The Face of Success*

Shannon N. Minehan† and Dennis Wesselbaum‡

June 7, 2021

Abstract

In this paper, we add to the literature on gender differences in labour market outcomes by studying the drivers of academic success. We extend the literature by incorporating otherwise unobserved personality traits. We compile a unique data set containing detailed information of all full-time faculty members of the Top 100 US Economic Departments (N=2,473). Data is taken from individual CVs and Google Scholar. We use the facial width-to-height ratio (fWHR) to proxy personality traits such as competitiveness, dominance, and risk-taking behaviour. Our results show that more dominant women are less successful (i.e. have a lower h-Index) compared to less dominant women, while more dominant men are more successful after controlling for various confounding variables. These findings emerge over the life-cycle of a researcher and survive various robustness checks.

Keywords: Academic Performance, Gender, Labour Market Outcomes, Personality.

JEL classification: A14, D91, J16, J24.

*We wish to thank Chris Barrett, Sandy Black, Juliane Hennecke, Sol Hsiang, and Andrea Menclova for helpful comments and suggestions. We also thank participants at the Microeconomics Study Group Meeting (online) for valuable feedback. All remaining errors are our own.

†Motu Economic and Public Policy Research Email: shannon.n.minehan@gmail.com

‡Department of Economics, University of Otago, Dunedin 9016, New Zealand. Email: dennis.wesselbaum@otago.ac.nz

A thesis

submitted to the Department of Economics

University of Otago

In partial fulfilment of the requirements

for the degree of Master in Economics

February 2021

Contents

Table of Abbreviations

AEA	American Economic Association
EEA	European Economic Association
FD	First Differences
fWHR	Facial Width-to-Height Ratio
GRE	Graduate Record Examination
OLS	Ordinary Least Squares
RES	Royal Economic Society
RMSE	Root Mean Square Error
SFD	Spatial First Differences
SSRN	Social Science Research Network

List of Figures

List of Tables

Table of Symbols

t	Time
d	Oster Test Coefficient
$\Phi_{i,t}$	Vector of skills at any point in time
i	Index of gender
M	Male
F	Female
$\Phi_{i,t}$	Cognitive skill vector
P_i	Stable component
I	Investment
C	Cognitive skills
N	Non-cognitive skills
j	Index of skills
$f_{i,t}^{j}$	Monotonically increasing function
Y	Outcomes
g	Monotonically increasing function
m	Specific outcome
e	Error term
l	Lag
n	Contemporaneous outcomes
R	Incentives
h	Function links inputs into the effort function to the task-level effort
$\mathbb{E}$	Mathematical expectation operator
Θ_i	Set of preferences
ϵ_i	Error term

β	Constant
$\mathbf{x}_i$	K-dimensional matrix of control variables
$\mathbf{c}$	M-dimensional vector of unobservables
Δ	Difference operator

1 Introduction

Persistent gender differences in labour market outcomes are a well-researched area in economics. Papers such as Blau and Kahn, (2017); Cortes and Pan, (2017); Risse et al. (2018); and Goldin et al. (2017) have highlighted a number of drivers of the divide, including but not limited to factors such as education, experience, and social norms and beliefs (Akerlof and Kranton, 2000). Despite these developments in knowledge and in the techniques used to raise awareness of and minimise the observable differences however, the gender gap continues to persist.

We speculate unobservable personality traits such as dominance may be contributing.[1] In particular, we postulate social norms surrounding desired gender-specific personality traits may be a key driver of gender specific labour market outcomes. Thus this paper assesses the impact of personality traits on performance across gender.

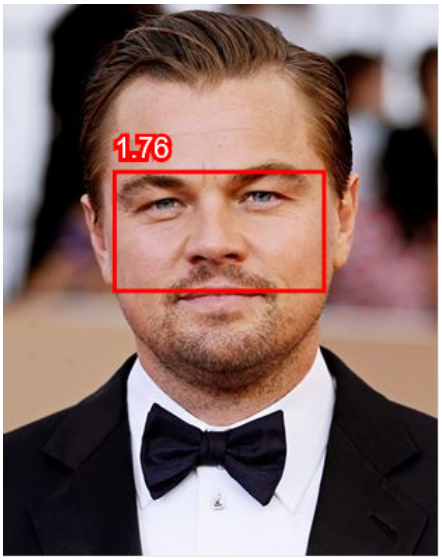
The facial width to height ratio (fWHR) is one physical feature which has been extensively linked to various personality traits such as dominance, aggression, and masculinity. Measuring the width between the ears and dividing it by the length between the eyebrows and upper lips to give a resulting value (as demonstrated in Figure 1), the fWHR is an ideal proxy for personality in that it can be used as a marker for certain heritable personality traits[2] such as competitiveness, risk taking behaviour, aggression, and dominance.[3] This is because the fWHR was once an evolutionary tool used to distinguish certain personality traits among cavemen (Haselhuhn et al, 2013). In this paper, we too use the fWHR as a proxy to measure personality traits. Individuals with values higher than the average are deemed dominant whilst those below the average are classed as non-dominant. Consequently, when assessing the images shown in Figure 1 below, we would class both Leonardo DiCaprio and

[1] Croson and Gneezy, (2009); and Bertrand, (2011) research similar hypotheses.

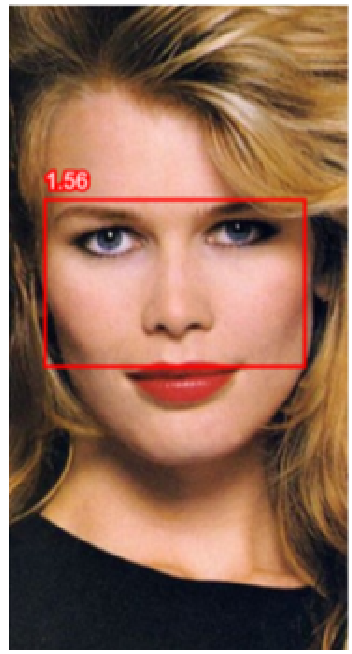
[2] See Eckel and Petrie, (2011).

[3] See Carre and McCormick, (2009); Carre et al. (2010); Niederle, (2017); Schweiser and Karami, (2018); Lefevre et al. (2014); Haselhuhn et al. (2015); Craig et al. (2019); Niederle and Vesterlund, (2007); Niederle and Vesterlund, (2011); and Valentine et al. (2014).

Figure 1: fWHR Examples



(a) Leonardo DiCaprio



(b) Claudia Schiffer

Claudia Schiffer as non-dominant individuals. Accordingly, by using the fWHR as a proxy for personality, we can assess the impact of personality traits on labour market outcomes.

In terms of the labour market outcomes we investigate, we have restricted our sample to academia. Primarily, we have analysed fWHR and academic success in the Economics industry, looking at faculty members in the top 100 US Economic departments. Academic performance is primarily measured through the h-Index, and a number of controls have also been incorporated into the model, including but not limited to variables such as ethnicity and time since PhD (a proxy for age).

The starting point for our analysis is a conceptual framework that links personality to skills and outcomes. To develop this model, we combine the model of cognitive and non-cognitive skill formation (Cunha and Heckman, 2007; and Cunha et al. 2010) with the field of heritability (Turkheimer, 2000; Power and Pluess, 2015; and Roysamb et al. 2018) and stability of personality (Ferguson, 2010; Cobb-Clark and Schurer, 2012; and Harris et al. 2016).

From this model, we derive two key testable hypotheses: more dominant women are less successful compared to non-dominant women, and the opposite relationship for men. We offer two main reasons for this: firstly, dominant men build larger networks while dominant women build smaller networks (McDowell et al. 2007; Ductor et al. 2018; and Lindenlaub and Prummer, 2020), and secondly, dominant women may be discriminated against (Burgess and Borgida, 1999; Maass et al. 2003; Parkins et al. 2006; and Berdahl, 2007) which could limit non-cognitive skill formation and, hence, lead to worse outcomes.

To test these hypotheses, we build a unique and novel data set of all faculty members of the Top 100 US University Economic departments. Data is taken from individual CVs, Google Scholar, and from the web-page picture (for the computation of the fWHR). Our sample covers 87 percent of the faculty members (N = 2,473) and splits into 1,941 males and 532 females.

Our key findings can be summarised as follows. Men with higher fWHRs, ergo more dominant, are more successful with higher h-Indexes. Conversely, women with higher fWHRs, ergo more dominant, are less successful (lower h-Indexes). Furthermore, this pattern appears to emerge over the academic life-cycle, becoming stronger as both men and women move up the academic ranks from assistant to associate to full professor.

We have highlighted a number of channels with which we speculate this may be the case. The first of these is a direct channel pertaining to personality. Personality traits such as risk taking, competitiveness, dominance, and achievement drive have all been associated with success. For males, these qualities are viewed as desirable. Therefore, we would speculate having a higher fWHR which implies an abundance of these qualities would in fact benefit a man for the aforementioned reasons. On the contrary, for females, these traits defy gender norms which favour compassion, nurturing, and empathy.[4] Non-conformity is often discriminated against, which may be a determining factor for our results.

We also highlight two indirect channels. The first of these is networking. Our results suggest that collaborating with three other academics will increase an individual's h-Index by 1 for both females and males. Our data shows dominant males network quite frequently, whilst dominant females do not. We speculate this may be the case for various reasons. Firstly, dominant women may feel less need to have co-authors out of personal preference. This is in line with McDowell et al. (2007); Ductor et al. (2018); and Lindenlaub and Prummer, (2020) who deduce that dominant women prefer to work alone or in smaller networks. Secondly, dominant women are discriminated against, and therefore unable to find willing collaborators. Discrimination has also previously been highlighted as a channel for the persistent gender differences by Burgess and Borgida, (1999); Maass et al. (2003); Parkins et al. (2006); and Berdahl, (2007), corroborating this driving channel. Nonetheless, the decreased networking of dominant women may explain their lower h-Indexes compared to dominant men and non-dominant women.

The paper is structured as follows. Firstly, we discuss existing literature pertaining to measures and drivers of academic success, women in economics, network formation, fWHR in biology and economics, and beauty. Secondly, we outline our model of personality, skills, and outcomes, and thirdly our data and econometric strategy. Next we analyse our main findings, and finally we conduct a number of robustness checks relating to potential sample selection bias, omitted variable bias, and other various robustness checks. Supplementary tables and figures can be found in the appendix.

[4] See Tinsley et al. (2009); Mueller et al. (2006); and Costrich et al. (1975).

2 Literature Review

2.1 Measuring Academic Performance

Measuring academic performance is necessary to inform decisions on hiring, promotion, tenure, grants, and awards. In addition, academic performance is the key outcome variable in this paper. We focus primarily on the h-Index (Hirsch, 2005) as a measure of performance.[5] The h-Index has been widely accepted and used in various fields (Haley, 2017; Geraci et al. 2015; and Delgadillo, 2016) including economics (Ellison, 2013), reflecting the number of papers, H, which have been cited H number of times.[6]

By design, the h-Index gives a single number that balances productivity and impact of a researcher in their field. This is preferable as using a simple publication count could be misleading. On one hand, a researcher could publish many papers which are not cited often or, on the other hand, a researcher could have few papers which are cited often, both of which would be ranked higher than they necessarily ought to be. In addition, the interpretation of the h-Index varies when comparing younger or senior researchers. For younger researchers, the h-Index informs whether the researcher has published papers which have had impact on the field. For senior researchers, the index will ignore papers without a substantial number of citations and only focuses on the number of papers that have had a large effect on the field.

Hilmer et al. (2011) determine the h-Index to be a strong predictor of salary and place of employment. Analysing the income and positions of 1,009 faculty members as well as their

[5]Other supplemental methods such as the use of citations and the i10-Index were assessed in section 6.

[6]For example, an h-index of 23 suggests that the researcher has 23 papers all of which have been cited at least 23 times. To move up a rank, the researcher must possess a 24th paper with 24 citations and all remaining 23 papers must obtain one additional citation each.

publications and citations, they determined the variables to be highly correlated. Research influence, as measured using the h-Index, was found to be an accurate predictor of salary and place of employment in academia, explaining 50 percent of the variation observed in log salaries.

Other papers, however, highlighted potential flaws in the h-Index as a measure of academic performance (see, for example, Ellison, 2013; Geraci et al. 2015; Symonds et al. 2006). Since the h-Index depends on time, it favours senior researchers, as younger researchers (especially in economics) have not had the time to publish 50 papers that have been cited 50 times each (Conley and Önder, 2014). It also favours researchers publishing in fields which have higher citations frequencies (Moed, 2009). Furthermore, the h-Index does not take into account whether the researcher is a single author or is part of a co-author team (Geraci et al. 2015), valuing both at the same rate.

Additionally, recent literature has stressed potential biases related to gender found in use of the h-Index (Geraci et al. 2015; Symonds et al. 2006). Holliday et al. (2014) study male and female academic performance across 82 radiation oncology departments in the US. They find that fewer women achieve senior ranks, but those who do have similar productivity compared to their male counterparts. Geraci et al. (2015) use a sample of 70 male and 70 female randomly selected professors from the Top 100 US Psychology departments. Similarly, they find a significant gender differences in the h-Index and salaries even after controlling for age. Symonds et al. (2006) too assess gender performance differences among a sample of 168 (39 female, 129 male) researchers at British and Australian Life Science University departments. They also find the h-Index to be biased against female researchers. In response, they propose a modified index measure to account for this so-called gender bias labelled research status. This variable is calculated as the residual obtained from regressing the h-Index on the number of publications. When computing the research status for their sample, Symonds et al. (2006) found no gender differences in performance.

2.2 Drivers of Academic Performance

The previous section reviews the literature on our preferred measure of academic performance. In this section, we want to review the literature discussing the driving forces of performance. We abstract from the effect of gender on academic performance in this section, as it is the focus of section 2.3.

The most obvious driver of academic performance (e.g. tenure or promotions) is publishing peer-reviewed articles that attract citations (which increase the h-Index). However, over time the publishing process has changed dramatically. Card and DellaVigna, (2013) analyse the level of annual submissions in economics to the Top 5 economic journals[7], and find that these nearly doubled between 1990 and 2012. This suggests either an increase in desirability to publish in these journals, as suggested by Heckman and Moktan, (2018), or an increase in industry knowledge. They also find that the number of articles published in these journals declined from 400 per year to 300 per year. This in turn drops the acceptance rate from 15 to 6 percent. They argue that this trend is driven, at least partially, by an increase in article length (on average articles are three times longer now), implying, ceteris paribus, that young authors are now less likely to be published, lowering their academic performance. However, averaging about 200 citations each, a publication in a Top 5 Journal is important in increasing one's h-Index.

Hamermesh's (2013) conclusions support these findings and also highlight that the fraction of older authors has increased (by a factor of four). He also documents a trend towards publishing papers which collect data (in the lab or the field) and away from theoretical papers and papers using readily available data.

[7] These are the American Economic Review, Econometrica, the Journal of Political Economy, the Quarterly Journal of Economics, and the Review of Economic Studies.

Ellison, (2002) studies delays in the publication process and tries to understand the causes of these delays. He shows that the submit-accept time at most top journals has increased by 12-18 months over the last 30 years. The reasons for this finding are threefold. First, paper length has increased and longer papers require greater review times. Second, the number of authors per paper has increased and these papers tend to have longer review times. Third, as discussed above, space in top journals became sparser. Along this line, Conley et al. (2011) document that productivity has decreased. Using a sample of 14,271 PhD graduates between 1986 and 2000 in US and Canadian economic departments, they found delays were linked to diminished productivity (measured by the number of AER equivalent publications).

Several other drivers of performance are documented in the literature. Einav and Yariv, (2006) link the effect of surname initials on professional outcomes in economics. They find earlier surname initialled scholars were more likely to receive tenure at top economic departments. These individuals were also more likely to receive awards such as the Nobel prize. They suggest that this is because economics papers typically order co-authors alphabetically. When the model was applied to psychology where co-authors are not alphabetised, there was no effect observed from surname initial on success.

Collins, (2000) also documents racial differences in economic PhDs and the membership of the AEA. She finds that minority individuals are more likely to exit and are more likely to not obtain a PhD compared to non-minority individuals. Similarly, Bayer and Rouse, (2016) document these differences and argue that the absence of diversity constrains the range of topics researched. They suggest that implicit attitudes and institutional practices are the primary catalysts of this problem.

Further, the importance of being linked to an editor of a journal is shown in Brogaard et al. (2014). Using data from 50,000 articles published in 30 economics and finance journals, they find that the editor's colleagues from the same University publish 100 percent more papers in the editor's journal compared to times when the editor was not in office.

Finally, the impact of attending conferences is documented in Gorodnichenko et al. (2019). Using more than 4,000 papers presented at three leading conferences (AEA, EEA, and RES) between 2006 and 2012, they show that presenting at these conferences increases the probability of the paper being published in a high-quality journal. In addition, attendance also increases citations and abstract views of the paper. The last two drivers (editorship and conference attendance) point toward an important channel: networks. Section 2.4 will focus exclusively on the effects of networks on academic performance.

2.3 Women in Economics

The average American woman's socio-economic outcomes have changed dramatically over the past fifty years (Blau and Kahn, 2017; Goldin and Mitchell, 2017; and Lundberg 2020). However, significant differences persist along various dimensions, including the labour market (Goldin et al. 2017; and Charles et al. 2018). Gender differences also exist in economics and have been documented as early as 1974 in a paper by Gordon et al. (1974), who showed that women earned about 11 percent less than corresponding men at an undisclosed University.

Kahn, (1995) documents similarities between men and women when it comes to undergraduate grades, admission rates to PhD programs, first job offers, and publication rates when controlling for rank of the PhD granting University or the current employer (see Hilmer and Hilmer, 2007 for contradicting results). Gender differences are found for GRE scores, application rates for PhD programs, drop-out rates from the PhD, salaries (see also Ginther, 2003; and Blackaby et al. 2005), and promotions. Along this line, Ginther, (2002) shows that women are under-represented in the upper ranks and are less likely to receive tenure. This was corroborated by Buckles, (2019) who finds that less than one third of economics majors are women and the proportion of women in economics decreases as they move towards tenure. Despite progress being made, Lundberg and Stearns, (2019) find that the proportion of women entering the economics market has stalled relative to other disciplines. They also document that the fields men and women conduct their PhD thesis in to be relatively stable over time. Further, they argue that women are held to higher standards than men.

Reasons for these findings are manifold. Boustan, (2019), for example, argues that early career success of women varies across PhD programs. The drivers are number of women in the department, advisor-student contact, and collegial research seminars. Card et al. (2020) show that male referees do not show a bias against female-authored papers and, therefore, rule out a gender bias in the publishing process (see also Abrevaya and Hamermesh, 2012; Hengel, 2017; Hengel, 2019; and Astegiano et al. 2019). In Babcock et al. (2017), experimental evidence is provided showing that women, in contrast to men, are more likely to volunteer for service positions, are more likely to be asked to volunteer for these positions, and are more likely to accept these requests. Since these service jobs crowd-out research time, they reduce the likelihood of promotion. Hilmer and Hilmer, (2010) document differences in the job mobility of men and women by tracking graduating PhD students in the US. While women tend to move downward, men move horizontally or upward. Further, Blackaby et al. (2005) show that men receive more outside offers than women and document perceptions by women to be discriminated against using a survey of UK economists. Leslie et al. (2015) postulate that expectations of brilliance explain the gender gaps in academia, as do Ginther and Kahn, (2015). Operating on the belief that to succeed, you must have an innate raw gift with the subject, cultural and gender beliefs discourage women who are stereotyped to not possess such "brilliance" (as they put it). This decreases female labour force participation. Maths-intensive courses are a prime example of this, with women receiving less PhD qualifications in these areas.

2.4 Network Formation

In the previous sections, we have ignored the importance of networks for performance. However, networks are vital for the exchange of information and the positive spill-over effects they create (See Adams, 2013; Bosquet and Combes, 2017; Borjas and Doran, 2015; and Bailey et al. 2018). Collaborating with others allows specialisation, insures against risks, and increases the number and size of projects.

Card and DellaVigna, (2013) document that the number of co-authors has increased from 1.3 in 1970 to 2.3 in 2012. Azoulay et al. (2010) use the death of a superstar economist as a source of exogenous variation in the co-authorship network. They find that after the death of the superstar, collaborators face a 5-8 percent drop in quality-adjusted publication rates. Along this line, Wuchty et al. (2007) show that teams have changed the way knowledge is generated. Using data on 19.9 million papers and 2.1 million patents from the last 50 years to show that teams dominate single-authored papers, they find networking produces papers which are cited more often and have higher impact. Similarly, Freeman and Huang, (2015) show that increased diversity in teams leads to publications in higher ranked journals and receives more citations. They argue that diversity increases the quality of the paper or the number of people the paper reaches (or both) via accessing different networks. Besides the effect on research impact, Combes et al. (2008) show that networks increase the probability of being hired. Using data from the centralized hiring process for economics professors in France, they show that not being linked to the jury requires a much better publication record as compensation.

In general, networks exhibit homophily, i.e. they are homogeneous according to characteristics such as behaviour, socio-demographic and intra-personal factors (McPherson et al. 2001; Lewis et al. 2012; and Krause et al. 2010). The most important factors pinning down the network are race and ethnicity, followed by other factors such as age, education, and gender (McPherson et al. 2001; Mayer and Puller, 2008; Apfelbaum et al. 2014; Freeman and Huang, 2014; Currarini et al. 2009; and Kerr, 2008). Fafchamps et al. (2010) study co-author relationships in economics over 20 years. They show that the likelihood of collaborating increases if the researchers are closer in an existing network. Mairesse and Turner, (2005) investigate factors that affect the collaboration of authors such as proximity of laboratories, productivity, and publications, finding collaborations within a single laboratory are 40 times higher than collaborations between laboratories in the same town and 100 times higher than collaborations between laboratories in different towns.

Gender differences in network formation of economists were first documented by Ferber and Teiman, (1980); and McDowell and Smith, (1992). They find that economists tend to co-author with colleagues of the same sex. For women, this contributes to a lower number of publications and, consequently, to a lower probability of being promoted compared to men. McDowell et al. (2006) use data from the AEA directories for six time observations from 1964 to 1997. They find that women in the top departments are less likely to co-author. However, when they only consider top journals, they find that women are more likely to co-author. They argue that networks affect the joint decision to co-author and publish. A similar result was obtained by Boschini and Sjögren, (2007), showing that women are more likely to single-author than men. Ductor et al. (2018) also find that women in economics have less co-authors compared with men and they collaborate more often with the same co-authors. Further, they co-author with more senior colleagues. Networks of men and women exhibit homophily. Along this line, Lindenlaub and Prummer, (2020) find that while men have more co-authors (larger networks), women have denser networks. Finally, Agarwal et al. (2016) find in terms of gender, networking with males can help females succeed further.

2.5 fWHR in Biology

The fWHR has been extensively researched in the field of biology and has been linked to various personality traits in humans and animals (See Kachur et al. 2020; and Wilson et al. 2014). This section gives an overview of this literature and key findings.

Bruce et al. (2012) determined that evolution aided in developing recognisable facial cues. These changed in ability to be perceived as the face became more familiar, and could be used to determine criminals and recognition of difficult situations. Similarly McGugin et al. (2013) deduced that facial recognition operated in the same manner as object recognition.

Altschul et al. (2019) study the link between fWHR and dominance in rhesus macaques (Macaca Mulatta). Specifically, they survey sex, age, facial morphology, dominance status, and personality of 109 monkeys from the California National Primate Research Centre. They find higher dominance levels among the primates with higher fWHRs. Additionally, lower fWHRs are correlated with higher confidence levels. They conclude that fWHR may be indicative of aggressive traits and assertiveness. In humans, fWHR has also been recognised as a marker of behavioural traits. Geniole et al. (2015) assess the impacts of fWHR on biology in a sample of 10,853 subjects. They find that on average males have higher fWHR than females. They also document that having a higher fWHR predicts threatening behaviour in men (N = 4,603), and a marker for dominance in both genders (N = 948). Hehman et al. (2015) investigate the overall effect of higher fWHRs on individuals by investigating how fWHR influences group membership selection decisions. Using competition as a proxy for conflict whereby there is a higher desire for dominance and aggression, they find with a sample size of 101 individuals, that higher fWHRs are more likely to be chosen for competitions as it is perceived as a marker of traits advantageous to competition. In contrast, Wang et al. (2019) investigate the relationship between fWHR and antisocial tendencies in a sample of 1,305 people but find only little evidence supporting this relationship.

Zilioli et al. (2015) study fWHR as a signal of increased physical prowess. They determine that fWHR is a facial cue that could be read and interpreted. Their first study, with a sample size of 241, finds that fWHR co-varies with actual physical formidability, while their second study, using 48 UFC fighters, finds fWHR to be a marker of formidability. These findings are corroborated by Sell et al. (2008). Along this line, Lefevre et al. (2014) relate the fWHR to alpha status. Employing a model and a sample size of 103 subjects (49 female), they assess fWHR with respect to masculinity. They find that fWHR can be linked to aggression (physical and verbal), self-reported dominance, and anger though not hostility. Further, there appears to be no difference pertaining to gender. Hence, they conclude that fWHR can be a cue for dominance and aggression in males and females. Haselhuhn et al. (2013) investigate the channels through which fWHR works. They determine that fWHR is linked to dominance via a self-fulfilling prophesy. Higher fWHRs developed evolutionarily as a signal and marker of dominance. This marker persists in the modern world however, causing individuals to react to higher fWHR individuals as if they will be dominant which in turn entices them to act dominantly. Zebrowitz et al. (1998) too found a similar channel, assessing the hypothesis that baby-faced boys compensate for their less dominant expectations by behaving in a childlike manner. They determined baby-faced individuals were typically smarter than mature-faced peers disproving claims of intellectual weakness. Furthermore, these individuals were found to be more likely to be delinquents, contrasting pre-conceived notions that baby-facedness is correlated with warmth, weakness, and submissiveness.

Carrying on in terms of self-perception, Watkins et al. (2010) determined a negative relationship existed between men's own perceived dominance and their perception of other men's dominance. They concluded this suggests less dominant men may be more perceptive of cues of dominance compared to more dominant men. Mileva et al. (2014) however found a positive association between perceptions of others' dominance and their fWHRs in men as well as self-perceived dominance and fWHR.

fWHR has also been found to vary across gender. The average American male's fWHR stands at 1.83 whilst the average American female's at 1.73. Carre and McCormick, (2008) investigate sexual dimorphism with respect to fWHRs, specifically regarding judgements pertaining to gender, emotion, and personality. Assessing variations among males and females, males and varsity hockey players, and males and professional hockey players with sample sizes of 88, 21, and 112 respectively, they find that men typically possess higher fWHRs compared to women. They also find that men possess higher dominance scores and greater reactive aggression. Thus higher fWHRs are positively related to aggressive behaviour, particularly in men.

fWHR has also been determined to vary among ethnicities. Fang et al. (2011) found different races possessed variability in facial dimensions. The most notable of these was the height of the forehead, though measurements of the mouth, eyes, and nose were significant also. No statistically significant difference pertaining to gender was observed.

Zhang et al. (2018) survey total, dyadic, and solitary sexual desire in a sample of 754 women. They find no link, suggesting fWHR is not related to women's sexual desires. They also find no evidence linking a woman's face shape sexual dimorphism to their socio-sexual orientation. Kramer, (2017) also found no evidence to suggest fWHR was a sexually dimorphic measure in skulls or faces.

Valentine et al. (2014) study the effect of fWHR on attractiveness to the opposite gender. Assessing the effect of male attractiveness to females specifically through speed dating, their sample included 78 men and 81 women. In contrast to Zhang et al. (2018), they find a positive correlation between perceived dominance and attractiveness as well as the likelihood of a follow-up date. This relationship persists in the short-term, but not long-term, and suggests higher personal perceived dominance, i.e. higher fWHR, is positively linked to mating desirability.

Finally, research has shown that exposure to testosterone has effects on females. Bütikofer et al. (2019) find that in utero exposure of female twins to testosterone reduces the probability of graduating from high school, completing college, and lowers life-cycle earnings. It needs to be stressed that fWHR and testosterone are not related. Many personality traits correlated with higher fWHRs among men have been speculated to be influenced by pubertal testosterone which contributes to facial structure as well as dominance and is higher in men than women. Furthermore, Bird et al. (2016) examine the relationship between fWHR, baseline testosterone, and competition-induced testosterone using a sample size of 780 men. They find no statistically significant relationship between any of the investigated variables.

2.6 fWHR in Economics

The fWHR - as a proxy variable for personality - has received recent attention in economics and finance. As the fWHR proxies for characteristics such as dominance, risk preferences, and competitiveness, we expect to see an effect of fWHR in certain fields.

He et al. (2019) link fWHR to achievement drive among 1,744 Chinese male financial analysts. They find analysts with higher fWHRs are more likely to exhibit improved performance. This relationship is more prominent among lower-status analysts compared to high-status analysts, as well as in firms with greater levels of uncertainty and in analysts with greater instances of competition. Similarly, Schweiser and Karami, (2018) assess the effect of fWHR among hedge fund managers. They focus on their risk-taking tendencies in a sample of 7,549 individuals, finding a positive correlation with fWHR and risk-taking behaviour. Further, Lu and Teo, (2018) link the impact of fWHR on hedge fund manager performance. In particular, they find hedge fund managers with lower fWHR outperform those with higher fWHR. They argue this may be because higher fWHR managers are more likely to terminate their funds, disclose violations, and perform with greater operational risk. High fWHR individuals in the sample are also more reluctant to sell under-performing stocks, suggesting personality traits associated with higher fWHR are not beneficial in this field. Kausel et al. (2018) suspect that whilst the aforementioned factors impact academic success, personality traits are an additional contributing factor. They find a link between fWHR and academic success through the channel of assertiveness. Using a sample of 231 students, they determine that fWHR accurately predicts academic performance in non-quantitative courses. It does not, however, predict it in either basic nor applied quantitative courses. Finally, Hahn et al. (2017) assess the effect of fWHR in companies. They find in a number of organisations, both profit and non-profit, fWHR increased positively with rank in the organisation. Higher fWHRs are also positively correlated with company's donations to charitable causes and environmental awareness campaigns. This suggests fWHR is linked to leadership capabilities and social rank, potentially through the mechanisms of dominance and aggression which could aid in success. Lin et al. (2018) also links fWHR to political corruption. In a sample of 325, they show images of unfamiliar politicians to the subjects and ask for behavioural perceptions. On average, those accused of corruption in real life are perceived in the study as more corrupt, aggressive, and dishonest. These individuals, however, are also discerned to be more masculine, competent, and ambitious. Additionally, when the facial structure of these individuals was adjusted, the corruptible perceptions change. This suggests perceived social behaviour is impacted by fWHR.

Haselhuhn et al. (2014) also related fWHR to negotiation performance in men. They found that more dominant men, those with higher fWHRs, were less cooperative negotiators meaning they were firmer in getting their wants, but less likely to engage in fair compromise than men with lower fWHRs.

Lewis et al. (2012) too assessed the behavioural traits associated with fWHR among former US presidents. They found a positive association between fWHR and achievement drive suggesting indirect correlations to dominance and aggression.

2.7 Beauty

There has been no research linking fWHR directly with beauty, but there is research showing the effect of beauty on labour market outcomes. Overall, it is speculated people benefit from being more attractive, though women may at times face a beauty penalty (Johnson et al. (2010).

Hamermesh and Biddle, (1994) investigate the impact of beauty on earnings. Using data about beauty obtained from interviewers' opinions of a sample of 5,000 survey respondents, they find a positive relationship between attractiveness and earnings, with perceived plainness costing individuals approximately 5-10 percent of their income. They find these effects are consistent across gender. More so, unattractive women have lower-labour force participation rates and the impact of attractiveness is independent of occupation, suggesting exogenous employer discrimination. Biddle and Hamermesh, (1995) confirm these results by assessing the impact of beauty on earnings of 3,750 lawyers. They find better looking attorneys earn more than less attractive attorneys. Along this line, Andreoni and Petrie, (2007) show that there is a beauty premium in most workforces, however this becomes a penalty when cooperation is explored. More attractive individuals are perceived to be more cooperative, but when this is not the case, they appear to be more selfish. In contrast to these findings, Kanazawa and Still, (2018) document that the very unattractive earn more than the highly attractive. Sparacino and Hansell, (1979), however, find no link between attractiveness and academic success, which is corroborated by Talamas et al. (2018).

There also exists a related literature documenting the effects of make-up or grooming on labour market outcomes. Hamermesh et al. (2002) find that the purchase of beauty products increases women's earnings by increasing women's perceived attractiveness. Palumbo et al.

(2017) link high self-esteem to greater academic performance. With previous literature suggesting self-esteem could be raised through the use of make-up, they find that women wearing make-up scored higher on tests than those without make-up. This suggests beauty, by raising self-esteem, positively impacts academic performance. Similarly, Wong and Penner, (2016) find that attractive individuals earn 20 percent more than those of average attractiveness, but this relationship lessens when grooming is controlled for.

In academia, Fidrmuc and Paphawasit, (2018) investigate the impact of physical attractiveness on productivity. Removing any impact of physical perceptions of beauty by looking at academic publications, they assessed if attractiveness is positively correlated with publications using a sample of 2,000 academics. They find a positive relationship between attractiveness and both journal quality and number of citations. Liu et al. (2018) investigate the effect of beauty on academic career success using a sample of professors at the top 50 US business schools. They find that more attractive professors find better first jobs and receive tenure earlier. The effects of beauty from associate to full professor, however, are insignificant. They argue that beauty is a proxy for intelligence and social competency.

3 A Model of Personality, Skills, and Outcomes

In this section, we want to introduce a framework to conceptualize our thinking about the effect of personality or personality traits on skill formation and outcomes.[8] We further formulate how skills evolve over time.[9]

We combine the model of cognitive and non-cognitive skill formation (Cunha and Heckman, 2007; Cunha et al. 2010; and Heckman and Mosso, 2014) with the work by Heckman et

[8]In the psychology literature, "traits" are relatively stable patterns of behaviour, thought, and emotion (see Roberts, 2009).

[9]Notice that we could also use the term "ability" rather than "skill" as they are interchangeable (Cunha and Heckman, 2007).

al. (2019) on personality psychology and the field of heritability (Turkheimer, 2000; Power and Pluess, 2015; Roysamb et al. 2018; and Almlund et al.(2011); and McAdams and Pals, (2006)) and stability of personality traits (Ferguson, 2010; Cobb-Clark and Schurer, 2012; and Harris et al. 2016). For cognitive skills, Bouchard et al. (1990); Devlin et al. (1997); Burt, (2008); Fletcher, (2013); and Plomin, (1999) show that general cognitive ability and IQ are heritable. For example, Bouchard et al. (1990), using a twin-study design, find that about 70 percent of IQ is explained by genes. Devlin et al. (1997) however find this to explain less than 50 percent of the variation. Similarly, Fowler et al. (2009) show that people are endowed with traits which affect network attributes, a non-cognitive skill. Importantly, these traits are heritable and genes explain about 46 percent of the variation in in-degree, 47 percent in node transitivity, and 29 percent of betweenness centrality. Further, in contrast to the modelling in the existing literature, we allow for gender differences in our model.

We begin by defining the vector of skills $\Phi_{i,t}$ at any point in time t. Time in the model could refer to age or the time since completing the PhD and joining a University at a junior position. The index i indicates whether the person is male, M, or female, F. Research has shown that gender differences in personality traits exist (Schmitt et al. 2008; Weisberg et al. 2011; Del Giudice et al. 2012; Braakmann, 2009; Buchan et al. 2008; and Flinn et al. 2018), suggesting that distinguishing between personality traits is important in our model for two reasons. Firstly, we argue that gender differences will affect the accumulation of skills, for example by affecting the formation of social and collaborative networks (McDowell et al. 2007; Boschini and Sjögren, 2007; Blau and Kahn, 2017; Bertrand, 2018; Lindenlaub and Prummer, 2020), and, hence, outcomes. Secondly, it allows for discrimination against people who violate (perceived) gender norms and ideals. This is relevant in that extensive literature documents that women who do not act according to gender norms or who do not fit gender ideals are discriminated against (Burgess and Borgida, 1999; Maass et al. 2003; Parkins et al. 2006; and Berdahl, 2007). Further, Blau and Kahn (2017) as well as Bertrand (2018) stress personality traits to be driving forces of gender differences in labour market outcomes. They identify differences in risk and time preferences, competitiveness, self-esteem, self-confidence, social norms, and gender identity as crucial factors to explain part of the existing gender gaps.

The skill vector $\Phi_{i,t}$ is multidimensional and we assume that it depends on cognitive, $\Phi_{i,t}^{C}$, and non-cognitive, $\Phi_{i,t}^{N}$, skills as in Cunha and Heckman (2007) and Cunha et al. (2010), such that

$$\forall t, i \in \{M, F\} : \Phi_{i,t} = \left(\Phi_{i,t}^{C}, \Phi_{i,t}^{N}\right). \tag{1}$$

The vector of cognitive skills covers IQ and inherent talents, while the vector of non-cognitive skills covers personality traits that can vary over time, such as neuroticism and agreeableness (Harris et al. 2016). Skills can vary with the accumulation of experience and people can acquire new skills as they get older, such that the dimensionality of these vectors can change over time (Heckman and Rubinstein, 2001; Heckman and Raut, 2016; and Heckman et al. 2013).

In contrast to the modelling approach in Cunha and Heckman, (2007); Cunha et al. (2010); Heckman and Mosso, (2014); and Heckman et al. (2019), we explicitly consider stable personality traits, which do not vary over time and are inherited rather than a product of the environment. Therefore, we distinguish between time-varying non-cognitive skills captured in the vector $\Phi_{i,t}^{N}$ and a stable component, P_i, which is time-invariant and which we refer to as personality (traits). Examples of the stable component include openness, conscientiousness, and extraversion (Cobb-Clark and Schurer, 2012; and Harris et al. 2016), as well as risk (Schildberg-Hörisch, 2018) and time preferences (Meier and Sprenger, 2015). Further, Bouchard et al. (1990) show that personality traits - to some degree - are heritable.[10]

Having discussed the two skill vectors, we define the skill formation technology that

[10] A different approach to model the stable, heritable component is to make assumptions about the initial values of the skill vectors. This would leave our results unaffected but make the notation more complicated.

determines how cognitive and non-cognitive skills evolve over time. Following Cunha and Heckman, (2007); and Cunha et al. (2010), we assume that the technology depends on the stock of cognitive and non-cognitive skills, personality, and investment, I, into specific skills. Formally,

$$\forall t, i \in \{M, F\}, j \in \{C, N\} : \Phi^j_{i,t+1} = f^j_{i,t}\left(\Phi^C_{i,t}, \Phi^N_{i,t}, P_i, I^j_{i,t}\right), \tag{2}$$

where $\Phi^C_{i,1}$ and $\Phi^N_{i,1}$ are given initial values. Further, the function f^j is monotone increasing in all its arguments, twice continuously differentiable, and concave. We think about the investment into cognitive skills as studying and learning new methods or subjects, and for non-cognitive skills as general personal development.[11] This formulation of skill formation technology allows for two crucial features. First, self-productivity, where skills learned at time t augment skills learned later. Second, there is cross-fertilization, i.e. cognitive skills affect the production of non-cognitive skills and vice versa.

This technology could be modelled as stage specific by introducing an index s as in Cunha et al. (2010), to allow the undergraduate or postgraduate periods to be critical periods of skill development. For example, investments during these time periods could have higher returns compared to investments later in the career. Since this is beyond the scope of our paper, we abstract from this extension.

Skills and personality are important drivers of outcomes, Y, in the model (Borghans et al. 2008; Dohmen et al. 2010; Heckman et al. 2019). However, outcomes also depend on effort and incentives to perform. Outcomes are produced according to

$$\forall m, t, i \in \{M, F\} : Y_{m,i,t} = g_{m,i,t}\left(\Phi^C_{i,t}, \Phi^N_{i,t}, P_i, e_{m,i,t}\right), \tag{3}$$

where m indicates the specific outcome, for example publishing articles, applying for promotion, or applying for external grants. The function g maps inputs to outcomes and is

[11] For example affected by life-events such as marriage or giving birth.

monotone increasing, concave in all arguments, and twice continuously differentiable. This formulation implies that skills can affect different outcomes with different weights. For example, cognitive skills might increase the probability to publish in a Top 5 journal, but non-cognitive skills might affect the probability of being promoted via, for example, network formation. This approach also allows to compensate for any shortcomings in achieving an outcome. A deficit in one dimension can be compensated for by an abundance of another dimension, e.g. motivation or effort.

Further, it is possible to generalize equation (3) and allow past outcomes to affect contemporaneous outcomes. For example, publishing an article in a Top 5 journal today could affect the probability of publishing in the same journal again in the future. In addition, publishing in a Top 5 journal could also increase the probability of being promoted in the future. This implies that there is a dynamic cross-fertilization between outcomes such that equation (3) could be written as

$$\forall m, t, i \in \{M, F\} : Y_{m,i,t} = g_{m,i,t}\left(\Phi^C_{i,t}, \Phi^N_{i,t}, P_i, e_{m,i,t}, Y_{m,i,t-l}, Y_{n,i,t-l}\right), \tag{4}$$

where l indicates the lag with which past outcomes affect contemporaneous outcomes and we have included the possibility that outcome n, e.g. receiving a grant, affects outcome m, e.g. publishing in a Top 5 Journal.

Effort is a function of skills, personality, incentives, and preferences as in Heckman et al. (2019). Incentives, R, are provided by the University or the market, e.g. job offers from other Universities. Incentives depend on the information about the uncertain return. For example, allocating effort into a research project where the outcome is highly uncertain, e.g. because it is unclear whether results will be obtained at all or where the paper will be published, might not be optimal given multiple available projects with different risk-reward profiles.

The effort supply function can be written as

$$\forall m, t, i \in \{M, F\} : e_{m,i,t} = h_{m,i,t}\left(\Phi_{i,t}^{C}, \Phi_{i,t}^{N}, P_i, \mathbb{E}_{i,t}\, R_{m,i,t}\left(\mathcal{I}_{m,i,t-1}\right) \mid \Theta_i\right), \tag{5}$$

where $\mathbb{E}$ denotes the mathematical expectation operator and $\mathcal{I}_{m,i,t-1}$ denotes the information set at time $t-1$ about the anticipated return of outcome m in period t. The function h links the inputs into the effort function to the task-level effort. The function is monotone increasing, concave in all arguments, and twice continuously differentiable. Further, following the approach by Heckman and Mosso, (2014) effort also depends on the set of preferences, Θ_i.

Given the assumptions on the skill formation technology, the effect of skills on outcomes is given by

$$\forall t, i \in \{M, F\}, j \in \{C, N\} : \frac{\partial Y_{m,i,t}}{\partial \Phi_{i,t}^{j}} > 0, \tag{6}$$

such that higher cognitive and non-cognitive skills will increase outcomes. While this effect is non-linear and exhibits decreasing marginal returns, the sign of the effect will be positive. The effect of personality on non-cognitive skills, e.g. the ability to form networks via openness or dominance, however is more interesting and at the heart of our model. We argue that men, who are more risk-taking, dominant, and who have a higher achievement drive will be more successful partially by acquiring greater non-cognitive skills, e.g. having larger networks (McDowell et al. 2007; Ductor et al. 2018; and Lindenlaub and Prummer, 2020). However, for women who do not act according to gender norms or who do not fit gender ideals, i.e. who are more risk-taking, dominant, and who have a higher achievement drive; could be less successful. There are various reasons for this. Firstly, they may endogenously invest less time and effort into building networks, preferring to work alone or in smaller networks (McDowell et al. 2007; Ductor et al. 2018; and Lindenlaub and Prummer, 2020). Secondly, they may prefer to work alone and make other decisions (e.g. where to send a paper for publication or which conferences to attend; Blau et al. 2010) due to different levels of risk aversion. Thirdly, they may be discriminated against, which would limit their ability to accumulate non-cognitive skills (Burgess and Borgida, 1999; Maass et al. 2003; Parkins et al. 2006; and Berdahl, 2007). Discrimination against women in economics has been documented, for example, by Blackaby et al. (2005); Blau and Kahn, (2017); and Bertrand (2018) who stress the role of personality traits and norms for non-cognitive skill formation is a viable explanation for gender gaps more generally.

Therefore, the effect of personality on non-cognitive skills depends on gender and is given by

$$\forall t : \frac{\partial \Phi_{i,t}^{N}}{\partial P_i} = \begin{cases} > 0 & \text{if } i = M, \\ < 0 & \text{if } i = F. \end{cases} \tag{7}$$

such that an increase in P is interpreted as becoming more dominant. Thus, this implies that the effect of changes in personality - via non-cognitive skills - on outcomes depends on gender. In our empirical analysis, we will exploit cross-sectional variation in outcomes, personality traits, and confounding variables to investigate the effect of personality on outcomes by gender.

4 Data and Econometric Strategy

4.1 Data Set Construction

The data set used in our analysis consists of information from all faculty members of the top 100 US (University) Economic departments. Universities are ranked according to the ideas.repec.org ranking of the top 25 percent of US Economics departments as of September 2020, and the data set was collected from July to November 2020.

We exclude economists working at Federal Reserve Banks, international organizations (e.g. IMF or World Bank), or private companies (e.g. Microsoft or Google) in the US. There are three reasons to exclude them. First, within the University world, the expectation is that faculty members contribute to (i) teaching, (ii) research, and (iii) service jobs. This will be different outside the University world and would affect the behaviour of individuals. Second, employers might set different incentives and, third, might measure performance differently. Therefore, while these exclusions reduce our sample size, they keep the sample homogeneous with respect to the expectations and incentives of employers and the measurement of performance. For similar arguments, we have excluded other countries and academic disciplines too.

Further, we define a "faculty member" as an assistant, associate, or full professor. This excludes adjunct, visiting, or emeritus professors and all teaching positions as well as Post-Doctorates and PhD students.[12]

Given the large variation in the structure of CVs, it is impossible to use a data scraping algorithm and we, therefore, collected the data manually. We collect a list of faculty members from the departments web-page. From this web-page, we gain access to each faculty member's CV. Since faculty members often have a private web-page, we check whether the CV on the departments web-page or the private web-page is more recent and use the more up to date document.[13] We exclude all CVs older than five years, as information is likely to be outdated. 1,497 (60.6 percent) of the total observations were from CVs most recently updated in 2020. 564 (22.8 percent) were from 2019, 210 (8.5 percent) from 2018, 89 (3.6 percent) from 2017, 64 (2.6 percent) from 2016, and 47 (1.9 percent) from CVs most recently updated in 2015.

For each faculty member we record the following information: rank, gender, ethnicity, nationality, time since completing a Bachelor degree, time since completing a PhD, PhD

[12] One individual in our sample was visiting an University and appeared in the faculty list of two Universities. We allocated this individual to the "home" University.

[13] For a very small number of faculty members we can not find CVs on the departments web-page or a private web-page and search for other web-pages with a CV. Often, we find a CV on web-pages of networks or institutions the faculty member is associated with, for example, the IZA network.

granting University, first job after completing the PhD, number of job switches since the first job, whether the individual holds or has held an editorial position, number of grants won, number of publications, number of Top 5 publications, number of publications in Nature and Science, number of distinct co-authors, research field, whether the research is theoretical (or empirical), and the date of the CV.

Some clarifications are in order. Importantly, to be counted as a publication, the output needs to be peer-reviewed and we exclude, for example, books, policy reports, and Op-Eds. We also exclude any reprints of published outcomes. Further, the AER:Papers & Proceedings and the AER:Insights do not count as Top 5. Papers conditionally accepted count as publications, but papers in revise and resubmit do not. The publications in Nature and Science only refer to the flagship journals and do not include the field journals, e.g. Science Advances or Nature Food.

Further, we do not collect information on age, as we found that date of birth is rarely stated in the CV. Instead we use time since bachelors and PhD as proxies. Ethnicity can take on four values: white, black, Asian, or Hispanic and is inferred from the picture, as ethnicity is typically not stated in the CV. While nationality is often stated on the CV, if this is not the case, we assume that it is identical to the location where the person obtained their Bachelor's degree.

We would prefer to collect data on the dollar value of grants received but, unfortunately, we found that these values are most often not stated. Further, and to make matters worse, we also observe that some faculty members state dollar values of only a subset of grants, i.e. they select which values to reveal. This runs into a selection problem, which we cannot reasonably address. Therefore, we only collect information on the total number of grants allocated to the individuals, as listed in their CVs.

Job switches are not promotions, rather they count how often an individual switches the employer. Often, these are horizontal movements, e.g. people moving from an assistant professor position at one University to another assistant professor position at a different

University. The research field was also recorded and can take three values: macroeconomics, microeconomics, and econometrics. We found that is impractical to have a finer definition of the research field as most people work in different sub-fields, for example, labour and health or macroeconomics and development at the same time. Therefore, making a decision on the sub-field would introduce noise and we therefore only rely on more precise information about the - broader - research field.
The date of the CV refers to the year of the last update and we either take this information directly from the "last updated" information provided in the CV, from the web-link, the name of the CV file, or from the last year that occurs in the CV, typically in the list of publications (excluding, e.g., grants or editorial positions that could potentially run into the future).

We then collect the following data from Google Scholar for each individual: citations, h-Index, and i10-Index. Notice that we only use Google Scholar and do not use other programs, such as SCOPUS, to collect this information. The reason for this is that the different programs can produce different values due to different search algorithms and databases. Hence, using both programs would create a bias in our measures of performance.

We extract three types of information from the picture of each individual: fWHR, beauty, and whether the person appears overweight.

Typically, we use the photo supplied on the University website to calculate the fWHR. However, if a photo was missing or if in the photo, the individual was on an angle, had their nose facing away from the lens, a tilted face, unaligned eyes, or was blurry, we searched for an alternative photo.[14] Using the picture on the private web-page or a Google image search gave a range of potential pictures. We then selected the picture which was most suitable (i.e. fulfils the requirements of the software: high quality, nose pointing towards the camera,

[14]Sometimes a problem was able to be fixed by rotating the entire picture if, for example, the original picture was taken on an angle.

aligned eyes, no angle) for the software to compute the fWHR.[15] We have compared the fWHR from different suitable pictures to reduce the likelihood of selecting a picture that does not represent the true fWHR. If a suitable photo was not able to be found, the subject was excluded from the sample.

Beauty is measured as in Hamermesh and Biddle, (1994); Biddle and Hamermesh, (1998); Hamermesh et al. (2002); Andreoni and Petrie, (2008); French et al. (2009); and Kanazawa and Still, (2018). As shown in this literature, there exists an agreed upon standard of beauty in a society at a given point in time which, additionally, is stable over time, i.e. changes only gradually. Accordingly, we assign a beauty value to each individual, which can take on five potential outcomes: 1 (Strikingly Beautiful), 2 (Above Average), 3 (Average), 4 (Below Average) and 5 (Homely). In the econometric analysis, in line with the literature, we only consider three potential outcomes: Above average (1+2), average (3), and below average (4+5). The reason is that few people are typically rated as either strikingly beautiful or homely (Hamermesh and Biddle, 1994).

We acknowledge that this measure of beauty is not ideal, as it only captures facial features and does not allow to take other features into account. However, relying on a picture also has an advantage, as the measure of beauty is not contaminated by other information usually obtained while meeting someone in real life, e.g. behaviour or socio-economic status (Biddle and Hamermesh, 1998). Nevertheless, our measure of beauty will be noisy. This noise is unlikely to be systematically related to our outcome variable and, therefore, should not affect our econometric analysis. Given the picture, we also create a dummy variable which classifies an individual as overweight (e.g. Hamermesh and Biddle, 1994).

Overall, the Top 100 Departments in our sample employ 2,838 economists of which 2,266

[15] We use the Python software package by de Kok, (2018) to compute the fWHR. `https://github.com/TiesdeKok/fWHR_calculator`

are male (80 %) and 572 are female (20 %).[16] Our data set covers 87 percent of these faculty members ($N = 2,473$). Our sample comprises 1,941 males (78 %) and 532 females (22 %). Put differently, we cover 86 % of male faculty members and 93 % of female faculty members.

The missing observations occur for various reasons. First, we exclude CVs older than five years. Second, some web-links are broken and it is not possible to find a CV elsewhere. Third, if it is not possible to find a suitable picture, we also exclude the faculty member. Therefore, we lose 365 individuals (Male: 325, Female: 40).

Finally, it is important to remember that the decision to have a Google Scholar account is voluntary and not every faculty member has an account. In our data set, 2,016 faculty members chose to have an account (Total: 82 %, Male: 83 %, Female: 77 %). This selection introduces a potentially important source of selection bias which we address in the robustness section using a Heckman selection model.

4.2 Descriptive Statistics

As shown in table 1, the male population had an average fWHR of 1.70. With respect to the female population, the average fWHR was 1.67.

In terms of ethnicity, the male population consisted of 76 percent white, 2 percent black, 17 percent Asian, and 5 percent Hispanic. Similarly, the female population was 71 percent white, 2 percent black, 24 percent Asian, and 3 percent Hispanic.

With respect to rank, 26 percent of the observed men were assistant professors as were 39 percent of the women. 19 percent of the men were associate professors and 23 percent of the women. Finally, over half of the male population, 56 percent, were full professors while less

[16] Tables 10, 11, and 12 in the appendix present the coverage by University.

than half of the female population, 38 percent, carried the same rank. Of the male sample, 39 percent worked in applied microeconomics, 23 percent in theoretical microeconomics, 30 percent in macroeconomics, and 8 percent in econometrics. In the female sample, 60 percent, 14 percent, 19 percent, and 7 percent worked in the same fields respectively. Overall, 32 percent of the women observed worked in theoretical economics compared to 50 percent of the male population.

In the male population, the average time since completing their PhD was 19.85 years meaning the average male age in the sample, assuming their PhD was completed by age 26, is approximately 46. In the female population, the average time since completing their PhD was 13.93 years meaning the average female age in the sample using the previous assumption is approximately 40. This suggests on average that males in the economics field are typically older than females. Furthermore, the sample shows that on average male economists collaborate with 22.10 co-authors, whereas female economists collaborate with 12.72, suggesting that on average, women collaborate with 9.38 less authors than men. Further, women appear to switch jobs less frequently than men with an average number of job switches of 0.88 for the female population and 1.05 for the male population. Similarly, the female population's average number of grants was 5.19 whereas the men's was 5.42. Furthermore, 47 percent of the female sample held an editorial position in her career compared to 57 percent of the male population.

With respect to education, 53 percent of both populations had a bachelors degree other than economics, and 5 percent of both samples had PhDs differing from economics as well. 46 percent of women observed received their PhDs from the top 10 universities, and 63 percent from the top 20. Similarly, 45 percent of men received their PhDs from the top 10 universities, and 60 percent from the top 20.[17] Following on from education, 12 percent of

[17] Table 13 provides these statistics.

the women sampled worked their first jobs at a top 10 university, as did 14 percent of the men observed.

Each researcher was also assessed with respect to their appearance. Of those observed, 3 percent of the female population were deemed overweight compared to 7 percent of the male population. In regards to perceived beauty, 18 percent of the women observed were categorised as of above average attractiveness, 46 percent average attractiveness, and 36 percent below average attractiveness. In the male population, these percentages were 12 percent, 63 percent, and 24 percent respectively. Essentially, half of both sub-samples were deemed as average, and a larger proportion of the remaining two groups were rated above-average as opposed to below-average. Our observations also disperse women predominantly in the middle category, consistent with much social-psychological literature which states that women's appearances evoke stronger reactions, both positive and negative, than men's (Hatfield and Sprecher, 1986).

With respect to beauty, the correlation with fWHR in men was 0.03 and 0.02 for women. A t-test comparing beauty by dominance found insignificant results for both females and males. Additionally, Spearman and Kendall rank correlation tests found no statistically significant correlation.

In regards to performance measures, table 1 further shows that for the female population, the average number of publications is 16.92 in the sample of 531. We find that within the 410 Google Scholar observations, women are cited on average 2,930.62 times. In the same population, the average h-Index for women is 15.33, and the average i10-index[18] and m-Index[19] are 21.16 and 1.14 respectively. Further, on average, a woman has 1.53 papers published in a top 5 publication, and 0.09 publications in Science or Nature.

[18] The number of publications with at least 10 citations.

[19] The h-Index divided by the number of years since the scientists' first paper was published.

For the male population, the average number of publications is 35.21 in the sample of 1,940. We also find men are cited on average 7,717 times. Of those with Google Scholar accounts (obs = 1,604), the average h-Index for men is 23.53, and the average i10-index and m-Index are 41.13 and 1.31 respectively. Further, on average, a man has 3.66 papers published in a top 5 publication, and 0.16 publications in Science or Nature.

Table 1: Descriptive Statistics.

	Female					Male				
Variable	Mean	Std. Dev.	Min	Max	Obs.	Mean	Std. Dev.	Min	Max	Obs.
fWHR	1.67	0.09	1.37	1.96	531	1.70	0.09	1.41	2.19	1,940
Ethnicity										
White	0.71	0.45	0	1	531	0.76	0.43	0	1	1,940
Black	0.02	0.14	0	1	531	0.02	0.13	0	1	1,940
Asian	0.24	0.43	0	1	531	0.17	0.37	0	1	1,940
Hispanic	0.03	0.17	0	1	531	0.05	0.22	0	1	1,940
Time since PhD	13.93	10.86	0	48	521	19.85	13.86	0	65	1,921
Co-Authors	12.72	15.09	0	120	531	22.10	30.06	0	400	1,940
Switches	0.88	1.00	0	5	526	1.05	1.17	0	8	1,916
Grants	5.19	6.95	0	42	530	5.42	8.10	0	102	1,931
Editor	0.47	0.50	0	1	530	0.57	0.50	0	1	1,938
Theory	0.32	0.47	0	1	531	0.50	0.50	0	1	1,940
Bachelor Different	0.53	0.50	0	1	524	0.53	0.50	0	1	1,916
PhD Different	0.05	0.22	0	1	530	0.05	0.23	0	1	1,935
PhD Top 10	0.46	0.50	0	1	531	0.45	0.50	0	1	1,940
PhD Top 20	0.63	0.48	0	1	531	0.60	0.49	0	1	1,940
First Job Top 10	0.12	0.32	0	1	531	0.14	0.35	0	1	1,940
Field										
Applied Micro	0.60	0.49	0	1	531	0.39	0.49	0	1	1,940
Theoretical Micro	0.14	0.34	0	1	531	0.23	0.42	0	1	1,940
Macro	0.19	0.40	0	1	531	0.30	0.46	0	1	1,940
Econometrics	0.07	0.25	0	1	531	0.08	0.28	0	1	1,940
Beauty										
Above Average	0.18	0.39	0	1	531	0.12	0.33	0	1	1,940
Average	0.46	0.50	0	1	531	0.63	0.48	0	1	1,940
Below Average	0.36	0.48	0	1	531	0.24	0.43	0	1	1,940
Overweight	0.03	0.17	0	1	531	0.07	0.25	0	1	1,940
Rank										
Assistant Prof.	0.39	0.49	0	1	531	0.26	0.44	0	1	1,940
Associate Prof.	0.23	0.42	0	1	531	0.19	0.39	0	1	1,940
Professor	0.38	0.49	0	1	531	0.56	0.50	0	1	1,940
Performance										
Publications	16.92	22.34	0	266	531	35.21	46.71	0	462	1,940
Citations	2,930.62	5,769.67	2	67,684	410	7,717.00	19,988.87	1	331,721	1,940
h-Index	15.33	13.20	1	89	410	23,53	22.22	1	221	1,604
i10-Index	21.16	26.68	0	193	410	41.13	66.35	0	1,262	1,604
m-Index	1.14	0.76	0.2	10	397	1.31	0.81	0.04	8	1,576
Top 5	1.53	2.79	0	21	531	3.66	6.50	0	94	1,940
Top 5 Share	0.11	0.18	0	1	483	0.13	0.18	0	1	1,868
SN	0.09	0.51	0	6	531	0.16	0.83	0	14	1,940
Research Status	0	8.19	-77.05	50.49	410	0	12.99	-120.48	136.49	1,604

From table 2's t-tests, by gender results, we observe a number of statistically significant differences between the sub-samples. We observe that on average women have a fWHR 0.04 lower than their male counterparts. This is statistically significant at the 1 percent level, and suggests the female sample tends to be less dominant than the male sample. Furthermore, females on average have 5.92 years less experience than males, and 9.38 less collaborators. This suggests in our sample, the women are younger, and less likely to network. Both results again are statistically significant at the 1 percent significance level. On average, females also switch jobs less, are less likely to adopt an editorial position during their careers, and are less likely to research theoretical economics than their male co-workers, with all differences statistically significant at the 1 percent level. The female subjects in the sample are also less likely on average to be white (statistically significant at the 5 percent level). In terms of the economic field each subject researches within, our t-tests find females are more likely on average to research applied microeconomics, and less likely to research theoretical microeconomics or macroeconomics; all results statistically significant at the 1 percent level. Importantly, we do not find statistically significant differences among those receiving a PhD from a top 10 or top 20 department, nor a first job at a top 10 department. This indicates that there is no selection bias before starting the career.

In regards to variables that mark success, we deduce that on average women have 18.29 less publications than their male counterparts, 4,786.38 less citations, an h-Index 8.2 lower,[20] an i10-index 19.97 lower, and an m-index 0.17 lower. Further, women have 2.13 less top 5 publications and 0.07 less Science and Nature publications than the male population also. All of these results are statistically significant at the 1 percent level except for the number of publications in Science and Nature which is significant at the 5 percent level.

Table 3 splits the female population into two groups, dominant and non-dominant pertaining to their fWHR rankings. Those greater than the sample average of 1.67 (244 observations overall) were classed as dominant. Those less than or equal to 1.67 (287 observations)

[20]Refer to Tables 13, 14, and 15 in appendix. When comparing these rankings, the implications of these suggest that Esther Duflo (89) highest ranked female would be at rank 30 on the male ranking

Table 2: Descriptive Statistics - T-Tests by Gender.

	Difference	Female	Obs.	Male	Obs.
fWHR	-0.04***	1.67 (0.09)	531	1.70 (0.09)	1,940
White	-0.05**	0.71 (0.45)	531	0.76 (0.43)	1,940
Time since PhD	-5.92***	13.93 (10.86)	521	19.85 (13.86)	1,921
Co-Authors	-9.38***	12.72 (15.09)	531	22.10 (30.06)	1,940
Switches	-0.17***	0.88 (1.00)	526	1.05 (1.17)	1,916
Grants	-0.24	5.19 (6.95)	530	5.42 (8.10)	1,931
Editor	-0.10***	0.47 (0.50)	530	0.57 (0.50)	1,938
Theory	-0.19***	0.32 (0.47)	531	0.50 (0.50)	1,940
PhD Top 10	0.01	0.46 (0.50)	531	0.45 (0.50)	1,940
PhD Top 20	0.03	0.63 (0.48)	531	0.60 (0.49)	1,940
First Job Top 10	-0.02	0.12 (0.32)	531	0.14 (0.35)	1,940
Field					
Applied Micro	0.20***	0.60 (0.49)	531	0.39 (0.49)	1,940
Theoretical Micro	-0.09***	0.14 (0.34)	531	0.23 (0.42)	1,940
Macro	-0.10***	0.19 (0.40)	531	0.30 (0.46)	1,940
Econometrics	-0.01	0.07 (0.25)	531	0.08 (0.28)	1,940
Performance					
Publications	-18.29***	16.92 (22.34)	531	35.21 (46.71)	1,940
Citations	-4,786.38***	2,930.62 (5,769.67)	410	7,717.00 (19,988.87)	1,604
h-Index	-8.20***	15.33 (13.20)	410	23.53 (22.22)	1,604
i10-Index	-19.97***	21.16 (26.68)	410	41.13 (66.35)	1,604
m-Index	-0.17***	1.14 (0.76)	397	1.31 (0.81)	1,576
Top 5	-2.13***	1.53 (2.79)	531	3.66 (6.50)	1,940
Top 5 Share	-0.02**	0.11 (0.18)	483	0.13 (0.18)	1,868
SN	-0.07**	0.09 (0.51)	531	0.16 (0.83)	1,940

Notes: Mean differences by gender (allowing for unequal variances). Standard deviations in parenthesis. Significance levels: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$.

were classed as non-dominant. The results are similar when using the mean plus one standard deviation as cut-off to classify individuals as dominant or non-dominant.

Our findings suggests that non-dominant women have 3.37 more co-authors, 4.12 more publications, 1,713.12 more citations, an h-Index 5.51 higher, an i10-Index 9.88 higher, an m-Index 0.10 higher, 0.24 more top 5 publications, and 0.02 more Science and Nature publications compared to their more dominant female colleagues. Further, table 3 also suggests dominant women are less often editors, typically younger, and tend to study less theoretical microeconomics and/or econometrics, all suggestive evidence for our theory.

Table 3: Descriptive Statistics - Dominant Females.

	Difference	Non-Dominant	Obs.	Dominant	Obs.
Socio-Demographic					
White	-0.06	0.68 (0.47)	287	0.74 (0.44)	244
Time since PhD	1.47	14.61 (10.95)	281	13.14 (10.72)	240
Co-Authors	3.37***	14.27 (16.54)	287	10.90 (12.99)	244
Switches	0.16*	0.96 (1.09)	284	0.80 (0.89)	242
Grants	0.42	5.38 (7.28)	286	4.96 (6.53)	244
Editor	0.05	0.49 (0.50)	286	0.44 (0.50)	244
Theory	0.08**	0.36 (0.48)	287	0.27 (0.45)	244
PhD Top 10	0.02	0.47 (0.50)	287	0.45 (0.50)	244
PhD Top 20	0.05	0.65 (0.48)	287	0.60 (0.49)	244
First Job Top 10	0.03	0.13 (0.34)	287	0.10 (0.30)	244
Field					
Applied Micro	-0.11***	0.55 (0.50)	287	0.66 (0.47)	244
Theoretical Micro	0.004	0.14 (0.35)	287	0.14 (0.34)	244
Macro	0.08**	0.23 (0.42)	287	0.15 (0.36)	244
Econometrics	0.03	0.08 (0.28)	287	0.05 (0.23)	244
Performance					
Publications	4.12**	18.82 (22.16)	287	14.70 (22.40)	244
Citations	1,713.12***	3,711.97 (6,719.00)	223	1,998.86 (4,208.65)	187
h-Index	5.51***	17.84 (14.68)	223	12.34 (10.46)	187
i10-Index	9.88***	25.66 (29.64)	223	15.79 (21.51)	187
m-Index	0.10	1.19 (0.89)	216	1.08 (0.57)	181
Top 5	0.24	1.64 (2.93)	287	1.39 (2.62)	244
Top 5 Share	-0.01	0.11 (0.17)	258	0.12 (0.19)	225
SN	0.02	0.10 (0.58)	287	0.08 (0.43)	244

Notes: Mean differences by dominance for females (allowing for unequal variances). Standard deviations in parenthesis. Dominance defined as having larger than average fWHR. Significance levels: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$.

Similarly, table 4 splits the male population into two groups, dominant and non-dominant pertaining to their fWHR rankings. Those greater than the sample average of 1.70 (899 observations) were classed as dominant. Those less than or equal to 1.70 (1,041 observations) were classed as non-dominant.

Non-dominant men have 4.06 less co-authors, 6.27 less publications, 3,694.27 less citations, an h-Index 5.17 lower, an i10-Index 11.42 lower, an m-Index 0.12 lower, 0.66 less top 5 publications, and 0.06 less Science and Nature publications than their more dominant male colleagues. Further, table 4 suggests dominant males switch jobs more often. Although the values given to publications, citations, top5, m-index, and co-authors are insignificant, this may be due to the small sample size. Nonetheless, this is again suggestive evidence for our theory. Importantly, for men or women there appears to be no selection into Top 10 or 20 PhD programmes based upon fWHR. We do not find a statistically significant differences across the two dominance groups. The same holds for a first job in the Top 10 Departments.

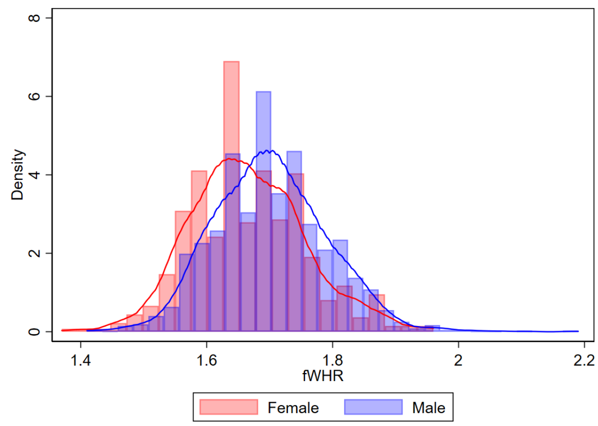
Table 4: Descriptive Statistics - Dominant Males.

	Difference	Non-Dominant	Obs.	Dominant	Obs.
Socio-Demographic					
White	-0.07***	0.73 (0.44)	1,041	0.80 (0.40)	899
Time since PhD	-1.49**	19.16 (13.82)	1,029	20.65 (13.88)	892
Co-Authors	-4.06***	20.22 (27.58)	1,041	24.28 (32.57)	899
Switches	-0.08	1.01 (1.14)	1,030	1.10 (1.19)	886
Grants	-0.58	5.16 (7.99)	1,038	5.74 (8.22)	893
Editor	-0.04*	0.55 (0.50)	1,040	0.59 (0.49)	898
Theory	0.05**	0.53 (0.50)	1,041	0.47 (0.50)	899
PhD Top 10	-0.02	0.44 (0.50)	1,041	0.46 (0.50)	899
PhD Top 20	0.005	0.60 (0.49)	1,041	0.59 (0.49)	899
First Job Top 10	0.01	0.14 (0.35)	1,041	0.13 (0.34)	899
Field					
Applied Micro	-0.05**	0.37 (0.48)	1,041	0.42 (0.49)	899
Theoretical Micro	0.03	0.24 (0.43)	1,041	0.21 (0.41)	899
Macro	-0.01	0.29 (0.45)	1,041	0.30 (0.46)	899
Econometrics	0.03**	0.10 (0.30)	1,041	0.07 (0.25)	899
Performance					
Publications	-6.27***	32.31 (43.37)	1,041	38.57 (50.10)	899
Citations	-3,694.27***	5,985.03 (16,229.25)	852	9,679.30 (23,390.71)	752
h-Index	-5.17***	21.11 (19.79)	852	26.28 (24.41)	752
i10-Index	-11.42***	35.77 (63.50)	852	47.19 (68.97)	752
m-Index	-0.12***	1.25 (0.75)	835	1.37 (0.86)	741
Top 5	-0.66**	3.35 (6.04)	1,041	4.01 (6.99)	899
Top 5 Share	0.01	0.14 (0.19)	1,004	0.13 (0.17)	864
SN	-0.06	0.13 (0.73)	1,041	0.19 (0.93)	899

Notes: Mean differences by dominance for males (allowing for unequal variances). Standard deviations in parenthesis. Dominance defined as having larger than average fWHR. Significance levels: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$.

The histogram shown in Figure 2 summarises our findings. fWHR, the dependent variable, is on the x-axis ranging from 1.35 to 2.2; and its respective density is observed on the y-axis. All female researchers observed are recorded in red, and male researchers in blue. For both sub-samples, we see a normal distribution with both respective means lower than the national averages of 1.73 for females and 1.83 for males.[21]

Figure 2: Histogram of fWHR.



Notes: Histogram and density estimates of female (red) and male (blue) fWHR.

Figure 3 next shows a scatter plot of the data.[22] We also find the highest h-Index in the female sample occurs at under half the value of the highest male h-Index in the sample. These findings may corroborate Symonds et al. (2006)'s findings.

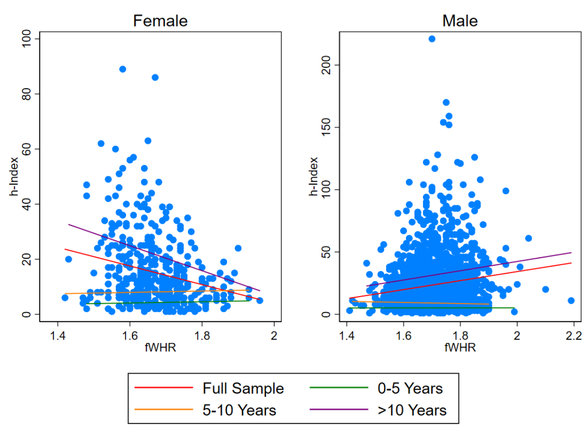
Adjusting for time since PhD and running linear regressions, we see that in both genders, in the 0-5 years and 5-10 years categories, there appears to be minimal effect from fWHR on h-Index, which is expected as people need time to publish and generate a high h-Index. Moving to the >10 year window, however, a relationship begins to form, with higher fWHRs negatively linked to h-Index in the female sample, and positively linked in the male sample. Overall, the graph shows that over time the two lines rotate around two different fix points.

[21]We speculate this difference may be because those who pursue economics may have smaller dominance levels already.

[22]Figure 4 in the appendix presents a heat map of the data.

The line for females rotates around a high fWHR value, whereas for males it rotates around a low fWHR value.[23]

Figure 3: Scatter Plot by time since PhD.



Notes: Scatter plot for the relationship between fWHR and the h-Index for female (red) and male (blue) academics by time since PhD.

4.3 Econometric Strategy

4.3.1 Main Strategy

We make use of a cross-sectional research design at the individual level, i, given by

$$Performance_i = \alpha + \beta fWHR_i + \gamma \mathbf{x}_i + \epsilon_i, \tag{8}$$

where $\alpha \in \mathbb{R}$ is a constant and ϵ_i is the error term.

The dependent variable is academic performance measured by the h-Index. In robustness checks, we use other measures of academic success. The key parameter we are interested in is β, which captures the effect of personality on performance. Personality in our analysis is

[23]Results from Figure 3 also hold when performing a cubic spline regression (See figures 5, 6, 7, and 8 in the appendix).

proxied by the value of the facial width-to-height ratio ($fWHR$).

Further, $\mathbf{x}_i$ is a K-dimensional matrix of control variables. The model controls for various confounding factors: age, ethnicity, the number of co-authors, editorship, and field. In the robustness section, we extend this set of control variables. We prefer to have a parsimonious model, which includes the main drivers, rather than to over-fit and to introduce bad control problems (Hsiang et al. 2013; and Burke et al. 2015). Finally, we use robust standard errors in all regressions.

Our key explanatory variable, fWHR, is plausibly exogenous and we do not expect reverse causality to be a problem. It should be noted that this reduced-form approach does not allow us to identify *how* the fWHR affects academic performance, but causal inference is obtained by the random variations in fWHR across scholars.

We are, however, concerned about the potential selection bias created by the self-selection into having a Google Scholar account and the omitted variable bias in our research design. To address the sample selection problem, we use a Heckman selection model (see section 6). To address the omitted variable bias, we use different approaches. While we attempt to control for as many potential confounders as possible, we also use different research designs to address this issue (see the section 6 for details). First, we run our regression model (eq. 8) on a matched sample. This addresses the issue because observed and unobserved characteristics are highly correlated (Stuart, 2010; and Ferraro and Miranda, 2014). Therefore, matching on observable variables implies at least some matching on the unobservable variables. The matched sample is generated using one-to-one nearest neighbour matching with replacement as well as entropy-balancing in an alternative specification.
Second, we use a first difference approach inspired by the work by Druckenmiller and Hsiang, (2019), which addresses the problem of unobservable heterogeneity in this type of cross-

sectional research design over rankings.

4.3.2 First Differences

The cross-sectional regression design discussed in the previous section (see eq. 8), suffers from potential omitted variable bias. The reason is that important covariates, which affect our outcome variable and are potentially correlated with the fWHR variable, are not observable (Clarke, 2005; and Angrist and Pischke, 2010). Examples for such unobserved variables are the number of children a scholar has, the usage of parental leave, or the general level of effort provided.

In order to address this problem, we adopt the novel cross-sectional research design, called *spatial first differences*, (SFD), developed by Druckenmiller and Hsiang, (2019). Intuitively, the research design uses the fact that observations can be organized in space (or in our applications, we can generate a ranking according to observable outcomes). The estimator then only compares observations to their immediately adjacent neighbours and, at the same time, compares all observations to a neighbour. The underlying assumption is that immediately adjacent observations are comparable to one another but are not comparable to observations far away; as it is assumed in a standard cross-sectional approach. Hence, the first difference estimator inspired by this paper removes all omitted variables common to neighbouring units.[24] In our application, we interpret locations as researchers.

Formally, consider i to be an index of location, which describes the (linear) rank order of spatial observations. Further, y is the outcome variable and $\mathbf{x}$ is a K-dimensional vector of observables. Then, the usual identifying assumption in a cross-sectional regression,

$$\forall i \neq j : \mathbb{E}\left[y_i \,|\mathbf{x}_j\right] = \mathbb{E}\left[y_j \,|\mathbf{x}_j\right], \tag{9}$$

[24] This design is similar to quasi-experimental research designs.

is replaced with the weaker assumption

$$\forall \{i, i-1\} : \mathbb{E}\left[y_i \left| \mathbf{x}_{i-1}\right.\right] = \mathbb{E}\left[y_{i-1} \left| \mathbf{x}_{i-1}\right.\right], \tag{10}$$

where only small differences in $\mathbf{x}$ and y are compared. This assumption is referred to as the *local conditional independence assumption.* Notice that this assumption is similar to the identifying assumption in time-series models, event-study designs, the assumption in panel models where sequential observations within a panel unit are comparable (e.g. fixed effects estimators), and regression discontinuity research designs (Wooldridge, 2010).

The spatial first difference estimator is then given by

$$\underbrace{y_i - y_{i-1}}_{\Delta y_i} = \underbrace{(\mathbf{x}_i - \mathbf{x}_{i-1})}_{\Delta \mathbf{x}_i} \beta + \underbrace{(\mathbf{c}_i - \mathbf{c}_{i-1})}_{\Delta \mathbf{c}_i} \theta + \underbrace{(\epsilon_i - \epsilon_{i-1})}_{\Delta \epsilon_i}, \tag{11}$$

where $\mathbf{c}$ is an M-dimensional vector of unobservables and Δ is the equivalent to the time-series difference operator. As the variables in $\mathbf{c}$ are unobservable, they are omitted from the spatial first difference regression

$$\Delta y_i = \Delta \mathbf{x}_i \beta_{FD} + \Delta \epsilon_i. \tag{12}$$

This regression can be estimated using various approaches such as ordinary least squares (OLS, for short) or maximum likelihood. Using OLS, the estimator is given by

$$\hat{\beta}_{FD} = \left[\Delta \mathbf{x'} \Delta \mathbf{x}\right]^{-1} \left[\Delta \mathbf{x'} \Delta y\right]. \tag{13}$$

If the assumption in equation (10) holds, then

$$\mathbb{E}\left[\Delta \mathbf{x'} \Delta \mathbf{c}\right] = \mathbf{0}_{K \times M}, \tag{14}$$

which is the key identifying assumption in the spatial first difference design implying that the covariance between changes in $\mathbf{x}$ and $\mathbf{c}$ between immediately adjacent observations is not systematically correlated within the entire population. As this equation can also be interpreted as the covariance between the local derivatives of $\mathbf{x}$ and $\mathbf{c}$ with respect to space, the condition is only violated if the second derivatives systematically generate changes in these derivatives in the same locations.

In sum, if the assumption (14) holds, then the estimator $\hat{\beta}_{FD}$ is unbiased. The idea is that if the unobservables variables $\mathbf{c}$ are common between neighbours, the influence on y will be differenced out and the term $\Delta\mathbf{c}$ is equal to zero (or at least very close to it). Further, even if there is a component of $\mathbf{c}$ which is not common between neighbours, the estimator will still be unbiased, if the non-zero component of $\Delta\mathbf{c}$ is not correlated with changes in $\mathbf{x}$ between neighbours $\Delta\mathbf{x}$.[25]

Compared to our main approach and the typical usage of the SFD, we order individuals in our data set by (i) the number of publications, (ii) the number of Top 5 publications, and (iii) the share of Top 5 publications. Therefore, we replace i as the location with the position in one of these rankings. The assumption is that adjacent individuals in these rankings are similar in observable and unobservable characteristics. For each (unique, linear) ranking, we then run a first difference regression (see section 6), which reduces the omitted variable bias.

4.4 Hypotheses

We hypothesise, based on the formula shown in equation (7), that in the male sample $\beta > 0$, but in the female sample, $\beta < 0$. That is to say, we contend that men with higher fWHRs will be more successful than men with lower fWHRs. Conversely, women with higher fWHRs will be less successful than those with lower fWHRs. We postulate this will be the case for three primary reasons: (i) behaviour, (ii) discrimination, and (iii) networking.

[25]Further details such as the asymptotic distribution or the estimated variance can be taken from Druckenmiller and Hsiang, (2019).

In terms of behaviour, we speculate that higher fWHR'ed individuals (more dominant) express greater behavioural tendencies regarding competitiveness, drive for success, and many other personality traits as outlined in section 2 (See He et al. 2019). Irrespective of gender, we suspect these personality traits will directly lead more dominant individuals to success as their perceived stubbornness and aggression likely mean they will push forward ideas and fight to be published (See Kausel et al. 2018). Furthermore, we contend these personality features will persevere across both females and males.

With respect to discrimination, we postulate dominant women and to a lesser extent non-dominant men may face prejudice for non-conformity in regards to (perceived) gender norms. This is because we contend that higher fWHRs in men are indicative of higher levels of aggression, dominance, competitiveness, and risk taking behaviours. As these are stereotypically masculine traits, they are well received leading to greater success for dominant men. In contrast, behavioural traits such as dominance and aggression are not stereotypically associated with femininity meaning the behaviour of such women may in fact buck gender norms. We theorise the impacts of this are twofold: firstly, higher fWHR women may face a greater level of workplace discrimination which may limit their success (See Sutherland et al. 2014; and Bohren et al. 2018), and secondly, said women may in fact be 'punished' for their non-conformity (See Phillips and MacKinnon, 2009; Dall'Ara and Maass, 1999; and Heilman et al. 2004).

Additionally, we hypothesise this links directly to networking which we contend will have a large effect on performance. We argue that networking depends on sociability, however, this may favour certain dominance dynamics. In particular, we deduce that there may be inherent biases working against dominant women. Due to the potential discrimination, dominant women may be collaborated with less because (i) higher fWHR and therefore more dominant women feel less need to collaborate (See McDowell et al. 2007; Ductor et al. 2018; and Lindenlaub and Prummer, 2020); and (ii) other academics may be less inclined to network with these individuals (See Phillips and MacKinnon, 2009). We believe these may outweigh the behavioural advantages of higher fWHRs that we would expect to observe for men.

5 Estimation Results

Our main results are shown in table 5. The first two columns pertain to the relationship between fWHR and h-Index for females and males with no controls, i.e. in the misspecified model. The female sample consists of 402 observations, and the male of 1,590. Overall, we observe that for females, if their fWHR increases by 1 unit, their h-index will fall by 33.39. That is to say, a one standard deviation shift for a female, equal to 0.09, would lower a female h-Index by 3. In the male sub-sample, we observe that increasing fWHR by 1 unit will result in an h-Index increase of 36.38. Thus, a one standard deviation shift (0.09) would result in an h-Index 3.28 units higher. Both of these results are statistically significant at the 1 percent significance level. Intuitively, this makes sense. In the misspecified model, we would expect the relationship between fWHR and success to be strongly positively related for males as we would expect men with higher fWHRs to demonstrate more personality traits that are more likely to lead them to success, in line with Lu and Teo, (2018), and Hahn et al. (2017)'s findings. Similarly, while we expect these qualities to also occur in females, we would expect discrimination and networking difficulties to negate the aforementioned behavioural benefits, resulting in the strong negative relationship we observe.

The third and fourth columns show the resulting specified relationship between fWHR and h-Index across gender when ethnicity, time since PhD, number of co-authors, editorial status, and field are controlled for. The controlled result suggests a one standard deviation

Table 5: Main Results.

Variable	Female	Male	Female	Male
fWHR	-33.39***	36.38***	-11.62***	9.45***
	(6.56)	(5.62)	(3.78)	(3.12)
Controls				
White			0.78	1.58***
			(0.77)	(0.54)
Time since PhD			0.46***	0.71***
			(0.07)	(0.04)
Co-Authors			0.40***	0.34***
			(0.07)	(0.03)
Editor			3.15***	4.01***
			(0.72)	(0.67)
Field				
Micro			0.18	-2.24**
			(1.34)	(0.99)
Macro			1.81	3.20***
			(1.45)	(1.23)
Obs.	402	1,590	402	1,590
R^2	0.05	0.02	0.73	0.69

Notes: Dependent variable is h-Index. Robust standard errors in parenthesis. Constant not shown. Significance levels: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$.

fWHR shift increase would cause a 1 unit drop in the corresponding h-Index for a female. For males, a one standard deviation fWHR shift increase would cause a 0.84 unit drop in the corresponding h-Index. Further, both of these results are statistically significant at the 1 percent significance level. Intuitively, these findings are in line with Kausel et al. (2018) in that they observe some effect on success dependent on fWHR. Further, these findings provide evidence in favour of our hypothesis (See Sutherland et al. 2014; and Phillips and MacKinnon, 2009).

The effect of fWHR is smaller compared to the misspecified model, which is due to the fact that, for example, age is a key driver of the h-Index. In terms of the control variables, the impact of being white results in a 0.70 (0.75) increase in h-Index for females and an increase of 1.68 (0.54) for males with the latter being statistically significant at the 1 percent significance level. Further, by design of the h-Index, with each additional year of time since completion of a PhD, females increase their h-index by 0.38 and males by 0.68. Additionally, collaborating with another author increases the h-index by 0.37 for females and 0.33 for males. Female editors also have an h-Index 2.47 higher than non-editors, and male editors 4.04 higher also. All results found for the variables time since PhD, co-authors, and editorial status are statistically significant at the 1 percent significance level for both males and females.

These findings are in line with existing literature. Race is observed to have an effect on performance as shown in McPherson et al. (2001). Furthermore and intuitively, each extra year working in the industry increases performance. Finally, the findings in terms of networking collaborate those found by Combes et al. (2008); Ferber and Teiman (1980); and McDowell and Smith, (1992). Networking is correlated with success, but factors such as discrimination can interfere with potential collaboration.

When split into the field sub-divisions of microeconomics and macroeconomics, females researching microeconomics were found to have an h-Index 0.31 higher than non-microeconomics researchers. Females researching macroeconomics had an h-Index 1.62 higher than non-macroeconomics researchers also. Males working in microeconomics had an h-Index 1.89 lower than non-microeconomics workers, statistically significant at the 10 percent significance level. Males working in macroeconomics, however, had an h-Index 3.35 higher than non-macroeconomics workers, statistically significant at the 1 percent significance level.

6 Robustness

Potential limitations of the model pertain primarily to sample selection and omitted variable bias. Thus a number of specifications to the model were implemented to address these limitations. We present the results below.

6.1 Sample Selection Bias

Due to the nature of the data collected, not all subjects in the sample possess Google Scholar accounts with which we could access their respective performance measures, creating a sample selection bias. We speculate this may be the case for one of three reasons. Firstly, the subject has not yet published enough papers to have a need for a Google Scholar account. Secondly, the subject has published so many publications that they feel there is no need to create an account. Or thirdly, that the subject is unaware or un-incentivised to create an account. As this means the outcomes are missing, this creates a sample selection bias towards publications for the first two reasons, and towards age for the latter.

To minimise the bias which may result from this, we employed two specifications to the model; a Heckman selection model, and propensity score matching.

A Heckman selection model fits a model where the desired outcome is sometimes missing using one equation to determine the dependent variable, and a second equation to determine the likelihood that the dependent variable value will be missing. When the model is specified as shown in table 6 with 523 female and 1,926 male observations, we find the instrument of number of publications is significant.[26] The estimate of the effect of the fWHR on the h-Index for females is -10.54 and for males is 8.32, both statistically significant at the 1 percent significance levels. This suggests a 0.1 increase in fWHR for females results in a 1.054 decrease in the observed h-Index, and a 0.1 increase in fWHR for males results in a 0.832 increase in the observed h-Index. These results are fairly consistent with the main findings found in the model.

In terms of the control variables, we find estimates for time since PhD, number of co-authors, and editorial status to be statistically significant at the 1 percent level for both females and males. The male coefficient for the control variable white is also statistically

[26] The results are robust to using time since PhD (<5 years as an instrument).

significant at the 5 percent level, and the coefficients for both the micro and macro fields in the male sample too are significant at the 10 percent level.

The second robustness check implemented to test the sample selection bias was propensity score matching which minimises the bias by matching any observables and unobservables, therefore creating a balanced sample.[27] Specifically, we matched the subjects on the following variables; rank, time since PhD, co-authors, field, white, and switches. Total observations were 183 for the female sample and 739 for the male sample. Figure 9 in the appendix presents the distribution before and after matching rather than performing a t-test, because the sample size can affect the t-test validation of the balanced sample.

Table 6 shows the results, with estimators of -8.85 and 6.32 for the fWHR of the female and male samples respectively. This implies a 0.1 increase in fWHR for a female will result in a decrease in the observed h-Index of 0.885, and an increase of 0.1 in the fWHR in a male will increase the observed h-Index by 0.632. These results are both also statistically significant at the 5 percent significance level.

In terms of the control variables, all coefficients for time since PhD, number of co-authors, and editorial status for both the female and male samples are statistically significant at the 1 percent level. The estimate for the male white variable is also significant at the 10 percent level, and the male macro field estimator is significant at the 5 percent significance level.

6.2 Omitted Variable Bias

Another potential limitation of the model is omitted variable bias. While we have controlled for a number of variables in the specified model compared to the misspecified model as shown

[27] The results are robust to using entropy balancing (first two moments).

Table 6: Robustness Checks.

Variable	Heckman Female	Heckman Male	First Difference Female	First Difference Male	Matching Female	Matching Male
fWHR	-10.54***	8.32***	-10.82***	8.93***	-8.85**	6.32**
	(3.46)	(2.99)	(3.27)	(3.01)	(4.39)	(3.13)
Controls						
White	0.65	1.46**	0.88	1.68**	0.83	1.31*
	(0.72)	(0.68)	(0.74)	(0.67)	(1.46)	(0.76)
Time since PhD	0.45***	0.59***	0.39***	0.66***	0.59***	0.74***
	(0.04)	(0.03)	(0.06)	(0.05)	(0.11)	(0.06)
Co-Authors	0.31***	0.27***	0.36***	0.33***	0.31***	0.33***
	(0.02)	(0.01)	(0.07)	(0.03)	(0.10)	(0.04)
Editor	4.18***	5.58***	1.53 **	1.66**	3.22***	2.90***
	(0.80)	(0.70)	(0.72)	(0.74)	(1.10)	(0.88)
Field						
Micro	0.32	-1.96*			-0.49	-1.38
	(1.30)	(1.02)			(3.44)	(1.46)
Macro	1.12	1.70			1.67	3.25**
	(1.40)	(1.07)			(3.42)	(1.64)
Selection						
Number Publications	0.70***	0.46***				
	(0.11)	(0.06)				
LR-Test χ^2	44.91	220.74				
Obs.	523	1,926	393	1,576	183	739
R^2			0.48	0.48	0.72	0.71

Notes: Dependent variable is h-Index. Heckman selection model uses a dummy for the number of publications larger than 5 as selection variable. First difference estimation uses ranking by top 5 share (citations as tie-breaker). First difference estimation also includes difference for field variable (not shown). Matching sample obtained using propensity score matching. Robust standard errors in parenthesis. Constant not shown. Significance levels: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$.

in table 5, unobserved variables may still confound the results. To minimise this bias, we have conducted a number of robustness checks.

First, we conducted a first differences model. As explained in subsection 4.3.2, this compares two subjects of similar status in the model at a time to minimise the effect of any sample selection bias. Using a sample of 393 female subjects, and 1,576 male subjects, we have compared these based on ranking, publications, and top 5 ranking.

The specified model shown in table 6 resulted in coefficients for female and male fWHRs of -10.82 and 8.93 respectively, both again statistically significant at the 1 percent significance level. This is consistent with our initial findings suggesting a 0.1 increase in fWHR for females and males will results in a 1.082 decrease and 0.893 increase in the observed h-Index respectively.

For both the female and male sub-samples, we find the coefficients estimated for control variables time since PhD and co-authors to be statistically significant at the 1 percent significance level. The estimates for both genders' editorial status are also significant at the 5 percent significance level as is the coefficient for the control variable white in the male sample. Conclusively, the first difference results are robust to publications, ranking, and top 5 ranking.

The next specification to the model was an Oster (2019) test which evaluates the importance of the unobserved variables with respect to the observed variables. We found a δ of 0.0004 for the female sample, and 0.0001 for the male sample. This means that the unobserved variables in the female sample are 0.004 percent as important as the observed variables. Similarly, in the male sample, the unobserved variables are 0.001 percent as important as the observed variables. This suggests any omitted variable bias is minimal.

We also included variables pertaining to beauty to see if these had any impact on performance.[28] Specifically, we included a binary variable for weight status (overweight), and a measure of beauty taken from Hamermesh et al. (2002). The resulting fWHR estimators were consistent with initial findings at -10.93 for females, and 9.58 for males. Both values were again statistically significant at the 1 percent significance level. In terms of the control variables, time since PhD, co-authors, and editorial status were statistically significant at

[28] Refer to Table 7.

the 1 percent level for both females and males, as was the male white variable. Both the micro and macro male field specific variables were found to be statistically significant too (5 percent level). Of the beauty specific variables, however, none proved to be statistically significant. Given this result as well as no major change in the fWHR estimators, we interpret this to suggest there is no beauty premium or penalty.

The next robustness check for potential omitted variable bias we conducted was to add a number of additional control variables. These included observations on the number of grants subjects had been awarded, what theoretical field they researched in, any variation in Bachelor or PhD majors other than Economics, ethnicity, university rank, nationality, surname (See Einav and Yariv, 2006), the interaction between theory and field, and the interaction between ethnicity and fWHR. The rank of the researcher here is used as a proxy for service workload. Despite the additional control variables, the effect of the fWHR remained consistent with our main findings.

Finally, we conducted a Lasso regression. This essentially hands all the data to Stata who run an algorithm to determine which variables it believes effect the dependent variable. When we perform the Lasso regression, the fWHR is selected for both the male and the female samples and the results are similar to our main findings presented in table 5.

6.3 Other Robustness Checks

Other robustness checks implemented included a Poisson regression specification to the model. The reason for this was that our dependent variable (h-Index) is best described as count data. Our resulting estimations[29] for both the female and male samples were statistically significant at the 1 percent level with the female fWHR estimator equal to -0.98

[29]Refer to Table 7.

and the male fWHR estimator equal to 0.46. This suggests less dominant women are on average more successful than more dominant women, and the converse relationship for the male sample; consistent with our main findings. With respect to the control variables, both the female and male time since PhD, number of co-authors, and editorial status were statistically significant at the 1 percent significance level, as was the male white variable.

Table 7: Additional Robustness Checks.

	Poisson		Research Status		Incl. Beauty	
Variable	Female	Male	Female	Male	Female	Male
fWHR	-0.98***	0.46***	-9.52***	7.68***	-10.93***	9.58***
	(0.23)	(0.16)	(3.13)	(2.93)	(3.84)	(3.14)
Controls						
White	0.10	0.10***	0.69	2.54***	0.59	1.43***
	(0.07)	(0.03)	(0.71)	(0.72)	(0.81)	(0.55)
Time since PhD	0.03***	0.03***	0.13**	0.15	0.49***	0.72***
	(0.003)	(0.001)	(0.06)	(0.04)	(0.08)	(0.04)
Co-Authors	0.01***	0.004***	0.13**	-0.02	0.40***	0.34***
	(0.003)	(0.001)	(0.06)	(0.04)	(0.07)	(0.03)
Editor	0.51***	0.56***	2.98***	4.91***	3.13***	4.00***
	(0.07)	(0.03)	(0.59)	(0.73)	(0.72)	(0.67)
Field						
Micro	0.02	-0.06	1.79*	-0.62	-0.14	-2.28**
	(0.11)	(0.05)	(1.07)	(0.89)	(1.36)	(1.00)
Macro	0.10	0.07	3.05**	2.92**	1.39	3.13**
	(0.11)	(0.06)	(1.18)	(1.20)	(1.49)	(1.23)
Beauty						
Above Average					0.28	1.08
					(0.68)	(0.87)
Below Average					-1.27	-0.81
					(0.90)	(0.76)
Overweight					-1.33	1.11
					(1.55)	(1.94)
Obs.	402	1,590	402	1,590	402	1,590
R^2	0.51	0.57	0.34	0.12	0.73	0.69

Notes: Dependent variable is h-Index for the Poisson regression and the regression including the beauty control. Poisson R^2 is Pseudo R^2. Robust standard errors in parenthesis. Constant not shown. Significance levels: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$.

Furthermore, Symonds et al. (2006) and Geraci et al. (2015) highlighted a number of concerns in using the h-Index as a measure of performance, namely inherent gender bias and a higher weighting towards the top 5 Economic journals. To combat potential bias arising from using the h-Index as our dependent variable, we re-ran the model using other measures of success; research status, number of publications, citations, i10-Index, and the number of top 5 publications. Results are shown in tables 7 and 8.

First, we consider the research status variable proposed by Symonds et al. (2006) outlined in section 2.1. as the dependent variable with the reason being that the h-Index may be biased. Calculated as the residual obtained from regressing the h-Index on the number of publications, this variable was found by Symonds et al. (2006) to better control for gender differences in performance. Thus, the coefficients for the fWHR for the female and male sample were -9.52 and 7.68 respectively with both results statistically significant at the 1 percent significance level. This result is robust to our main findings, suggesting fWHR affects performance the same regardless of whether success is measured using the h-Index or research status. Furthermore, it suggests our initial results are not skewed by inherent gender bias in the dependent variable. Of the control variables, again we see 1 percent statistical significance for the editorial status coefficients for both females and males. The male white variable is statistically significant at the 1 percent level, the female time since PhD, female number of co-authors, and both female and male macro at the 5 percent level, and the female micro at the 10 percent significance level.

Maintaining the control variables from the original model, we next adjusted the dependent variable to number of publications which produced estimators for the fWHR for the female and male sample of -7.14 and 3.77 respectively. This suggests that increasing a female fWHR by 0.1 would decrease their respective number of publications by 0.714, a result statistically significant at the 10 percent significance level. Increasing a male's fWHR would increase their number of publications by 0.377. Overall this indicates that dominant women tend to be less successful on average compared to non-dominant women, and dominant men tend to be more successful on average than non-dominant men when measuring success through publications.

With citations as the dependent variable, we observe that increasing the fWHR for females by 0.1 results in a decrease in the number of citations by 419.534. This result is statistically significant at the 5 percent significance level, and suggests more dominant women will be cited less often on average than less-dominant women. For the male sub-sample, increasing fWHR by 0.1 results in an increase in citations of 701.163. Again statistically significant at the 5 percent level, these results indicate dominant males are more likely to be cited on average than non-dominant males.

Using the i10-Index as the dependent variable, our specified model results in estimators for the fWHR of -18.99 for the female sample, and 20.65 for the male sample. The female result is statistically significant at the 1 percent level, suggesting a 0.1 increase in fWHR for a female will result in a drop in their respective i10-Index of 1.899. The male result, statistically significant at the 5 percent level, suggests a 0.1 increase in fWHR for a man results in an increase in the i10-Index of 2.65. This means on average, dominant women are less successful than non-dominant women. Conversely and consistent with the main results, dominant men are on average more successful than non-dominant men.

Finally, using the number of top 5 Economic journal publications as the dependent variable resulted in estimators of -0.97 and 0.78 for the fWHR for the female and male subsamples respectively. This suggests increasing the fWHR by 0.1 results in 0.097 less top 5 publications for women, and 0.078 more top 5 publications for men, however these are statistically insignificant.

Ultimately, these results suggest fWHR has a significant effect on academic performance

Table 8: Additional Robustness Checks - Different Dependent Variables.

	Publications		Citations		i10-Index		Top 5	
Variable	Female	Male	Female	Male	Female	Male	Female	Male
fWHR	-7.14*	3.77	-4,195.34**	7,011.63**	-18.99***	20.65**	-0.97	0.78
	(4.28)	(5.72)	(1,822.43)	(3,235.55)	(7.11)	(9.44)	(1.03)	(1.29)
Controls								
White	0.57	-2.77**	203.69	1,856.80***	0.61	0.64	0.17	0.54**
	(0.62)	(1.28)	(430.14)	(475.16)	(1.49)	(1.45)	(0.19)	(0.24)
Time since PhD	0.83***	1.35***	92.71*	366.42***	0.80***	1.61***	0.02	0.12***
	(0.09)	(0.08)	(52.23)	(47.04)	(0.16)	(0.14)	(0.02)	(0.01)
Co-Authors	0.68***	0.93***	178.47***	241.14***	0.90***	1.15***	0.05**	0.05***
	(0.09)	(0.07)	(60.42)	(46.26)	(0.17)	(0.14)	(0.02)	(0.01)
Editor	-0.23	-0.74***	581.39***	-698.71	2.08	-2.11	1.44***	1.77***
	(1.08)	(1.20)	(332.79)	(863.18)	(1.55)	(2.62)	(0.23)	(0.29)
Field								
Micro	-2.24*	-3.87**	-11.11	-4,238.51***	-1.78	-7.34**	-0.58	0.31
	(1.23)	(1.55)	(878.40)	(1,220.03)	(2.98)	(2.94)	(0.37)	(0.36)
Macro	-1.29	1.93	1,117.57	2,392.51	1.77	9.96**	0.23	1.64***
	(1.39)	(1.93)	(949.93)	(1,660.16)	(3.39)	(4.25)	(0.45)	(0.44)
Obs.	520	1,921	402	1,590	402	1,590	520	1,921
R^2	0.78	0.73	0.43	0.30	0.69	0.55	0.27	0.26

Notes: Robust standard errors in parenthesis. Constant not shown. Significance levels: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$.

for both males and females irrespective of how academic performance is measured. Our initial findings persist with fWHR positively driving academic success for males, and negatively driving it for females. Our model is robust to other dependent variable specifications.

We also speculate outliers in the sample may too confound results. Accordingly, we restricted our sample in various ways to test if our results were consistent. These sample restrictions included excluding assistant professors on the grounds that their careers are just beginning and they are therefore less likely to have published many articles, acquired many citations, or earned a higher h-Index. Similarly, we also excluded any academics in the sample with zero publications. In another restriction, we excluded those in the early stages of their careers (<5 years) on the basis that they may have lower recorded h-Indexes due to less experience. Our final sample restriction pertained to only including CVs less than 3 years old. This would mean they were more likely to be accurate. Ultimately, despite restrictions to the sample, our results are robust.[30]

A different way to assess whether fWHR is relevant for performance is to perform an out-

[30] Results available upon request.

of-sample prediction. We conducted out-of-sample predictions for each sub-sample using 70 percent and 80 percent of the original sample data respectively. These samples are drawn randomly and individually by gender. Results are shown in table 9.

For the 70 percent sample predictions, using only the initial control variables found in the main estimation results and employing the h-Index as the dependent variable, we find a fit of 4.051 for the female sample, and 8.107 for the male sample. Adding in the independent variable, fWHR, lowers both of these RMSE values to 3.994 and 8.085 respectively, suggesting fWHR creates a better fit for both genders.

In the 80 percent sample predictions, the RMSE values are lower. With only controls, the female RMSE value is 3.170 and the male RMSE value 6.000. With the addition of the fWHR variable, these drop to 3.125 and 5.968 respectively, again suggesting the fWHR better fits the data.

Table 9: Out-of-Sample Predictions.

	70 % Sample		80 % Sample	
	Female	Male	Female	Male
Only Controls	4.051	8.107	3.170	6.000
With fWHR	3.994	8.085	3.125	5.968

Notes: Table shows RMSE values. The samples are randomly drawn from the population. Dependent variable is h-Index. Regressions use the same control variables as before.

7 Conclusion

In conclusion, we found that men with higher fWHRs, ergo more dominant, tended to have higher h-Indexes than men with lower fWHRs, a finding consistent across all model estimations and robustness checks. Conversely, women with higher fWHRs and therefore more dominant personality traits tended to have lower h-Indexes than women with lower fWHRs, a finding again consistent across all model estimations and robustness checks. In sum, personality matters and it influences labour market outcomes such as performance.

These findings are vital in adding to an ever growing narrative regarding inherent biases in the work place as well as across genders. They suggest that innate features of our personality (traits) have direct implications on our performance in the labour market. Evolutionarily, we have developed to recognise and favour specific facial ratios and features. While individuals cannot change their faces and bone structure to adjust their fWHRs, evidence of this bias can aid in becoming more aware of such prejudices. Furthermore, the behavioural implications of this research could aid in educating and helping to hire a more diverse staff as well as also assisting in collection of certain desirable skills or traits. Employers might now consider personality when hiring (if they aren't already) as they are made aware of the impacts of personality on performance.

Limitations of this paper include unobservable variables not considered in the model such as outside influences like family, marriage, maternity leave, personal experience, and socio-demographic upbringing. We postulate these may have an impact on level of academic success in that they may affect the amount of time able to be dedicated to research, as well as intrinsic drive and devotion. Finally, the significance of findings pertaining to the undesirable nature of dominance in women may be decreasing with recent movements towards equality. Whilst currently social movements such as 'Me Too' evidence the fact that preferences towards gender norms exist, their ultimate goal is to bring awareness to such preferences and ideally minimise them. This would suggest as they gain traction, the underlying trends for dominant men and non-dominant women to succeed may in fact dissipate.

This paper could be extended along the following lines. First, we think that it would be interesting to compare the field of Economics to other fields. Second, comparing scholars in the same field but across countries could offer interesting insights also. Finally, controlling more explicitly for the quality of the co-author network could reduce any remaining omitted variable bias in our main estimation strategy.

A References

Abrevaya, J., and Hamermesh, D. (2012). Charity and Favoritism in the Field: Are Female Economists Nicer (To Each Other)? *The Review of Economics and Statistics*, **94(1)**.

Adams, J. (2013). The Fourth Age of Research. *Nature*, **497**.

Agarwal, S., Qian, W., Reeb, D., and Sing, T. (2016). Playing the Boys Game: Golf Buddies and Board Diversity. *American Economic Review*, **106(5)**, 272-276.

Apfelbaum, E., Phillips, K., and Richeson, J. (2014). Rethinking the Baseline in Diversity Research: Should we be Explaining the Effects of Homogeneity? *Perspectives on Psychological Science*, **9(3)**.

Almlund, M., Duckworth, A., Heckman, J., and Kautz, T. (2011). Personality Psychology and Economics. *Discussion Paper*, 5500.

Altschul, D., Robinson, L., Coleman, K., Capitanio, J., and 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(4-5)**, 532-552.

Andreoni, J., and Petrie, R. (2007). Beauty, Gender and Stereotypes: Evidence from Laboratory Experiments. *Journal of Economic Psychology*, **29(1)**, 73-93.

Angrist, J., and Pischke, J. (2010). The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics. *SSRN Electronic Journal*.

Astegiano, J., Sebastián-González, E., and Castanho, C. (2019). Unravelling the Gender Productivity Gap in Science: A Meta-Analytical Review. *Royal Society Open Science*, **6(6)**, 181566.

Azoulay, P., Zivin, J., and Wang, J. (2010). Superstar Extinction. *Quarterly Journal of Economics*, **125(2)**, 549-589.

Babcock, L., Recalde, M., Vesterlund, L., and Weingart, L. (2017). Gender Differences in Accepting and Receiving Requests for Tasks with Low Promotability. *American Economic Review*, **107(3)**.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018). Social Connectedness: Measurement, Determinants, and Effects. *Journal of Economic Perspectives*, **32(3)**, 259-280.

Bayer, A., and Rouse, C. (2016). Diversity in the Economics Profession: A New Attack on an Old Problem. *Journal of Economic Perspectives*, **30(4)**, 221-242.

Berdahl, J. (2007). The Sexual Harassment of Uppity Women. *Journal of Applied Psychology*, **92(2)**, 425-437.

Bertand, M. (2018). Coase Lecture – the Glass Ceiling. *Economica*, **85**.

Bertrand, M. (2011). *Handbook of Labor Economics* (4th ed., pp. 1543-1590). Elsevier.

Biddle, J., and Hamermesh, D. (1995). Beauty, Productivity, and Discrimination: Lawyers' Looks and Lucre. *Journal of Labor Economics*, **16(1)**, 172-201.

Bird, B., Cid Jofré, V., Geniole, S., Welker, K., Zilioli, S., and Maestripieri, D. et al. (2016). Does the Facial Width-to-Height Ratio Map onto Variability in Men's Testosterone Concentrations? *Evolution and Human Behavior*, **37(5)**, 392-398.

Blackaby, D., Booth, A., and Frank, J. (2005). Outside Offers and the Gender Pay Gap: Empirical Evidence from the UK Academic Labour Market. *The Economic Journal*, **115(501)**.

Blau, F., Currie, J., Croson, R., and Ginther, D. (2010). Can Mentoring Help Female Assistant Professors? Interim Results from a Randomized Trial. *American Economic Review*, **100(2)**, 348-352.

Blau, F., and Kahn, L. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, **55(3)**, 789-865.

Bohren, A., Imas, A., and Rosenberg, M. (2018). The Language of Discrimination: Using Experimental versus Observational Data. *AEA Papers and Proceedings*, **108**, 169-174.

Borghans, L., Duckworth, A., Heckman, J., and ter Wheel, B. (2008). The Economics and Psychology of Personality Traits. *The Journal of Human Resources*, **43(4)**.

Boschini, A., and Sjogren, A. (2007). Is Team Formation Gender Neutral? Evidence from Co-authorship Patterns. *Journal of Labor Economics*, **25(2)**.

Bosquet, C., and Combes, P. (2017). Sorting and Agglomeration Economies in French Economics Departments. *Journal of Urban Economics*, **101**.

Borjas, G., and Doran, K. (2015). Which Peers Matter? The Relative Impacts of Collaborators, Colleagues, and Competitors. *Review of Economics and Statistics*, **97(5)**, 1104-1117.

Bouchard Jr, T., Lykken, D., McGue, M., Segal, N., and Tellegen, A. (1990). Sources of Human Psychological Differences: The Minnesota Study of Twins Reared Apart. *Science*, **250(4978)**.

Boustan, L., and Langan, A. (2019). Variation in Women's Success across PhD Programs in Economics. *Journal of Economic Perspectives*, **33(1)**, 23-42.

Braakmann, N. (2009). The Role of Psychological Traits for the Gender Gap in Full-Time Employment and Wages: Evidence from Germany. *SSRN Electronic Journal*.

Brogaard, J., Engelberg, J., and Parsons, C. (2014). Networks and Productivity: Causal Evidence from Editor Rotations. *Journal of Financial Economics*, **111(1)**, 251-270.

Bruce, V., and Young, A. (2012). Face Perception. New York, NY: *Psychology Press.*

Buchan, N., Croson, R., and Solnick, S. (2008). Trust and Gender: An Examination of Behavior and Beliefs in the Investment Game. *Journal of Economic Behavior and Organization*, **68(3-4)**, 466-476.

Buckles, K. (2019). Fixing the Leaky Pipeline: Strategies for Making Economics Work for Women at Every Stage. *Journal of Economic Perspectives*, **33(1)**, 43-60.

Burgess, D., and Borgida, E. (1999). Who Women are, who Women Should be: Descriptive and Prescriptive Gender Stereotyping in Sex Discrimination. *Psychology, Public Policy, and Law*, **5(3)**, 665-692.

Burt, S. (2008). Genes and Popularity Evidence of an Evocative Gene-Environment Correlation. *Psychological Science*, **19(2)**.

Bẗikofer, A., Figlio, D., Karbownik, K., Kuzawa, C., and Salvanes, K. (2019). Evidence that Prenatal Testosterone Transfer from Male Twins Reduces the Fertility and Socioeconomic Success of their Female Co-Twins. *Proceedings of the National Academy of Sciences*, **116(14)**.

Card, D., and DellaVigna, S. (2013). Nine Facts about Top Journals in Economics. *Journal of Economic Literature*, **51(1)**, 144-161.

Card, D., DellaVigna, S., Funk, P., and Iriberri, N. (2020). Are Referees and Editors in Economics Gender Neutral? *The Quarterly Journal of Economics*, **135(1)**, 269-327.

Carré, J., and McCormick, C. (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.

Carré, J., McCormick, C., and Mondloch, C. (2009). Facial Structure is a Reliable Cue of Aggressive Behavior. *Psychological Science*, **20(10)**, 1194-1198.

Carré, J., Morrissey, M., Mondloch, C., and McCormick, C. (2010). Estimating Aggression from Emotionally Neutral Faces: Which Facial Cues are Diagnostic? *Perception*, **39(3)**, 356-377.

Charles, K., Guryan, J., and Pan, J. (2018). The Effects of Sexism on American Women: The Role of Norms vs. Discrimination. *SSRN Electronic Journal*.

Cimpian, A., and Leslie, S. (2015). Response to Comment on "Expectations of Brilliance Underlie Gender Distributions across Academic Disciplines". *Science*, **349(6246)**, 391-391.

Clarke, K. (2005). The Phantom Menace: Omitted Variable Bias in Econometric Research. *Conflict Management and Peace Science*, **22(4)**, 341-352.

Cobb-Clark, D., and Schurer, S. (2012). The Stability of Big-Five Personality Traits. *Economics Letters*, **115(1)**, 11-15.

Collins, S. (2000). Minority Groups in the Economics Profession. *Journal of Economic Perspectives*, **14(2)**, 133-148.

Combes, P., Linnemer, L., and Visser, M. (2008). Publish or Peer-Rich? The Role of Skills and Networks in Hiring Economics Professors. *Labour Economics*, **15**.

Conley, J., and Önder, A. (2014). The Research Productivity of New PhDs in Economics: The Surprisingly High Non-Success of the Successful. *Journal of Economic Perspectives*, **28(3)**, 205-216.

Conley, J., Crucini, M., Driskill, R., and Önder, A. (2011). The Effects of Publication Lags On Life-Cycle Research Productivity in Economics. *Economic Inquiry*, **51(2)**, 1251-1276.

Cortes, P., and Pan, J. (2018). *The Oxford Handbook of Women and the Economy: Occupation and Gender* (pp. 425-452). Oxford University Press.

Costrich, N., Feinstein, J., Kidder, L., Marecek, J., and Pascale, L. (1975). When Stereotypes Hurt: Three Studies of Penalties for Sex-Role Reversals. *Journal of Experimental Social Psychology*, **11(6)**, 520-530.

Craig, B., Nelson, N., and Dixson, B. (2019). Sexual Selection, Agonistic Signaling, and the Effect of Beards on Recognition of Men's Anger Displays. *Psychological Science*, **30(5)**, 728-738.

Croson, R., and Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*, **47(2)**, 448-474.

Cunha, F., Heckman, J., and Schennach, S. (2007). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, **43(4)**, 738-782.

Currarini, S., Jackson, M., and Pin, P. (2009). An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica*, **77(4)**.

Dall'Ara, E., and Maass, A. (1999). Studying Sexual Harassment in the Laboratory: Are Egalitarian Women at Higher Risk? *Sex Roles*, **41(9/10)**.

Del Giudice, M., Booth, T., and Irwing, P. (2012). The Distance Between Mars and Venus: Measuring Global Sex Differences in Personality. *Plos ONE*, **7(1)**.

Delgadillo, L. (2016). Quantifying the Impact of Research for Tenure and Promotion in Family and Consumer Sciences. *Family and Consumer Sciences Research Journal*,

44(3), 255-263.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review*, **100**.

Druckenmiller, H., and Hsiang, S. (2019). Accounting for Unobservable Heterogeneity in Cross Section Using Spatial First Differences. *NBER*, Working Paper(25177).

Ductor, L., Goyal, S., and Prummer, A. (2018). Gender and Collaboration. *Working Paper*, **856**.

Eckel, C., and Petrie, R. (2011). Face Value. *American Economic Review*, **101(4)**, 1497-1513.

Einav, L., and Yariv, L. (2006). What's in a Surname? The Effects of Surname Initials on Academic Success. *Journal of Economic Perspectives*, **20(1)**, 175-188.

Ellison, G. (2013). How Does the Market Use Citation Data? The Hirsch Index in Economics. *American Economic Journal: Applied Economics*, **5(3)**, 63-90.

Ellison, G. (2002). The Slowdown of the Economics Publishing Process. *SSRN Electronic Journal*.

Fafchamps, M., Goyal, S., and van der Leij, M. (2010). Matching and Network Effects. *Journal of the European Economic Association*, **8(1)**.

Fang, F., Clapham, P., and Chung, K. (2011). A Systematic Review of Interethnic Variability in Facial Dimensions. *Plastic and Reconstructive Surgery*, **127(2)**, 874-881.

Ferber, M., and Teiman, M. (1980). Are Women Economists at a Disadvantage in Publishing Journal Articles? *Eastern Economic Journal*, **6(3/4)**.

Ferguson, C. (2010). A Meta-Analysis of Normal and Disordered Personality across the Life Span. *Journal of Personality and Social Psychology*, **98(4)**, 659-667.

Ferraro, P., and Miranda, J. (2014). The Performance of Non-Experimental Designs in the Evaluation of Environmental Programs: A Design-Replication Study using a Large-Scale Randomized Experiment as a Benchmark. *Journal of Economic Behavior and Organization*, **107**, 344-365.

Fidrmuc, J., and Paphawasit, B. (2018). Beautiful Minds: Physical Attractiveness and Research Productivity in Economics. *Semantic Scholar*.

Fletcher, J. (2013). The Effects of Personality Traits on Adult Labor Market Outcomes: Evidence from Siblings. *Journal of Economic Behavior and Organization*, **89**, 122-135.

Flinn, C., Todd, P., and Zhang, W. (2018). Personality Traits, Intra-Household Allocation and the Gender Wage Gap. *European Economic Review*, **109**, 191-220.

Fowler, J., Dawes, C., and Christakis, N. (2009). Model of Genetic Variation in Human Social Networks. *Proceedings of the National Academy of Sciences of the United States of America*, **106(6)**.

Freeman, R., and Huang, W. (2014). Strength in Diversity. *Nature*, **513**.

Freeman, R., and Huang, W. (2015). Collaborating with People Like Me: Ethnic Coauthorship within the United States. *Journal of Labor Economics*, **33(3)**.

French, M., Robins, P., Homer, J., and Tapsell, L. (2009). Effects of Physical Attractiveness, Personality, and Grooming on Academic Performance in High School. *Labour Economics*, **16(4)**, 373-382.

Geniole, S., Denson, T., Dixson, B., Carré, J., and McCormick, C. (2015). Evidence from Meta-Analyses of the Facial Width-to-Height Ratio as an Evolved Cue of Threat. *PLos ONE*, **10(7)**, e0132726.

Geraci, L., Balsis, S., and Busch, A. (2015). Gender and the H- Index in Psychology. *Scientometrics*, **105(3)**, 2023-2034.

Ginther, D. (2002). Women in the Economics Pipeline: How do we Compare to Other Disciplines? *Committee On the Status of Women in the Economics Profession*, 6-7.

Ginther, D. (2003). Is MIT an Exception? Gender Pay Differences in Academic Science. *Bulletin of Science, Technology and Society*, **23(1)**, 21-26.

Ginther, D., and Kahn, S. (2015). Comment on "Expectations of Brilliance Underlie Gender Distributions across Academic Disciplines". *Science*, **349(6246)**, 391-391.

Goldin, C., Kerr, S., Olivetti, C., and Barth, E. (2017). The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census. *American Economic Review*, **107(5)**, 110-114.

Gordon, N., Morton, T., and Braden, I. (1974). Faculty Salaries: Is There Discrimination by Sex, Race, and Discipline? *The American Economic Review*, **64(3)**.

Gorodnichenko, Y., Pham, T., and Talavera, O. (2019). Conference Presentations and Academic Publishing. *IZA*, 12573.

Hahn, T., Winter, N., Anderl, C., Notebaert, K., Wuttke, A., Clément, C., and Windmann, S. (2017). Facial Width-to-Height Ratio Differs by Social Rank across Organizations, Countries, and Value Systems. *PLos ONE*, **12(11)**, e0187957.

Haley, M. (2017). On the Inauspicious Incentives of the Scholar-Level H-Index: an Economist's Take on Collusive and Coercive Citation. *Applied Economics Letters*, **24(2)**, 85-89.

Hamermesh, D., and Biddle, J. (1994). Beauty and the Labour Market. *The American Economic Review*, **84(5)**.

Hamermesh, D., Meng, X., and Zhang, J. (2002). Dress for Success—does Primping Pay? *Labour Economics*, **9(3)**, 361-373.

Hamermesh, D. (2013). Six Decades of Top Economics Publishing: Who and How? *Journal of Economic Literature*, **51(1)**, 162-172.

Harris, M., Brett, C., Johnson, W., and Deary, I. (2016). Personality Stability from Age 14 to Age 77 Years. *Psychology and Aging*, **31(8)**, 862-874.

Haselhuhn, M., Ormiston, M., and Wong, E. (2015). Men's Facial Width-to-Height Ratio Predicts Aggression: A Meta-Analysis. *PLos ONE*, **10(4)**, e0122637.

Haselhuhn, M., Wong, E., and Ormiston, M. (2013). Self-Fulfilling Prophecies as a Link between Men's Facial Width-to-Height Ratio and Behavior. *PLos ONE*, **8(8)**, e72259.

Haselhuhn, M., Wong, E., Ormiston, M., Inesi, M., and Galinsky, A. (2014). Negotiating Face-to-Face: Men's Facial Structure Predicts Negotiation Performance. *The Leadership Quarterly*, **25(5)**, 835-845.

He, X., Yin, H., Zeng, Y., Zhang, H., and Zhao, H. (2019). Facial Structure and Achievement Drive: Evidence from Financial Analysts. *Journal of Accounting Research*, **57(4)**, 1013-1057.

Heckman, J., and Moktan, S. (2018). Publishing and Promotion in Economics: The Tyranny of the Top Five. *SSRN Electronic Journal*.

Heckman, J., and Raut, L. (2016). Intergenerational Long-Term Effects of Preschool-Structural Estimates from a Discrete Dynamic Programming Model. *Journal of Econometrics*, **191(1)**, 164-175.

Heckman, J., and Rubinstein, Y. (2001). The Importance of Noncognitive Skills:

Lessons from the GED Testing Program. *American Economic Review*, **91(2)**, 145-149.

Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, **103(6)**, 2052-2086.

Hehman, E., Leitner, J., Deegan, M., and Gaertner, S. (2015). Picking Teams: When Dominant Facial Structure is Preferred. *Journal of Experimental Social Psychology*, **59**, 51-59.

Heilman, M., Wallen, A., Fuchs, D., and Tamkins, M. (2004). Penalties for Success: Reactions to Women who Succeed at Male Gender-Typed Tasks. *Journal of Applied Psychology*, **89(3)**, 416-427.

Hengel, E. (2017). Publishing while Female. Are Women Held to Higher Standards? Evidence from Peer Review. *Cambridge Working Papers in Economics.*

Hengel, E. (2019). Gender Differences in Citations at Top Economics Journals. *Mimeo.*

Hilmer, C., and Hilmer, M. (2007). Women Helping Women, Men Helping Women? Same-Gender Mentoring, Initial Job Placements, and Early Career Publishing Success for Economics PhDs. *The American Economic Review*, **97(2)**.

Hilmer, C., and Hilmer, M. (2010). Are there Gender Differences in the Job Mobility Patterns of Academic Economists? *American Economic Review: Papers and Proceedings*, **100**.

Hilmer, M., Ransom, M., and Hilmer, C. (2011). Fame and the Fortune of Academic Economists: How the Market Rewards Influential Research in Economics. *Southern Economic Journal*, **82(2)**, 430-452.

Hirsch, J. (2005). An Index to Quantify an Individual's Scientific Research Output. *Proceedings of the National Academy of Sciences*, **102(46)**, 16569-16572.

Holliday, E., Jagsi, R., Wilson, L., Choi, M., Thomas, C., and Fuller, C. (2014). Gender Differences in Publication Productivity, Academic Position, Career Duration, and Funding Among U.S. Academic Radiation Oncology Faculty. *Academic Medicine*, **89(5)**, 767-773.

Johnson, S., Podratz, K., Dipboye, R., and Gibbons, E. (2010). Physical Attractiveness Biases in Ratings of Employment Suitability: Tracking Down the "Beauty is Beastly" Effect. *The Journal of Social Psychology*, **150(3)**, 301-318.

Kachur, A., Osin, E., Davydov, D., Shutilov, K., and Novokshonov, A. (2020). Assessing the Big Five Personality Traits Using Real-Life Static Facial Images. *SSRN Electronic Journal*.

Kahn, S. (1995). Women in the Economics Profession. *Journal of Economic Perspectives*, **9(4)**, 193-205.

Kanazawa, S., and Still, M. (2018). Is there Really a Beauty Premium or an Ugliness Penalty on Earnings? *Journal of Business and Psychology*, **33(2)**, 249-262.

Kausel, E., Ventura, S., Vargas, M., Díaz, D., and Vicencio, F. (2018). Does Facial Structure Predict Academic Performance? *Personality and Individual Differences*, **129**, 1-5.

Kerr, W. (2008). Ethnic Scientific Communities and International Technology Diffusion. *The Review of Economics and Statistics*, **90(3)**.

Kramer, R. (2017). Sexual Dimorphism of Facial Width-to-Height Ratio in Human Skulls and Faces: A Meta-Analytical Approach. *Evolution and Human Behavior*, **38(3)**, 414-420.

Krause, J., James, R., and Croft, D. (2010). Personality in the Context of Social Networks. *Philosophical Transactions of the Royal Society*, 365.

Lefevre, C., Etchells, P., Howell, E., Clark, A., and Penton-Voak, I. (2014). Facial Width-to-Height Ratio Predicts Self-Reported Dominance and Aggression in Males and Females, but a Measure of Masculinity does not. *Biology Letters*, **10(10)**, 20140729.

Leslie, S., Cimpian, A., Meyer, M., and Freeland, E. (2015). Expectations of Brilliance Underlie Gender Distributions across Academic Disciplines. *Science*, **347(6219)**, 262-265.

Lewis, K., Gonzalez, M., and Kaufman, J. (2012). Social Selection and Peer Influence in an Online Social Network. *Proceedings of the National Academy of Sciences*, **109(1)**, 68-72.

Lewis, G., Lefevre, C., and Bates, T. (2012). Facial Width-to-Height Ratio Predicts Achievement Drive in US Presidents. *Personality and Individual Differences*, **52(7)**, 855-857.

Lin, C., Adolphs, R., and Alvarez, R. (2018). Inferring Whether Officials are Corruptible from Looking at their Faces. *Psychological Science*, **29(11)**, 1807-1823.

Lindenlaub, I., and Prummer, A. (2020). Network Structure and Performance. *The Economic Journal*, forthcoming.

Liu, Y., Lu, H., and Veenstra, K. (2018). Beauty and Academic Career. *SSRN Electronic Journal*.

Lu, Y., and Teo, M. (2018). Do Alpha Males Deliver Alpha? Testosterone and Hedge Funds. *SSRN Electronic Journal*.

Lundberg, S. (2020). Women in Economics. *CEPR Press*.

Lundberg, S., and Stearns, J. (2019). Women in Economics: Stalled Progress. *Journal of Economic Perspectives*, **33(1)**, 3-22.

Maass, A., Cadinu, M., Guarnieri, G., and Grasselli, A. (2003). Sexual Harassment under Social Identity Threat: The Computer Harassment Paradigm. *Journal of Personality and Social Psychology*, **85(5)**, 853-870.

Mairesse, J., and Turner, L. (2005). Measurement and Explanation of the Intensity of Co-Publication in Scientific Research: An Analysis at the Laboratory Level. *NBER Working Paper*, 11172.

Mayer, A., and Puller, S. (2008). The Old Boy (and Girl) Network: Social Network Formation on University Campuses. *Journal of Public Economics*, **92(1-2)**, 329-347.

McAdams, D., and Pals, J. (2006). A New Big Five: Fundamental Principles for an Integrative Science of Personality. *American Psychologist*, **61(3)**, 204-217.

McDowell, J., and Kiholm Smith, J. (2020). The Effect of Gender-Sorting on Propensity to Coauthor: Implications for Academic Promotion. *Economic Inquiry*, forthcoming.

McDowell, J., Singell JR, L., and Stater, M. (2007). Two to Tango? Gender Differences in the Decisions to Publish and Coauthor. *Economic Inquiry*, 44(1).

McGugin, R., and Gauthier, I. (2013). Face Recognition. *The Oxford Handbook of Cognitive Neuroscience*, Volume 2: The Cutting Edges., 2, 165.

McPherson, M., Smith-Lovin, L., and Cook, J. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, **27(1)**, 415-444.

Mileva, V., Cowan, M., Cobey, K., Knowles, K., and 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.

Mueller, G., and Plug, E. (2006). Estimating the Effect of Personality on Male and Female Earnings. *ILR Review*, **60(1)**, 3-22.

Niederle, M. (2017). A Gender Agenda: A Progress Report on Competitiveness. *American Economic Review*, **107(5)**, 115-119.

Niederle, M., and Vesterlund, L. (2007). Do Women Shy Away from Competition? Do Men Compete too Much? *The Quarterly Journal of Economics*, **122(3)**, 1067-1101.

Niederle, M., and Vesterlund, L. (2011). Gender and Competition. *Annual Review of Economics*, **3**.

Palumbo, R., Fairfield, B., Mammarella, N., and Di Domenico, A. (2017). Does Make-Up Make you Feel Smarter? The "Lipstick Effect" Extended to Academic Achievement. *Cogent Psychology*, **4(1)**.

Parkins, I., Fishbein, H., and Ritchey, P. (2006). The Influence of Personality on Workplace Bullying and Discrimination. *Journal of Applied Social Psychology*, **36(10)**, 2554-2577.

Phillips, A., and MacKinnon, C. (2009). *Feminism and Politics* (pp. 295-313). Oxford University. Press.

Power, R., and Pluess, M. (2015). Heritability Estimates of the Big Five Personality Traits Based on Common Genetic Variants. *Translational Psychiatry*, **5(7)**, e604-e604.

Risse, L., Farrell, L., and Fry, T. (2018). Personality and Pay: Do Gender Gaps in Confidence Explain Gender Gaps in Wages? *Oxford Economic Papers*, **70(4)**, 919-949.

Roberts, B. (2009). Back to the Future: Personality and Assessment and Personality Development. *Journal of Research in Personality*, **43(2)**, 137-145.

Røysamb, E., Nes, R., Czajkowski, N., and Vassend, O. (2018). Genetics, Personality and Wellbeing. A Twin Study of Traits, Facets and Life Satisfaction. *Scientific Reports*, **8(1)**.

Schmitt, D., Realo, A., Voracek, M., and Allik, J. (2008). Why Can't a Man be More Like a Woman? Sex Differences in Big Five Personality Traits across 55 Cultures. *Journal of Personality and Social Psychology*, **94(1)**.

Schweizer, D., and Karami, M. (2018). Do Risk Takers Look Different? The Effect of Facial Masculinity on Hedge Fund Managers' Risk-Taking Behavior. *Preliminary Version.*

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., and Gurven, M. (2008). Human Adaptations for the Visual Assessment of Strength and Fighting Ability from the Body and Face. *Proceedings of the Royal Society B: Biological Sciences*, **276(1656)**, 575-584.

Sparacino, J., and Hansell, S. (1979). Physical Attractiveness and Academic Performance: Beauty is not always Talent. *Journal of Personality*, **47(3)**, 449-469.

Stuart, E. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science*, **25(1)**, 1-21.

Sutherland, C., Young, A., Mootz, C., and Oldmeadow, J. (2014). Face Gender and Stereotypicality Influence Facial Trait Evaluation: Counter-Stereotypical Female Faces are Negatively Evaluated. *British Journal of Psychology*, **106(2)**, 186-208.

Symonds, M., Gemmell, N., Braisher, T., Gorringe, K., and Elgar, M. (2006). Gender Differences in Publication Output: Towards an Unbiased Metric of Research Performance. *PLos ONE*, **1(1)**, e127.

Talamas, S., Mavor, K., and Perrett, D. (2016). Blinded by Beauty: Attractiveness Bias and Accurate Perceptions of Academic Performance. *Public Library of Science ONE*, **11(2)**, e0148284.

Tinsley, C., Cheldelin, S., Schneider, A., and Amanatullah, E. (2009). Women at the Bargaining Table: Pitfalls and Prospects. *Negotiation Journal*, **25(2)**, 233-248.

Turkheimer, E. (2000). Three Laws of Behavior Genetics and What They Mean. *Current Directions in Psychological Science*, **9(5)**, 160-164.

Valentine, K., Li, N., Penke, L., and Perrett, D. (2014). Judging a Man by the Width of His Face. *Psychological Science*, **25(3)**, 806-811.

Wang, D., Nair, K., Kouchaki, M., Zajac, E., and Zhao, X. (2019). A Case of Evolutionary Mismatch? Why Facial Width-to-Height Ratio may not Predict Behavioral Tendencies. *Psychological Science*, **30(7)**, 1074-1081.

Watkins, C., Jones, B., and DeBruine, L. (2010). Individual Differences in Dominance Perception: Dominant Men are Less Sensitive to Facial Cues of Male Dominance. *Personality and Individual Differences*, **49(8)**, 967-971.

Weisberg, Y., DeYoung, C., and Hirsh, J. (2011). Gender Differences in Personality across the Ten Aspects of the Big Five. *Frontiers in Psychology*, **2(178)**.

Wilson, V., Lefevre, C., Morton, F., Brosnan, S., Paukner, A., and Bates, T. (2014). Personality and Facial Morphology: Links to Assertiveness and Neuroticism in Capuchins (Sapajus [Cebus] Apella). *Personality and Individual Differences*, **58**, 89-94.

Wong, J., and Penner, A. (2016). Gender and the Returns to Attractiveness. *Research in Social Stratification and Mobility*, **44**, 113-123.

Wooldridge, J. (2010). Econometric Analysis of Cross Section and Panel Data. *MIT Press*.

Wuchty, S., Jones, B., and Uzzi, B. (2007). The Increasing Dominance of Teams in Production of Knowledge. *Science*, **316(5827)**, 1036-1039.

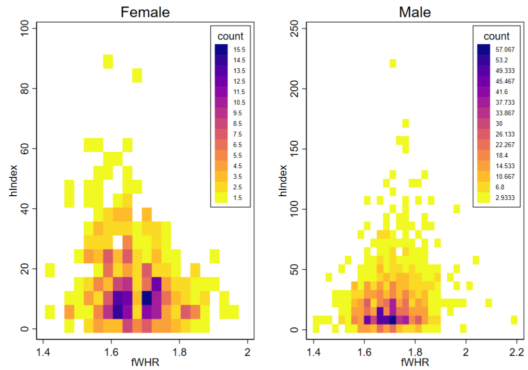
Zebrowitz, L., Andreoletti, C., Collins, M., Lee, S., and Blumenthal, J. (1998). Bright, Bad, Babyfaced Boys: Appearance Stereotypes do not Always Yield Self-Fulfilling Prophecy Effects. *Journal of Personality and Social Psychology*, **75(5)**, 1300-1320.

Zhang, W., Hahn, A., Cai, Z., Lee, A., Holzleitner, I., DeBruine, L., and Jones, B. (2018). No Evidence that Facial Width-to-Height Ratio (fWHR) is Associated with Women's Sexual Desire. *Public Library of Science ONE*, **13(7)**, e0200308.

Zilioli, S., Sell, A., Stirrat, M., Jagore, J., Vickerman, W., and Watson, N. (2015). Face of a Fighter: Bizygomatic Width as a Cue of Formidability. *Aggressive Behavior*, **41(4)**, 322-330.

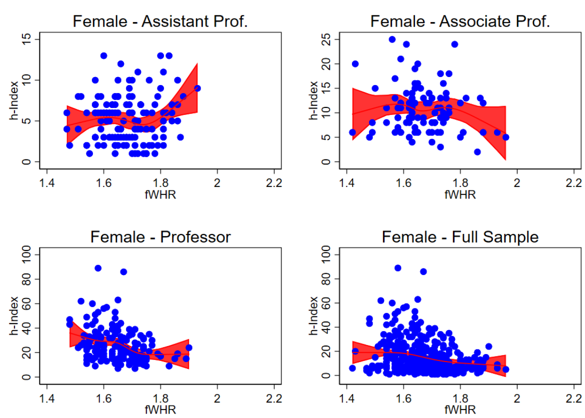
B Tables and Figures

Figure 4: Density Graph.



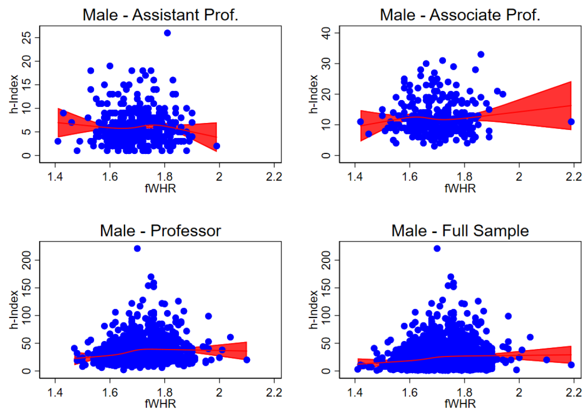
Notes: Heat map for the relationship between fWHR and the h-Index for female (left panel) and male (right panel).

Figure 5: Restricted Cubic Spline - Rank - Female.



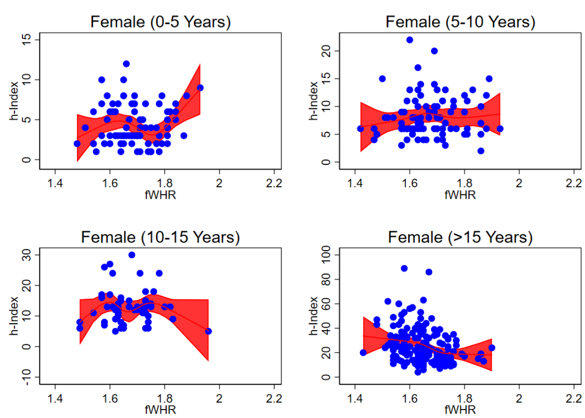
Notes: Restricted cubic spline estimation for the relationship between fWHR and the h-Index for female academics by rank.

Figure 6: Restricted Cubic Spline - Rank - Male.



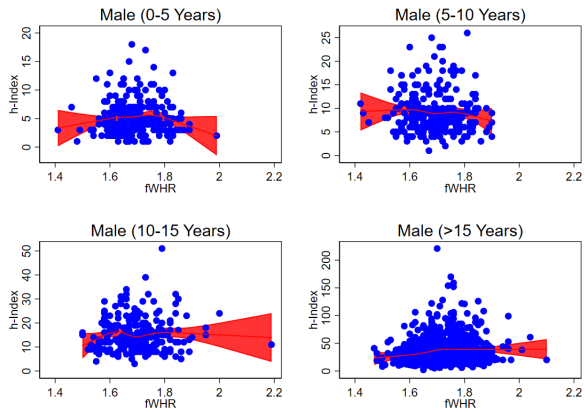
Notes: Restricted cubic spline estimation for the relationship between fWHR and the h-Index for male academics by rank.

Figure 7: Restricted Cubic Spline - Time - Female.



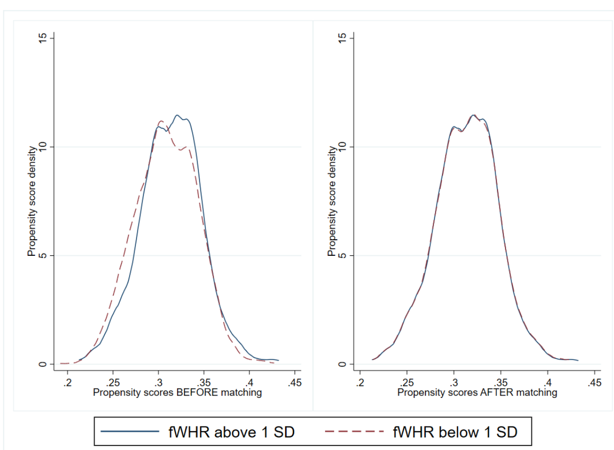
Notes: Restricted cubic spline estimation for the relationship between fWHR and the h-Index for female academics by time since PhD.

Figure 8: Restricted Cubic Spline - Time - Male.



Notes: Restricted cubic spline estimation for the relationship between fWHR and the h-Index for male academics by time since PhD.

Figure 9: Balance checks for total sample and matched sample.



Notes: Kernel density estimates of treated and non-treated groups, before and after matching. Propensity score matching on the following variables: Time since PhD, Co-Authors, White, Switches, Field, and Rank.

Table 10: Population and Sample Coverage by University.

		Population			Sample		
Rank	University	Total	Male	Female	Total (%)	Male (%)	Female (%)
1	Harvard	51	44	7	43 (84)	37 (84)	6 (86)
2	MIT	37	32	5	35 (95)	30 (94)	5 (100)
3	Berkeley	52	42	10	44 (85)	34 (81)	10 (100)
4	Chicago	35	33	2	30 (86)	28 (85)	2 (100)
5	Princeton	60	50	10	46 (77)	37 (74)	9 (90)
6	Stanford	43	35	8	39 (91)	31 (89)	8 (100)
7	Columbia	48	40	8	42 (88)	34 (85)	8 (100)
8	Brown	35	28	7	30 (86)	25 (89)	5 (71)
9	Yale	53	43	10	47 (89)	37 (86)	10 (100)
10	NYU	50	45	5	33 (66)	28 (62)	5 (100)
11	Boston University	43	37	6	38 (88)	32 (86)	6 (100)
12	U Pennsylvania	33	28	5	32 (97)	27 (96)	5 (100)
13	Dartmouth	32	25	7	29 (91)	22 (88)	7 (100)
14	UCSD	41	34	7	40 (98)	33 (97)	7 (100)
15	Northwestern	36	31	5	35 (97)	30 (97)	5 (100)
16	UCLA	43	33	10	40 (93)	30 (91)	10 (100)
17	U Michigan	42	35	7	37 (88)	30 (86)	7 (100)
18	Boston College	31	30	1	31 (100)	30 (100)	1 (100)
19	USC	29	23	6	26 (90)	20 (87)	6 (100)
20	UC Davis	35	25	10	34 (97)	24 (96)	10 (100)
21	Wisconsin-Madison	38	33	5	34 (89)	29 (88)	5 (100)
22	Michigan State	48	39	9	46 (96)	37 (95)	9 (100)
23	Duke	53	47	6	46 (87)	41 (87)	5 (83)
24	Georgetown	33	23	10	27 (82)	18 (78)	9 (90)
25	Vanderbilt	34	26	8	32 (94)	24 (92)	8 (100)
26	U Maryland	33	23	10	32 (97)	22 (96)	10 (100)
27	UC Irvine	29	23	6	28 (97)	22 (96)	6 (100)
28	Cornell	43	35	8	33 (77)	25 (71)	8 (100)
29	Penn State	40	33	7	35 (88)	30 (91)	5 (71)
30	Arizona State	32	26	6	26 (81)	21 (81)	5 (83)
31	U Texas (Austin)	34	25	9	29 (85)	21 (84)	8 (89)
32	UCSB	29	23	6	28 (97)	22 (96)	6 (100)
33	U Virginia	34	28	6	32 (94)	26 (93)	6 (100)
34	Rutgers	33	25	8	29 (88)	21 (84)	8 (100)
35	Johns Hopkins	21	18	3	18 (86)	15 (83)	3 (100)
36	Notre Dame	35	26	9	35 (100)	26 (100)	9 (100)
37	Ohio State	27	23	4	21 (78)	18 (78)	3 (75)
38	Washington U St. Louis	24	20	4	19 (79)	15 (75)	4 (100)
39	U Colorado	30	23	7	22 (73)	18 (78)	4 (57)
40	U Minnesota	22	18	4	21 (95)	18 (100)	3 (75)

Table 11: Population and Sample Coverage by University (cont'd).

		Population			Sample		
Rank	University	Total	Male	Female	Total (%)	Male (%)	Female (%)
41	UC Santa Cruz	23	15	8	20 (87)	13 (87)	7 (88)
42	Georgia State	28	22	6	24 (86)	20 (91)	4 (67)
43	George Washington	33	24	9	18 (55)	12 (50)	6 (67)
44	Williams College	28	21	7	23 (82)	16 (76)	7 (100)
45	U Washington	22	16	6	21 (95)	15 (94)	6 (100)
46	Texas A&M	31	21	10	29 (94)	19 (90)	10 (100)
47	U Pittsburgh	27	16	11	27 (100)	16 (100)	11 (100)
48	Chapman U	15	14	1	6 (40)	6 (43)	0 (0)
49	Caltech	18	14	4	15 (83)	11 (79)	4 (100)
50	Iowa State	32	27	5	32 (100)	27 (100)	5 (100)
51	U Oregon	26	22	4	26 (100)	22 (100)	4 (100)
52	George Mason	33	30	3	23 (70)	20 (67)	3 (100)
53	Indiana	28	21	7	21 (75)	15 (71)	6 (86)
54	Carnegie Mellon	22	16	6	21 (95)	15 (94)	6 (100)
55	Brandeis	15	9	6	11 (73)	6 (67)	5 (83)
56	U Illinois Urbana Champaign	26	21	5	24 (92)	19 (90)	5 (100)
57	U Kentucky	21	16	5	19 (90)	14 (88)	5 (100)
58	Clemson	26	23	3	20 (77)	18 (78)	2 (67)
59	Tufts	23	15	8	23 (100)	15 (100)	8 (100)
60	Purdue	29	26	3	26 (90)	23 (88)	3 (100)
61	U Missouri	19	15	4	17 (89)	13 (87)	4 (100)
62	U Connecticut	27	19	8	24 (89)	16 (84)	8 (100)
63	Emory	28	21	7	24 (86)	18 (86)	6 (86)
64	U Houston	21	15	6	20 (95)	14 (93)	6 (100)
65	Syracuse	23	17	6	23 (100)	17 (100)	6 (100)
66	U Rochester	22	18	4	19 (86)	15 (83)	4 (100)
67	UC Riverside	21	15	6	18 (86)	12 (80)	6 (100)
68	Florida State	28	24	4	21 (75)	17 (71)	4 (100)
69	UNC Chapel Hill	33	27	6	25 (76)	19 (70)	6 (100)
70	U Wyoming	14	12	2	13 (93)	11 (92)	2 (100)
71	U Arizona	22	18	4	19 (86)	15 (83)	4 (100)
72	Rice	23	16	7	23 (100)	16 (100)	7 (100)
73	UIC	17	13	4	17 (100)	13 (100)	4 (100)
74	Stony Brook SUNY	17	13	4	17 (100)	13 (100)	4 (100)
75	Southern Methodist U	19	17	2	15 (79)	14 (82)	1 (50)
76	U Alabama	20	17	3	19 (95)	16 (94)	3 (100)
77	Wellesley College	18	10	8	16 (89)	9 (90)	7 (88)
78	CUNY	51	38	13	35 (69)	23 (61)	12 (92)
79	Tulane	16	12	4	16 (100)	12 (100)	4 (100)
80	Binghamton	21	17	4	16 (76)	13 (76)	3 (75)

Table 12: Population and Sample Coverage by University (cont'd).

		Population			Sample		
Rank	University	Total	Male	Female	Total (%)	Male (%)	Female (%)
81	Drexel	19	15	4	18 (95)	14 (93)	4 (100)
82	Wisconsin-Milwaukee	17	14	3	16 (94)	13 (93)	3 (100)
83	U Georgia	20	16	4	19 (95)	15 (94)	4 (100)
84	U Hawaii	19	14	5	17 (89)	13 (93)	4 (80)
85	Appalachian State	18	17	1	14 (78)	14 (82)	0 (0)
86	American U	29	21	8	21 (72)	15 (71)	6 (75)
87	Brigham Young	23	21	2	23 (100)	21 (100)	2 (100)
88	U Kansas	16	13	3	14 (88)	11 (85)	3 (100)
89	U Texas (Dallas)	11	9	2	8 (73)	7 (78)	1 (50)
90	Middlebury College	25	15	10	20 (80)	12 (80)	8 (80)
91	NC State	18	15	3	16 (89)	13 (87)	3 (100)
92	U Florida	17	16	1	13 (76)	12 (75)	1 (100)
93	College of William and Mary	22	18	4	21 (95)	18 (100)	3 (75)
94	Claremont McKenna	29	22	7	27 (93)	21 (95)	6 (86)
95	New School	11	8	3	9 (82)	6 (75)	3 (100)
96	U California (Merced)	15	9	6	12 (80)	6 (67)	6 (100)
97	U Miami	13	11	2	13 (100)	11 (100)	2 (100)
98	Auburn	16	11	5	10 (63)	7 (64)	3 (60)
99	West Virginia	14	13	1	11 (79)	10 (77)	1 (100)
100	SUNY Albany	19	15	4	19 (100)	15 (100)	4 (100)
	Total	2,838	2,266	572	2,471 (87)	1,940 (86)	531 (93)

Notes: Ranking taking from ideas.repec in September 2020. Percentages shown in parenthesis.

Table 13: Additional Summary Statistics.

Rank	First Job	Obs. (%)	PhD University	Obs. (%)	Nationality	Obs. (%)
1	Princeton	81 (3.3)	Harvard	193 (7.8)	US	1,318 (53.5)
2	Harvard	78 (3.2)	MIT	172 (7.0)	China	144 (5.8)
3	Chicago	53 (2.2)	Berkeley	139 (5.6)	Italy	80 (3.3)
4	Yale	53 (2.2)	Stanford	137 (5.5)	Germany	75 (3.0)
5	Northwestern	50 (2.1)	Yale	116 (4.7)	Canada	71 (2.9)
6	Stanford	50 (2.1)	Princeton	111 (4.5)	UK	71 (2.9)
7	U Pennsylvania	50 (2.1)	Chicago	106 (4.3)	India	69 (2.8)
8	Berkeley	45 (1.8)	Northwestern	86 (3.5)	South Korea	51 (2.1)
9	Columbia	45 (1.8)	U Pennsylvania	74 (3.0)	Argentina	47 (1.9)
10	U Michigan	42 (1.7)	U Michigan	70 (2.8)	France	47 (1.9)
11	MIT	41 (1.7)	Minnesota	63 (2.6)	Spain	41 (1.7)
12	UCSD	33 (1.4)	Wisconsin-Madison	60 (2.4)	Israel	37 (1.5)
13	UCLA	33 (1.4)	Columbia	53 (2.1)	Russia	35 (1.4)
14	Wisconsin-Madison	33 (1.4)	UCSD	51 (2.1)	Turkey	35 (1.4)
15	NYU	32 (1.3)	NYU	47 (1.9)	Australia	26 (1.1)
16	Boston U	30 (1.2)	Rochester	43 (1.7)	Brazil	24 (1.0)
17	U Virginia	30 (1.2)	UCLA	42 (1.7)	Japan	22 (0.9)
18	Penn State	29 (1.2)	Cornell	40 (1.6)	Iran	17 (0.7)
19	FED Board	28 (1.2)	Duke	37 (1.5)	Chile	14 (0.6)
20	Michigan State	28 (1.2)	Brown	34 (1.4)	Netherlands	14 (0.6)

Notes: List of the top 20 Departments for first job after the PhD, the PhD granting Universities, and the top 20 Nationalities.

Table 14: Additional Summary Statistics.

	Female		Male	
Rank	Name	h-Index	Name	h-Index
1	Esther Duflo	89	Joseph Stiglitz	221
2	Janet Currie	86	James Heckman	170
3	Francine Blau	63	Jeffrey Sachs	159
4	Caroline Hoxby	62	Andrei Shleifer	154
5	Rachel Croson	60	Daron Acemoglu	152
6	Nora Lustig	57	Robert Barro	128
7	Maureen Cropper	56	Martin Ravallion	126
8	Amy Finkelstein	53	Richard B. Freeman	122
9	Catherine Eckel	53	Colin Camerer	122
10	Serena Ng	51	John B. Taylor	121

Notes: List of the top 10 Economists ranked by the h-Index.

Table 15: Auxiliary Regression Results.

	Co-Authors		Switches	
Variable	Female	Male	Female	Male
fWHR	-13.16**	18.97**	-0.96*	0.36
	(6.33)	(8.93)	(0.53)	(0.31)
White	-0.82	0.40	0.19*	0.05
	(1.32)	(1.01)	(0.10)	(0.06)
Time since PhD	0.70***	0.87***	0.04***	0.03***
	(0.11)	(0.08)	(0.006)	(0.003)
Editor	5.02***	8.63***	0.20*	0.34***
	(1.44)	(1.20)	(0.12)	(0.06)
Switches	1.88**	1.03		
	(0.84)	(0.78)		
Theory	-3.37***	-7.51***	-0.001	0.18***
	(1.13)	(1.34)	(0.13)	(0.06)
Field				
Micro	-0.70	-2.40	0.12	-0.05
	(2.03)	(1.51)	(0.24)	(0.11)
Macro	-5.21**	-6.97***	0.27	-0.01
	(2.11)	(1.55)	(0.24)	(0.11)
Obs.	401	1,578	401	1,578
R^2	0.41	0.28	0.24	0.20

Notes: Results for Co-Authors are robust to including ethnicity and beauty. The results for switches are robust to including ethnicity and having a first Job at a Top 10 Department. Robust standard errors in parenthesis. Constant not shown. Significance levels: ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$.

I hereby declare that this dissertation 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. No part of this dissertation was previously presented for another degree at this or any other institution.