from A* to C, with A* representing the highest-quality journal ranking. On the other hand, JIF ranks journals by calculating year-adjusted citations. Moreover, it also reports self-citation adjusted scores. As a result, I have decided to choose JIF to gauge scholars' academic performance in my research. [39]

[39] There are other journal ranking indices constructed in existing finance literature. For example, Chen and Huang's (2007) Author Affiliation Index (AAI) and Jarnecic et al.'s (2008) Research Productivity Dependency index (RPD). These indices do capture the author's affiliations. However, they might have been out of date as it has been over ten years since the methods were developed; further, the relative ranking of journals does change over time (Arnold et al., 2003).

Appendix D. Data collection details

D.1. fWHR data collection

I first define the institution list by following one of the best-known international university ranking systems - the *Times Higher Education Supplement (THE) World University Rankings* to select the top 50 institutions 2021 of *accounting and finance* subjects in the US region.[40] There are three reasons to select THE ranking over the other ranking systems. First, the THE ranking is one of the most well-known ranking systems. Second, the research performance tends to be more focused than the other performance indicators when it comes to the global university rankings (Pavel, 2015). Last, the THE ranking offers broader and more diverse performance measures than the other rankings (Pavel, 2015).

Meanwhile, I record the ranks – the variable THE_RANK in an ascending order based on the US regional ranking.[41] Next, I identify all the finance scholars from the finance departments on each institution's faculty website. However, an institution may not have a separate finance department. Instead, it may have a conjugate finance department with other disciplines, such as an accounting and finance department, finance and economics department, banking and finance department, or finance, insurance and real estate department. Under these circumstances, I screen out the finance scholars by narrowing the speciality or expertise down to the finance area.

Afterwards, I start collecting these scholars' information from each institution's website. The personal information includes a scholar's full name, gender, academic title, and the origin country of the scholar's highest qualification. In particular, I measure a scholar's fWHR from the person's profile photo on the institution's website by using my Python program. Specifically, by following Kamiya et al.'s (2019) suggestion, the facial width is measured as the distance between the highest points of the left and right zygomatic bone, whereas the facial height is measured as the distance between the midpoints of two inner ends of the eyebrows and the upper lip.[42] Consequently, the fWHR is calculated as the facial width divided by the facial height. Further, I classify the picture into two categories: all pictures and high-quality pictures, based on the three criteria by following Jia et al.'s (2014) suggestions. They are picture resolution (e.g., less than 120x150 pixels),

[40] I select the *overall* indicator instead of the *research* indicator alone because the research performance measure of each institution has been spread out across other indicator components. For example, the *Research,* the *Citations,* the *International Outlook* and the *Industry Income* categories all contain the research performance measures (please see https://www.timeshighereducation.com/world-university-rankings/world-university-rankings-2021-methodology). I understand this is an imperfect choice but believe it is a trivial matter since I measure the research productivity on an individual scholar level.
[41] THE provides only the world ranking. Thus, I adjust the ranks manually.
[42] Jia et al. (2014) propose a slightly different measure of the upper facial height, which is the distance between the upper lip and the highest point of the eyelids. However, this measurement difference is neglectable as it produces trivial matters (Kamiya et al., 2019).

unnatural expression (e.g., laughing) and face tilting (e.g., more than 45 degrees). In these cases, I consider the pictures as low-quality.

Furthermore, I classify the academic titles into a total of six levels according to the implied research abilities. For example, a *full professor* title implies higher research ability than that an *assistant professor* title implies. Thus, I classify these two titles into two separated levels. Albeit the US has a matured and well-structured academic title ranking series, and the fundamental hierarchy is *professor*, *associate professor*, *assistant professor*, and *instructor*, the variations of academic titles in use for each US institution are substantial (Shamos, 2011). For example, the title *senior research scientist* is equivalent to *associate professor* in terms of the title hierarchy but only tenure differs in *Carnegie Mellon University* (Shamos, 2011). In addition, the mixed use of official titles, working titles and functional titles can produce noises in the classification. For instance, the title *teaching professor* is a type of functional title, which implies only the teaching function but not a research ability. On the other hand, the title *head of the department* is a type of working title, which implies only the person's working position but not a research ability. Furthermore, the existing title prefixes and suffixes can distort the classification. Take the title *professor of practice* or *clinical professor* as examples; although both titles have "professor" in them, they are not necessarily implying the same level of research ability as the title *full professor* does. In order to best filter the potential noises mentioned previously, I decide to manually classify each scholar's academic title by following the *Handbook of Academic Titles* in the US institutions created by Shamos (2011) along with carefully examining the resume of each scholar.

The six levels of the title classification are professors – the variable: PROF, associate professors – the variable: ASSO_PROF, assistant professors – the variable: ASSIS_PROF, senior lecturer – the variable: SEN_LEC, other – the variable: OTHER and industry professionals – the baseline dummy PRAC. In particular, the level of professors includes distinguished, endowed or university professors and full professors. The levels of associate and assistant professors include associate professors and assistant professors respectively. The level of senior lecturers includes senior, distinguished or master lecturers and some teaching professors. The level of other includes lecturers, instructors, adjunct faculty, research associates and PhD students. Last, clinical professors, professors of practice, associate professors of practice, assistant professor of practice, senior or junior professional instructors are included in the level of industry professionals. Admittedly, these are not hard lines but provide guidelines for the classification. I make adjustments in accordance with the career history from the scholars' resumes. For example, for a scholar with a *clinical professor* title has worked in the research department of a reserve bank for thirty years, I classify this person into the level of professors. Likewise, for a scholar with a *senior lecturer* title has worked in the Wall Street for over a quarter of century, I classify this person into the level of industry professionals.

Additionally, I adjust a scholar's full name, which is listed on the institution's website by following the given publication information in the person's resume in order to better prepare for the publication data collection of the scholar at the next stage. This adjustment is mainly to add or delete the middle name(s) depending on the publication name the scholar commonly uses. For example, if a scholar commonly uses first and last name to publish articles, and this person's full name is recorded from the institution's website, I will delete the middle names to match that. Conversely, the middle name(s) is(are) added.[43] Furthermore, by reading the information listed both on the institutions' websites and in scholars' personal resumes, I find out the origin countries where scholars were granted the highest qualifications from. A dummy variable US with one is assigned for US domestic graduates and zero for otherwise. Last, a dummy variable MALE with one is for marking male scholars while zero for females.

D.2. Scholars' publication data collection

In this data collection phase, I collect each scholar's publication records from the EBSCO database by searching the scholar's publication name, which I have obtained from the previous data collection phase. Especially the publication titles, the publication types, the subjects, the authors and co-authors, the publication designations (mainly the journal titles) and the publication years are recorded. Also, the total number of authors and co-authors of each publication is recorded accordingly.

As the name is the only searching criterion, I deploy several filtering mechanisms in order to minimise the potential measurement errors. First, I design an automation Python program from the Python module *Selenium* for collecting the data. This helps avoid potential human errors by manually collecting. Second, the preferred publication name used by each scholar I have identified previously helps me target the scholar accurately when searching. Despite that, some of the common names can still create great distortion. For example, it has a total of 78,804 records showing up when searching by the name *An Li*. Clearly, the search result of 78,804 records is not sensible for this scholar. Therefore, I skip this scholar. I set up a maximum of 2,000 search results to be the elimination line for this filtering purpose. Moreover, both the information of the publication types and the subject of each publication provide me means to reduce the potential over-collecting issue due to the duplications of author names. Particularly, I exclude the publication types other than academic journal and periodical. In addition, the subjects other than finance, accounting, economics, management, and business are excluded as well. I also filter out the duplicate records by using the publication titles. Last, the journal list from the JCRs, which I collect next, also helps me reduce the over-collecting issue.

[43] The added middle name(s) is(are) often initial(s).

D.3. JCR data collection

Next, I collect the JCR data from the Clarivate's JCR website. The JCR data serves two main purposes for this research. First, I use the journal lists provided by the Clarivate's JCRs to identify the years of each scholar's very first and latest publications. The reason that I choose the JCR journal list over the FT50 journal list for this purpose is that JCRs offer a more comprehensive journal list, which covers the majority of academic journals over all the world. Due to the interdisciplinary nature of the finance discipline, and the multidisciplinary expertise that some finance scholars have, the combination of the journal lists from the four categories: finance, economics, business, and management, is considered.[44] Furthermore, I download the complete journal list of each category for every year from the Clarivate's JCRs website and combine all the lists into a final journal list. To do this, I avoid potential under-collecting issue by just selecting the latest journal lists from each category because of the possible name changes or issue changes of any journals. For example, the *Journal of Business* ceased publication in 2006, and thus is not included in any journal lists from 2007 onwards under any JCR categories. Another example is that the *Industrial Management Review* changed its title to *MIT Sloan Management Review* in 2001. As a result, any publications in the *Industrial Management Review* before 2001 would not be recognised if the journal list of any single year after 2001 were used alone. At the end, I use the variable 1st_YEAR to denote the year of each scholar's very first publication and the variable EXP to represent the total number of years between each scholar's first and latest publications, which proxies each scholar's research experience.

The second reason for collecting JCR data is to quantify the quality of each journal from the FT50 journal list. The JCRs provide three main citation-based journal quality measurements: the Journal Impact Factor (JIF), the Journal Impact Factor adjusted for self-citations (JIF_adj) and the five-year Journal Impact Factor (JIF_five). Simply put, the JIF is the previous two-year average citations of a journal as a whole, whereas the JIF_adj is the JIF adjusted after self-citations being removed, and the JIF_five is the previous five-year average citations of a journal as a whole. However, the JCRs have the JIF data only available from the earliest 1997 to the latest 2019 at the time of writing. By using the available JIF data only, I will under-estimate the research performances of some senior scholars. Taking Professor *William F. Sharpe* as an example, most of his work was published in the 70s and 80s. Hence, I decide to extrapolate the absent JIF values from the available JCR data. More specifically, I set the missing JIF of a journal at a particular year equal to the available JIF of that journal at the nearest year. For instance, the JIF of the *Journal of Finance* at the year 1991 is unavailable. Consequently, I assign the available JIF of the *Journal of Finance* at the nearest year 1997, that is 2.173, to the missing JIF value. Similarly, the JIF of the *Journal of Financial Economics* at the year 2021 is unavailable and I set this missing

[44] The Clarivate's JCRs does not have the accounting category. Consequently, any accounting academic journals have been classified into other categories.

JIF value equal to the available JIF of the same journal at the nearest year 2019, that is 5.731. The same rules are applied to all the missing values of the other two JIF alternatives.

Appendix E. Additional model fitting

E.1. The OLS of the publication rate on fWHR

In previous models, I assume that a scholar's publication experience is not linearly related with the scholar's productivity. This is because a scholar's publication experience at the early stage might have accelerated effects on the productivity, and these positive effects have a diminishing return on the increase of productivity along the increase of experience. In other words, the publication experience elasticity of productivity is decreasing along the timeline. In order to address this relationship, I add a quadratic term of the variable EXP. Also, the coefficients of the first order EXP with positive signs and the coefficients of the second order EXP^2 with negative signs validate the diminishing return relationship. However, although the variable EXP is highly significant across all the models, the quadratic term EXP^2 is only significant in some of the models. As a result, one may propose that instead of adding the quadratic term of EXP, I should consider regressing the publication rates of scholars, that is $productivity\_score_i/EXP_i$ for scholar i. Following this idea, I re-fit an OLS model by regressing the publication rates on all other variables. The specification is:

$$\log\left(\frac{productivity\_score_i}{EXP_i}\right) = \alpha + \beta_1 * FWHR\_HIGH_i + Control\ variales + \beta_{10} * THE\_RANK_i + \varepsilon_i \quad [6]$$

where $\log(PRODUCITVITY\_SCORE_i/EXP_i)$ is the natural log transformation of the productivity rate for scholar i and $FWHR\_HIGH_i$ is a dummy variable equal to one if scholar i's fWHR is greater than the median and zero otherwise; the control variables are the aforementioned variables, which include $MALE_i$, US_i, the title categorical variables $TITLE_i$, and $1st\_YEAR_i$; THE_RANK_i is the control variable for controlling the institutional differences.

The results in *Panel A* of *Table 11* show that the coefficients of FWHR_HIGH are all positive and statistically significant. However, the interpretations of these coefficients are different from the ones in the previous models. For example, first considering an average baseline scholar: a male with a low-fWHR, the highest qualification attained in the US and a *professor* title from *Boston University* (ranked at 23), is predicted to have a publication quantity rate of 0.37 per year and a publication quality rate of 1.12 per year on average. In contrast, a high-fWHR scholar from the exact same background will have a publication quantity rate and a publication quality rate of 0.42 and 1.29 per year respectively.[45] As a result, there are 13.5% increase in publication quantity rate per year and 15.2% in publication quality rate per year.[46] The MALE and US dummies are

[45] 0.37 = [exp (-2.373+0.091+0.045+0.754) + exp (-1.525+0.123-0.033+0.755)]/2 and
1.12 = [exp (-1.703+0.14+0.09+1.089) + exp (-0.841+0.164+0.02+1.099)]/2,
while 0.42 = [exp (-2.373+0.133+0.091+0.045+0.754) + exp (-1.525+0.137+0.123-0.033+0.755)]/2 and
1.29 = [exp (-1.703+0.134+0.14+0.09+1.089) + exp (-0.841+0.138+0.164+0.02+1.099)]/2.

[46] 13.5% = (0.42-0.37)/0.37 and 15.2% = (1.29-1.12)/1.12.

either insignificant or marginally significant due to the possible type I error. Interestingly, the coefficients of ASSIS_PROF are higher than those of ASSO_PROF in all four models. The possible explanation is that the scholars with an *assistant professor* title are at the early stage of the career trajectory and thus benefiting from the accelerated effects from the experience. By contrast, the scholars with an *associate professor* title are entering the mature stage of the career trajectory and therefore suffering from a diminishing return of the previous experience (e.g., creativity bottle necks). On the other hand, the scholars with a *professor* title are at the stable stage of the career trajectory with dominating numbers of publications and possible maximum resource support comparing to an assistant professor. As a result, their publication rates remain the highest among the academic title rankings.

Table 11. The full regression summaries of the two more alternative models

Panel A. OLS of the publication rate on fWHR

This table shows the regression results of the OLS of the publication rate on fWHR. In particular, FWHR_HIGH is the fWHR dummy; MALE is the gender dummy; US is the US domestic highest qualification dummy; the TITLE variables are the academic title dummies; 1st_YEAR represents the year of the very first publication of a scholar; EXP is the publication experience of a scholar; and THE_RANK represents the THE ranking for the corresponding institutions. Moreover, the dependent variables are the publication rates which are LOG(PRODUCTIVITY_SCORES/EXP). The standard errors are the HAC robust standard errors. *, ** and *** indicate the significances at 10%, 5% and 1% levels individually.

	log(N_coauth/exp)	log(N_auth/exp)	log(JIF_coauth/exp)	log(JIF_auth/exp)
	(1)	(2)	(3)	(4)
FWHR_HIGH	0.133***	0.137***	0.134**	0.138**
	(0.051)	(0.048)	(0.056)	(0.054)
MALE	0.091	0.123*	0.140*	0.164**
	(0.077)	(0.068)	(0.084)	(0.077)
US	0.045	-0.033	0.090	0.020
	(0.099)	(0.088)	(0.106)	(0.096)
PROF	0.754***	0.755***	1.089***	1.099***
	(0.155)	(0.134)	(0.170)	(0.155)
ASSO_PROF	0.545***	0.424***	0.954***	0.834***
	(0.161)	(0.139)	(0.179)	(0.163)
ASSIS_PROF	0.606***	0.464***	1.102***	0.951***
	(0.164)	(0.145)	(0.180)	(0.167)
SEN_LEC	-0.087	-0.030	-0.029	0.010
	(0.303)	(0.295)	(0.310)	(0.327)
OTHER	-0.100	-0.064	-0.038	-0.018
	(0.327)	(0.305)	(0.339)	(0.314)
1ST_YEAR	0.020***	0.028***	0.038***	0.045***
	(0.003)	(0.003)	(0.003)	(0.003)
THE_RANK	-0.011***	-0.009***	-0.015***	-0.013***
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	-2.373***	-1.525***	-1.703***	-0.841***
	(0.186)	(0.158)	(0.203)	(0.179)
N	935	935	935	935
Adjusted R^2	0.188	0.250	0.380	0.431

Panel B. The regression summaries of the mixed effects models

This table shows the regression results of the linear mixed effects models. In particular, the dependent variables are LOG(PRODUCTIVITY_SCORES); FWHR_HIGH is the fWHR dummy; MALE is the gender dummy; US is the US domestic highest qualification dummy; the TITLE variables are the academic title dummies; 1st_YEAR represents the year of the very first publication of a scholar; EXP is the publication experience of a scholar; EXP^2 represents the second order of the variable EXP; and THE_RANK represents the THE ranking for the corresponding institutions. However, the mixed effects models do not have the ordinary HAC robust standard errors. Instead, I apply the bootstrapping technique to derive the standard errors. The results are consistent and not presented in the report but available upon request. *, ** and *** indicate the significances at 10%, 5% and 1% levels individually.

	log(N_coauth) (1)	log(N_auth) (2)	log(JIF_coauth) (3)	log(JIF_auth) (4)
FWHR_HIGH	0.114**	0.127***	0.122**	0.136***
	(0.049)	(0.046)	(0.054)	(0.052)
MALE	-0.001	0.039	0.048	0.081
	(0.069)	(0.065)	(0.075)	(0.073)
US	0.128	0.046	0.171*	0.096
	(0.083)	(0.078)	(0.091)	(0.088)
PROF	0.994***	0.945***	1.257***	1.210***
	(0.105)	(0.099)	(0.115)	(0.112)
ASSO_PROF	0.751***	0.606***	1.105***	0.955***
	(0.117)	(0.110)	(0.129)	(0.125)
ASSIS_PROF	0.131	0.009	0.534***	0.403***
	(0.127)	(0.120)	(0.140)	(0.135)
SEN_LEC	-0.315	-0.236	-0.207	-0.144
	(0.202)	(0.190)	(0.222)	(0.214)
OTHER	-0.182	-0.136	-0.144	-0.113
	(0.223)	(0.210)	(0.245)	(0.237)
1ST_YEAR	0.015***	0.027***	0.044***	0.056***
	(0.004)	(0.004)	(0.004)	(0.004)
EXP	20.577***	22.972***	26.173***	29.018***
	(1.673)	(1.572)	(1.836)	(1.776)
EXP^2	-2.055**	-3.197***	-1.118	-2.396***
	(0.849)	(0.798)	(0.931)	(0.901)
Constant	0.167	1.045***	0.891***	1.790***
	(0.136)	(0.127)	(0.150)	(0.145)
INSTITUTION (effects)	Random	Random	Random	Random
N	935	935	935	935
Akaike Inf. Crit.	2,153.897	2,036.535	2,329.967	2,266.433

E.2. Linear mixed effect model

One may argue that the institution fixed effects control variable is not appropriate here. First, a scholar may be affiliated with multiple institutions. Particularly, for a scholar having the qualification granted in an institution but finding a job in another institution, the scholar may keep the affiliation with the previous institution. Second, a scholar from an institution may work with the scholars from other institution. This situation is very common in co-authorship networks. Specifically, Besancenot et al. (2017) and Seibert et al. (2017) find that co-authoring with scholars from top-ranked institutions increases research productivity. These two possibilities create a crossed effect among the institution categories and thus, the institutional fixed effects are inappropriate. Lastly, there are repeatedly measured data points existing in the sample. For example, the Adjunct Full Professor of Finance *Karin S. Thorburn* from *University of Pennsylvania* is visiting *Dartmouth College* as a visiting professor at the time I am collecting the data. As a result, Professor *Karin S. Thorburn* is measured twice in the sample. Another example is that Professor *Yael V. Hochberg* from *Rice University* is visiting *Northwestern University* as a visiting professor at the time of writing. These repeatedly measured data points make the institution fixed effects invalid.

In these cases, a random-effects model comes to rescue. A random-effects model is also called a variance components model (Gelman, 2005). In contrast to the fixed effects model controlling the mean differences, it controls the variance instead (Gelman, 2005). It will properly address the cross-effects between each level of the INSTITUTION categorical variable. Therefore, by following this idea, I fit a linear mixed model. The specification is:

$$\log(productivity\_score_i) = \alpha + \beta_1 * FWHR\_HIGH_i + Control\ variales + random\_effects(INSITUTION_i) + \varepsilon_i \qquad [7]$$

where $\log(PRODUCITVITY\_SCORE_i)$ is the natural log transformation of the productivity scores for scholar i and $FWHR\_HIGH_i$ is a dummy variable equal to one if scholar i's fWHR is greater than the median and zero otherwise; the control variables are the aforementioned variables, which include $male_i$, US_i, the title categorical variable $TITLE_i$, $1st\_YEAR_i$, EXP_i and $(EXP_i)^2$; and $INSITUTION_i$ here is the random effects control variable.

The results are displayed in *Panel B* of *Table 11*. Evidently, the coefficients of FWHR_HIGH still show significances but increased slightly in magnitudes from the previous fixed effects models. The results of other variables are similar to those in the fixed-effects models. However, the random effects explain only 8.9%, 7.2%, 11.7% and 10.3% of the total residual variation in four models, respectively.[47] Such low proportions of variation explained by the random effects added suggest that the institution random effects might not be efficient for controlling the variation.

[47] I use the variance of INSTITUTION random effects and residual reported in the regression summaries to calculate the proportion of residual variance explained by the INSITUTION random effects. Particularly, for the model (1), it is calculated as 8.9% = 0.05349/ (0.05349+0.52686); for the model (2), it is calculated as 7.2% = 0.03633/ (0.03633+0.4672); for the model (3), it is calculated as 11.7% = 0.08335/ (0.08335+0.63118) and for the model (4), it is calculated as 10.3% = 0.06784/ (0.06784+0.59209).