

On the Importance of Social Status for Occupational Sorting*

Paweł Gola[†]

May 2022

Abstract

Models of self-selection predict that occupations with flat wage schedules attract workers of low average skill. Yet, in academia wages are flat but the average skill level is high. In this paper, I examine whether social status concerns can explain this puzzle. I find that within-occupation status can ensure that academia attracts mainly high-skilled workers, but only at the cost of attracting few workers overall. If, however, workers care both about within- and between-occupation status, then academia can be arbitrarily large and attract workers of high average skill. I conclude that within- and between-occupation status concerns act as complements.

JEL Codes: D91, J24.

Keywords: occupational sorting, self-selection, social status, occupational prestige, relative concerns.

*I would like to thank Charles Brendon, Peter Eso, Dawei Fang, Sanjeev Goyal, Claudia Herresthal, Ed Hopkins, Ian Jewitt, Coen Teulings, Adrien Vigier, and Yuejun Zhao, as well as seminar participants in Cambridge, Gotheborg, Helsinki, Huddersfield, Oslo, Oxford, Trondheim, Warsaw, and Xiamen, for their comments and suggestions. All errors are mine.

[†]University of Edinburgh and BI Norwegian Business School, 30 Buccleuch Place, Edinburgh, EH8 9JT.

1 Introduction

There is abundant evidence that people care about their social status (Huberman, Loch, and Onculer, 2004; Bursztyn, Ferman, Fiorin, Kanz, and Rao, 2018) and their relative position among their peers (Luttmer, 2005; Card, Mas, Moretti, and Saez, 2012; Perez-Truglia, 2020). In particular, people are willing to forgo substantial pecuniary benefits (Cardoso, 2012; Bottan and Perez-Truglia, 2020) and accept high risks (Ager, Bursztyn, Leucht, and Voth, 2021) in exchange for higher status and/or rank. These findings strongly suggest that the desire for higher status should also affect workers' occupational choices. How profoundly, however, can this desire for status affect sorting patterns in equilibrium?

To make this question more precise, consider sorting into academia. Compared to other professional occupations, the wage schedule in academia is relatively flat: For example, wages of academic economists have been found to be less differentiated than wages of economists in the private sector (Machin and Oswald, 2000); more generally, overall wage dispersion in academia is much lower than in, for example, finance.¹ The literature on self-selection famously teaches us that occupations with flat wage schedules are less likely to attract workers of high skill than occupations with steep schedules (Roy, 1951; Heckman and Sedlacek, 1985; Borjas, 1987; Heckman and Honore, 1990). And yet, in an apparent refutation of Roy's model, even though academia pays relatively flat wages, it attracts workers of very high skill on average: The average cognitive score among academic economists and political scientists is higher than among parliamentarians, CEOs, and lawyers and judges (Bó, Finan, Folke, Persson, and Rickne, 2017, Table II) and academically-oriented research jobs (i.e., once that incentives publication of results) attract better researchers than commercially-focused research jobs (Stern, 2004).²

¹In Norway the 90/10 wage ratio is between 1.25 and 1.35 (depending on specification) times greater in finance than in academia, whereas the standard deviation of wages is between 1.8 and 2.25 times as large in finance as in academia. These numbers are based on my own investigation for Norway, conducted with invaluable help from Yuejun Zhao: The details are provided in Online Appendix D.

²Specifically, Stern (2004) finds that the compensating differential paid by more academic jobs in research becomes negative only after controlling for ability, which implies positive selection.

At the same time, rank plays a very important role in academia and the information about anyone’s rank is extremely easily accessible, to a degree that is comparable only with professional sports: The vast majority of academics make their entire publication and citation records publicly available, and there exist specialized websites that rank academics—globally, within their countries and within their departments.³ In general, easier access of information about relative standings translates into greater differences in happiness between high- and low-rank individuals (Perez-Truglia, 2020), which raises the possibility that it is precisely academia’s obsession with rank that provides the additional differentiation of rewards necessary for attracting high quality workers.

In this paper, I examine the circumstances under which within- and between-occupation relative concerns can explain why academia is able to attract high-skilled workers. My model builds on Roy (1951), which is the standard model in the literature on occupational sorting. There is a continuum of workers, who freely join one of the two occupations: finance or academia. Each worker is endowed with some level of financial skill and some level of academic skill. The wage in finance is an increasing function of the financial skill, whereas academia pays all workers the same, flat wage. As a result, in the no-status benchmark (i.e., if wages were the workers’ only reward) all workers with high enough financial skill join finance, with academia attracting only workers with low financial skill. Finally, I assume that the academic and financial skills are positively interdependent, which implies that in the benchmark academics have low academic skill as well.

I depart from Roy’s setting by assuming that, apart from wages, workers also care about their social status, which consists of two components: *occupational prestige* (i.e., between-occupation relative concerns) and *local status* (i.e., within-occupation relative concerns), each determined endogenously. First, I consider the impact of occupational prestige only. Following the economics literature on social status (see, for example, Weiss and Fershtman (1992); Fershtman and Weiss (1993); Mani and Mullin (2004)), I

³For example, for physics there exists <http://rtorre.web.cern.ch/rtorre/PhysRank/index.html>. In economics, rankings are compiled and regularly updated on <https://ideas.repec.org/top/>. There exist also countless articles which provide rankings of academics (e.g. Ioannidis, Boyack, and Baas, 2020) and academic departments (e.g. Amir and Knauff, 2008).

assume that occupational prestige is determined by the average skill in the two professions and thus enters the reward function in each occupation as an endogenous constant.⁴ Because of that, prestige does not make the reward schedule in academia any steeper, and thus cannot, on its own, improve selection into academia. On top of that, if workers care only about wages and prestige, then academia attracts fewer workers than in the benchmark: The steep wage schedule in finance makes finance the prestigious occupation, which then, in turn, attracts a larger number of workers than the wage level itself would warrant.⁵

Next, I consider the impact of local status only, which is modeled as a linear function of the worker's rank within her chosen profession. For example, an academic's local status depends on how her academic skill compares to that of other academics. The linear function is chosen so that the average local status in a profession is always equal to 0: If local status becomes more important in an occupation, then the top workers are rewarded more but the lowest-ranked workers are rewarded less. Finally, local status is allowed to enter workers' rewards with a different weight in each occupation: In occupations with more rigidly defined and more precisely observable notions of achievement, rank is more salient and thus influences workers' well-being strongly; in occupations where workers have very little idea about their own or anyone else's rank, local status has little scope to operate.

I find that while local status rewards can indeed overcome the impact of flat wage schedules in academia on selection—that is, can ensure positive selection into academia—they also introduce a trade-off between the number and the quality of workers attracted by academia. The more local status matters within academia, the higher is the punishment inflicted on the lowest-ranked academics, *regardless of their skill*. If the difference in the weight put on local status in academia and finance exceeds the difference between the academic wage and the lowest wage in finance, then no agent is willing to be the lowest-ranked academic and academia unravels (attracts

⁴Specifically, I assume that occupational prestige depends positively on the difference between the average academic skill in academia and the average financial skill in finance.

⁵This result is similar to the argument initially outlined in Chapter X of Adam Smith's *The Wealth of Nations* (Smith, 1776) and later formalized in Weiss and Fershtman (1992) and Fershtman and Weiss (1993).

a zero measure of workers). Accordingly, if local status matters sufficiently more in academia than in finance, then academia is the smaller occupation. At the same time, if local status matters similarly across occupations, but is much more important than wages, then an occupation can be large only if it attracts the workers who are bad at both jobs; naturally then, the smaller occupation attracts workers of higher skill on average. Overall, therefore, if workers from both occupations care sufficiently strongly about local status, and yet its importance is much greater in academia than in finance, then academia will be the smaller occupation, and will thus attract workers of higher skill (on average) than finance.

A similar reasoning implies also that if academic wages are low, then academia can attract workers of higher skill than finance *only if* it is the smaller occupation. If academia is large, it can attract workers of higher skill than finance only if the local status rewards are sufficiently more important in academia than they are in finance. However, if the academic wage is low, then even a slightly greater weight put on local status in academia than in finance will cause academia to be very small.

Finally, I examine what happens if workers care both about local status and occupational prestige. Strikingly, the trade-off between the size and quality of workers joining academia disappears in that case, suggesting that local status and occupational prestige act as complements. Specifically, if local status is sufficiently important in academia compared to finance and workers' taste for occupational prestige is sufficiently strong, then academia can attract an arbitrarily large number of workers while maintaining a higher average quality of workforce than finance. The intuition is novel: Suppose for now that the government strives to maintain a fixed size of the academic sector and achieves this goal by adjusting the academic wage. In such a case, if local status becomes more important in academia, then academia attracts workers who are more skilled on average, which increases academia's prestige. The greater the taste for prestige, the more this higher prestige means to the lowest-ranked academic, and thus the lower the wage level needed to maintain the desired size of academia. Returning to the case where the academic wage is constant and the size of academia varies, if the taste for prestige is arbitrarily high, then—regardless of how low academic wages are—local status can be much more important in academia than in

finance, without having an adverse effect on academia’s size; and high local status rewards for skilled workers allow academia to attract talent.

The rest of the paper is structured as follows. Section 2 discusses the related literature. Section 3 develops the model and motivates my modeling choices. Section 4 derives the main results. Section 5 discusses the policy implications of my results, provides examples of other occupations in which the results may be relevant, and discusses the appeal (or lack thereof) of two alternative explanations of the motivating puzzle. Appendix A contains the proofs of all propositions and lemmas. Online Appendix B discusses why none of the simplifying assumptions are critical and Online Appendix D compares the empirical distributions of wages in academia and finance.

2 Related Literature

There is only a handful of papers addressing the impact that social status has on sorting into occupations, most of them written by Chaim Fershtman and Yoram Weiss. In the model examined in Weiss and Fershtman (1992) and Fershtman and Weiss (1993), workers are *ex ante* homogenous in skill but can choose how much education to acquire: The prestige of each occupation depends on the average wage and average educational level in that occupation. Fershtman, Murphy, and Weiss (1996) embed an extension of that model into an endogenous growth model and show that the preference for social status may crowd out high-ability/low-wealth workers from the growth-enhancing occupation.

Mani and Mullin (2004) develop a Roy’s model with log-normally distributed skills, in which workers care only about social status.⁶ Social status is a weighted sum of the absolute, rather than relative, level of the worker’s occupation-specific skill and the occupational prestige (modeled as occupation-specific average skill). Crucially, the weight with which the absolute level of skill matters is equal to the proportion of workers who joined that occupation. Therefore, larger occupations tend to have steeper

⁶Albornoz, Cabrales, and Hauk (2020) is also relevant, even if it is not explicitly concerned with social status. The authors develop a Roy’s model with independently distributed skills, endogenous choice of effort, and productivity spillovers within occupations that act similarly to occupational prestige.

reward schedules, which provides an alternative mechanism through which social status can cause occupations with flat wage schedules to be both larger and attract higher quality talent in equilibrium. However, as there are no within-occupation relative concerns in [Mani and Mullin \(2004\)](#), the insights that local status introduces a trade-off between size and quality, and that local status and occupational prestige act as complements, are absent.

To the best of my knowledge, the only other article that examines the impact of local status on occupational sorting is the companion of this paper ([Gola, 2015](#)), which introduces both components of social status into the two-sector assignment model from [Gola \(2021\)](#) and derives the distributional consequences of an increase in the importance of local status. However, in that paper the number of jobs in each sector is fixed, which means that the occupational prestige component of the reward is competed away and has no impact on sorting.

Robert Frank has examined (in [Frank \(1984\)](#) and [Frank \(1986\)](#)) how local status affects workers' sorting into firms. However, the impact of social status on sorting across firms is fundamentally different from its impact on occupational sorting. A firm takes into account the effect of its hiring decisions on the well-being of its other employees, and thus internalizes the externalities produced by within-firm local status. An occupation consists of workers employed by many independent firms, none of which considers the effect of its hiring decisions on everyone else in that profession. Thus within-firm local status influences mostly internal wage structures, whereas within-occupation local status affects mostly occupational sorting and only indirectly wage structures.

There are a number of papers which allow for the presence of within- and between-group relative concerns but examine sorting across entities other than occupations or firms. Among these, [Damiano, Li, and Suen \(2010, 2012\)](#) are particularly relevant.⁷ In those papers workers choose between

⁷The papers by [de Bartolome \(1990\)](#), [Becker and Murphy \(2000\)](#), and [Morgan, Sisak, and Várdy \(2018\)](#) are also related, but less so. [de Bartolome \(1990\)](#) and [Becker and Murphy \(2000\)](#) consider the impact of between-group relative concerns on residential sorting in models with binary ability. [Morgan et al. \(2018\)](#) examine sorting into contests, in a setting where the success in each contest depends only on one's relative position among the participants.

two organizations, and their only concerns are their own rank and the average skill within their chosen organization. The complementarity between within- and between-group relative concerns is not explicitly pointed out by the authors, but it is present: For example, in [Damiano et al. \(2012\)](#) the authors show that if between-group relative concerns become more important, then the two organizations design steeper within-group reward schedules. The critical difference between these models and the present paper is that [Damiano et al. \(2010, 2012\)](#) assume that the two organizations have a fixed capacity, in order to “circumvent the issue of size effect” ([Damiano et al. \(2012\)](#), pp 2213). Conversely, the size effect is critical for my work, as my focus is on occupations rather than organizations. Accordingly, the insights about the trade-off between size and quality created by within-group relative concerns are absent in [Damiano et al. \(2010, 2012\)](#), as is the insight that between-group relative concerns alleviate said trade-off.

There is a small literature concerned with the role that fame plays in steepening the reward schedule in academia. Many authors (e.g., [Merton, 1973](#); [Dasgupta and David, 1987, 1994](#); [Stephan, 1996](#)) have discussed informally the crucial role played by research priority in motivating researchers: Being the first person to make a scientific discovery brings fame and respect, which creates incentives to exert effort and presumably attracts talented workers to academia. This reasoning is formalized by [Jeon and Menicucci \(2008\)](#); in their model the quality of the peer-review process determines whether fame accrues to the authors of actual scientific achievements: If this is the case and workers care about fame sufficiently strongly, then academia is able to attract superior talent. In that model, one receives the same fame reward whether there are many or just a few discoveries being made; thus there is no trade-off between the quality and the size of the workforce in academia. More recently, [Hill and Stein \(2021\)](#) make the intriguing point that the desire for research priority incentivizes researchers to put less care into their research; however, they abstract from the question on how research priority affects selection into academia.

3 The Model

There is a unit measure of workers, and there are two occupations: academia and finance. Each agent is fully described by her skill vector $(x_A, x_F) \in [a_x, b_x]^2$, where x_A and x_F are the skills used in academia and finance, respectively. The distribution H is symmetric, twice continuously differentiable and has a strictly positive, finite density in its support. Symmetry implies, among other, that both skills have the same marginal distribution; its cdf is denoted as H_M and its pdf as h_M . Finally, I assume that $H(x_A, x_F) > H_M(x_A)H_M(x_F)$ for all $(x_A, x_F) \in (a_x, b_x)^2$, which means that the two skills are positively interdependent.

Each worker joins the occupation which maximizes her reward: A worker's reward consists of her occupation-specific wage, the prestige of the occupation she joins and her position within that occupation (local status); the last two components are endogenous (i.e., dependent on workers' sorting). Entry into each occupation is free: Workers who join academia will be called academics, and workers who join finance will be called bankers. The reward function is specified in detail below; for now, it suffices to know that, keeping constant the occupational choices of other workers, the total reward received by a worker in occupation i is increasing in the occupation-specific skill x_i and does not depend on the other skill. This property implies that for any $x''_A \geq x'_A$, $x''_F < x'_F$ if worker (x'_A, x'_F) joins academia, then worker (x''_A, x''_F) must also prefer to join academia. For that reason, and without any loss in generality, I will restrict attention to such *sorting* of workers to occupations which can be characterized by the means of some increasing separation function $\psi : [a_x, b_x] \rightarrow [a_x, b_x]$, such that a worker (x_A, x_F) joins academia if $x_F < \psi(x_A)$ and joins finance if $x_F > \psi(x_A)$.⁸ A sorting will be called non-degenerate if the set $D = \text{cl}(\{x_A \in [a_x, b_x] : \psi(x_A) \in (a_x, b_x)\})$ has a strictly positive measure (where cl denotes the closure of a set). Finally, I will denote $\min D$, $\psi(\min D)$, $\max D$, and $\psi(\max D)$ by x_A^m , x_F^m , x_A^s and x_F^s , respectively. Figure 1 depicts how a separation function determines the sorting of workers into occupations.

I will now introduce the three components of rewards, and then define

⁸Workers for whom $x_F = \psi(x_A)$ are of measure zero, and can thus be ignored without loss of generality.

the total reward function and the equilibrium. Finally, in Section 3.1, I will motivate my modeling choices.

Occupational Prestige Occupational prestige can be thought of as the component of social status which is common to all members of a given profession. Following the literature, the prestige of a profession depends on the occupational average of skill (Fershtman et al., 1996; Mani and Mullin, 2004). Specifically, in any non-degenerate sorting the occupational prestige of a profession is proportional to the difference between the averages of the occupation-specific skills in the two professions, with

$$o_A = \frac{1}{M_A} (\bar{x}_A^A - \bar{x}_F^F) \quad \text{and} \quad o_F = \frac{1}{M_F} (\bar{x}_F^F - \bar{x}_A^A), \quad (1)$$

where \bar{x}_A^A is the average academic skill among academics, \bar{x}_F^F is the average financial skill among bankers and M_i denotes the measure of workers who joined occupation i . In other words, academia is the prestigious occupation if the average academic is better at research than the average banker is at finance.

Local Status Local status depends on the agent's rank in the occupation-specific skill among other members of her profession. Specifically, the local status of agent (x_A, x_F) who joins occupation i is

$$s_i(x_i) = 2G_i(x_i) - 1, \quad (2)$$

where $G_i(\cdot)$ denotes the distribution of the skill specific to occupation i among the workers who have joined occupation i .

Wages An agent (x_A, x_F) earns wage $w_F(x_F)$ if she joined finance and a flat wage $w_A(x_A) = w_A \in (w_F(a_x), w_F(b_x))$ if she joined academia. The wage function in finance is twice continuously differentiable with a strictly positive first derivative $w'_F > 0$.

Rewards and (Compensated) Equilibrium Given a sorting ψ , the *reward* of an agent (x_A, x_F) from joining occupation $i \in \{A, F\}$ is a weighted

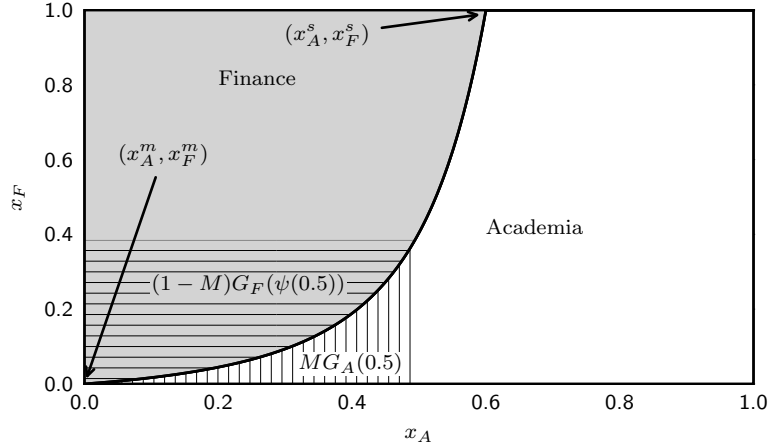


Figure 1: The Separation Function and Sorting

Notes: The solid black curve depicts the separation function ψ . The white (light gray) area below (above) ψ depicts the space of workers that join academia (finance). The vertically (horizontally) hatched area represents the space of workers with skill $x_A \leq 0.5$ ($x_F \leq \psi(0.5)$) who join academia (finance); their number depends on the number of workers who reside in this space, with $M_A G_A(x_A) + M_F G_F(\psi(x_A)) = H(x_A, \psi(x_A))$.

sum of her wage, the prestige of occupation i , and her local status within it:

$$t_i(x_i; \psi) = w_i(x_i) + l_i s_i(x_i; \psi) + k o_i(\psi), \quad (3)$$

where $l_i \geq 0$ is the importance of local status rewards in occupation i and k is the population-wide taste for prestige. In my analysis, I will be interested either in symmetric changes to l_A and l_F or in changes to l_A only. For that reason, it will be convenient to rewrite l_A as the sum of the *overall importance of local status* (relative to wages) l_F and the *importance of local status in academia* (relative to finance) $\delta \equiv l_A - l_F$.

Before I define what constitutes an equilibrium in this model, let me first introduce the more general concept of a *compensated equilibrium*.

Definition 1. A sorting ψ^c constitutes a *compensated equilibrium* if and only if (a) ψ^c is non-degenerate and (b) there exists some *compensating differential* $c \in \mathbf{R}$ such that for all $(x_A, x_F) \in [a_x, b_x]^2$

$$\begin{aligned} \psi^c(x_A) > x_F &\Rightarrow t_A(x_A; \psi^c) + c > t_F(x_F; \psi^c), \\ \psi^c(x_A) < x_F &\Rightarrow t_A(x_A; \psi^c) + c < t_F(x_F; \psi^c). \end{aligned} \quad (4)$$

Condition (4) stipulates that the economy is in a compensated equilibrium if there exists a compensating differential which, after adding it to the academic wage, would ensure that each academic receives at least as high a reward in academia as the reward she would receive in finance (and *vice versa*).

Compensated equilibria are closely related to the equilibria of this model. In an *equilibrium*, workers join the occupation that maximizes their reward, taking the sorting decisions of all other workers as given. Thus a non-degenerate sorting ψ^e constitutes an equilibrium if and only if it constitutes a compensated equilibrium for $c = 0$. In the case of degenerate equilibria, I will adopt the following convention: The sorting $\psi(x_A) = a_x(\psi(x_A) = b_x)$ constitutes an equilibrium if there exists an $\epsilon > 0$ such that for all $m \leq \epsilon$ there exists a compensated equilibrium ψ^c such that $M_A(\psi^c) = m(M_F(\psi^c) = m)$ and $c \geq 0(c \leq 0)$. In plain English, academia attracts no workers in equilibrium if and only if we would need to increase the academic wage in order to attract a small number of workers into academia.⁹

Furthermore, the taste for prestige k determines only which compensated equilibria constitute an equilibrium, but it leaves the set of compensated equilibria unaffected. That is, if a sorting ψ constitutes a compensated equilibrium for some $k' \geq 0$, then it constitutes a compensated equilibrium for all $k \geq 0$. The reason is that occupational prestige enters rewards as a constant, and thus as an endogenous compensating differential.

3.1 Discussion

In this section, I briefly discuss my modeling choices. A more detailed discussion is provided in Online Appendix B.

Occupational Prestige Two of my modeling choices regarding occupational prestige may seem somewhat *ad hoc*: (a) that occupational prestige

⁹Alternatively, one could adapt the Divinity Criterion to the current model; however, as the workers who would join academia in compensated equilibria with very small M_A are clearly those that are most likely to join academia when $M_A = 0$, these two definitions should yield the same results.

is inversely proportional to the size of the occupation and (b) that the taste for occupational prestige is the same in the two occupations. These two assumptions jointly normalize the sum of occupational prestige rewards to zero, that is, they ensure that $M_{AOA} + M_{FOF} = 0$.¹⁰ This implies that (a) any change in sorting leaves the sum of occupational prestige rewards unchanged and (b) that changes in the taste for prestige affect welfare only indirectly, through their impact on sorting. Crucially, this normalization leaves all formal results unaffected (see Online Appendix B.1). In fact, even the assumption that occupational prestige depends on the difference between the academic skill of academics and the financial skill of bankers is *not critical* for the results. Online Appendix B.2 explores a range of alternative assumptions, all of which result in the same message.

Finally, one could wonder whether occupational prestige should not be modeled as backward looking: Is it not plausible to think that present-day academics are attracted by the accomplishments of past greats, like Einstein or Skłodowska-Curie? However, the fact that occupational prestige depends on average skill in my *static model* is perfectly consistent with the fact that occupational prestige may depend on past achievements in a *dynamic model*, because average skill will be unchanged over time in a steady state of a dynamic model (see, for example, Mani and Mullin (2004) or Online Appendix OA.5.4. in Gola (2021)).

Local Status Local status is usually defined as the esteem one receives from one’s reference group (Frank, 1984). In this model, occupation is the only possible reference group, and esteem is modeled as one’s rank: As rank is always standard uniformly distributed, it follows that local status is $U[-1, 1]$ distributed within each occupation and thus of zero-sum. The assumption that the within-occupation ranking is based on the occupation-specific skill is natural, as the esteem received from peers is likely to be strongly related to how well the agent performs her job. A common al-

¹⁰They can also be straightforwardly micro-founded. Suppose that workers receive a utility from their occupational prestige whenever they meet someone from the other profession, in which case a worker from occupation i receives a utility boost (or decrease) of $(\bar{x}_i - \bar{x}_j)k$. Clearly, the sum of the two workers rewards is 0 on a meeting level; however, workers from the smaller occupation will participate in a larger number of between-occupation meetings, and thus occupational prestige enters their reward function with weight k/M_i .

ternative is to assume that the ranking depends on income (Hopkins and Kornienko, 2004). This is equivalent to my assumption if $w'(x_A) > 0$; the main results of this paper are robust to setting $w'(x_A)$ to be strictly positive (but small).

Unlike the taste for prestige, the taste for local status is allowed to be occupation dependent. This assumption is plausible, because the extent to which people care about local status depends on the intensity of social interactions among peers (Ager et al., 2021)—that is, on how socially hermetic the profession is—and how easily observable ranks are (Perez-Truglia, 2020)—that is, it depends on the precision and availability of information about ranks—both of which differ across occupations.¹¹ Nevertheless, it is worth stressing that the condition $l_A > l_F$ is necessary for academia to be both larger and to attract workers of (on average) higher skill than finance (as in Theorem 3). If the importance of local status was symmetric across sectors, then academia could attract workers of (on average) higher skill only if it was the smaller occupation.

Other Assumptions The remaining assumptions are all made to ease exposition, and are not critical. In particular, in the Online Appendix I discuss why the main message of the article remains unchanged if we allow for (a) non-constant (but still flat!) wage schedules in academia (OA B.3) and (b) endogenous wage functions in both occupations (OA B.4). I also explain that if skills are strongly negatively interdependent, then the smaller occupation always attracts better workers, regardless of how steep the reward schedules are (OA B.5).

4 Impact of Social Status on Sorting

4.1 The No-Status Case

Let us first consider, as a benchmark, what happens if there are no social status rewards, that is, if $t_i(x_i) = w_i(x_i)$. The equilibrium is trivial: The separation function is constant, that is, there exists a single cutoff value

¹¹ In a companion paper (Gola, 2015), I provide a microfoundation of the social status reward function in which the weight with which local status enters the utility function depends precisely on these two facts.

$\psi(x_A) = \psi^b = w_F^{-1}(w_A)$ such that all workers with $x_F > \psi^b$ join finance and all workers with $x_F < \psi^b$ join academia. This is because rewards are constant in academia but differ with skill in finance, and thus any agent who would earn less than the academic wage w_A in finance joins academia, and everyone else becomes a banker. As only workers with $x_F \leq \psi^b$ join academia and because $H(x_A, x_F) > H_M(x_A)H_M(x_F)$, it follows that $G_A(x_A) = H(x_A, \psi^b)/H_M(\psi^b) > H_M(x_A)$. This implies that \bar{x}_A^A is smaller than the population-wide average skill, $\bar{x} = \int_{a_x}^{b_x} x h_M(x) dx$. As finance attracts only workers with $x_F > \psi^b$, it must be that $\bar{x} < \bar{x}_F^F$, which means that, on average, academia attracts less skilled workers than finance does.

4.2 Prestige-Only Equilibrium

To see the effect of occupational prestige on sorting, let us find the equilibrium in the case where workers care only about wages and prestige, but not about local status; that is, in the case where $t_i(x_i) = w_i(x_i) + k o_i(\psi)$.

Theorem 1. If $l_A = l_F = 0$, then $M_A(\psi^e) < H_M(\psi^b)$ in all equilibria, and the set of equilibria is non-empty. Furthermore, $\bar{x}_A^A < \bar{x}_F^F$ in any non-degenerate equilibrium.

Proof. Because $k o_i$ does not depend on the worker's type, any compensated equilibrium must be of the form $\psi^c(x_A) = \psi^p$. Again, because $H(x_A, x_F) > H_A(x_A)H_F(x_F)$ we have $\bar{x}_A^A < \bar{x} < \bar{x}_F^F$ in any compensated equilibrium and hence also in any non-degenerate equilibrium. From this follows immediately that $c = w_F(\psi^p) - w_A + k \frac{\bar{x}_F^F - \bar{x}_A^A}{H_M(\psi^p)H_M(1-\psi^p)}$, so that $c > 0$ if $\psi^p \geq \psi^b$, which implies that $M_A(\psi^e) < H_M(\psi^b)$ in all equilibria. Finally, by continuity of c with respect to ψ^p , a non-degenerate equilibrium will not exist only if $c > 0$ for all ψ^p ; but this implies the existence of a degenerate equilibrium in which $\psi^p = 0$. \square

As in the no-status equilibrium, there exists a cutoff value of x_F that fully determines sorting. Because all academics benefit from prestige in equal measure, rewards are still constant in academia but differentiated in finance. Thus all workers with high financial skill join finance, making it necessarily more prestigious than academia. This in turn implies that the introduction of taste for prestige makes academia even less rewarding than

before, decreasing its size. With constant rewards in academia, prestigious academia simply cannot be sustained: High prestige would predominantly lure in workers of low financial skill, making academia less prestigious than finance.

4.3 Local Status Equilibrium

In this section, I consider what happens if workers care about local status but not about occupational prestige, in which case rewards are given by $t_A(x_A) = w_A + (l_F + \delta)(2G_A(x_A) - 1)$ and $t_F(x_F) = w_F(x_F) + l_F(2G_F(x_F) - 1)$. First, in Section 4.3.1 I characterize the unique equilibrium. Then, in Section 4.3.2, I focus on the compensated equilibrium and (a) prove that it is unique for a *given* size of academia and (b) examine how it depends on δ and l_F . Finally, in Section 4.3.3, I establish how much of an impact local status concerns have on occupational sorting in equilibrium.

4.3.1 Characterizing the Equilibrium

Suppose that $w_A \in (w_F(a_x) + \delta, w_F(b_x) + \delta)$.¹² Because the reward functions $t_F(\cdot), t_A(\cdot)$ are continuous in skill, it follows from Condition (4) and the definition of an equilibrium that a sorting ψ^e constitutes an equilibrium if and only if, for all $x_A \in D$,

$$t_F(\psi^e(x_A)) = t_A(x_A). \quad (5)$$

Recall that x_A^m denotes $\min D$ and x_F^m denotes $\psi^e(x_A^m)$. As x_i^m is simply the skill of the lowest ranked worker in occupation i , we have $G_i(x_i^m) = 0$. Therefore, $x_F^m = w_F^{-1}(w_A - \delta) > a_x$ by Equation (5). However, if the skill of the lowest ranked banker greater than a_x , then the skill of the lowest ranked academic must be equal to a_x , so that $x_A^m = a_x$.

Differentiating Equation (5) reveals that for all $x_A \in D$, ψ^e must satisfy

$$2(l_F + \delta)g_A(x_A) = \psi^{e'}(x_A) [w_F'(\psi^e(x_A)) + 2l_F g_F(\psi^e(x_A))], \quad (6)$$

¹²Otherwise the equilibrium is degenerate; see Proposition 2 and the discussion that follows it.

where $g_i(x_i)$ denotes the density of skill x_i in occupation i , with

$$g_A(x_A) = \frac{h_M(x_A)\Pr(X_F < \psi^e(x_A)|X_A = x_A)}{M_A} = \frac{\frac{\partial}{\partial x_A}H(x_A, \psi^e(x_A))}{M_A},$$

$$g_F(\psi^e(x_A)) = \frac{h_M(\psi^e(x_A))\Pr(X_A < x_A|X_F = \psi^e(x_A))}{M_F} = \frac{\frac{\partial}{\partial x_F}H(x_A, \psi^e(x_A))}{M_F}.$$

Equation (6) can be thus rewritten as the following initial value problem (IVP):

$$\psi^{e'}(x_A) = F(x_A, \psi^e(x_A)) \quad \text{and} \quad \psi^e(a_x) = w_F^{-1}(w_A - \delta) \quad (7)$$

where

$$F(x_A, x_F) \equiv \frac{1 - M_A}{M_A} \frac{(l_F + \delta) \frac{\partial}{\partial x_A}H(x_A, x_F)}{0.5(1 - M_A)w'_F(x_F) + l_F \frac{\partial}{\partial x_F}H(x_A, x_F)}.$$

However, the size of academia, M_A , clearly depends on ψ^e as well, and thus ψ^e must also satisfy:

$$M_A(\psi^e) \equiv \int_{a_x}^{b_x} \frac{\partial}{\partial x_A}H(r, \psi^e(r))dr. \quad (8)$$

Overall, to find the equilibrium, we need to solve the IVP defined by Equation (7) for each M_A , and then solve for M_A using Equation (8).

Proposition 1. Suppose $k = 0$. (i) If $w_A \in (w_F(a_x) + \delta, w_F(b_x) + \delta)$, then there exists a unique non-degenerate equilibrium and the size of academia increases with w_A in equilibrium. (ii) If $w_A \notin (w_F(a_x) + \delta, w_F(b_x) + \delta)$ then there exists no non-degenerate equilibrium.

In the local-status-only case, the equilibrium is unique. Naturally, the size of academia increases with the wage in academia, as higher pay attracts more workers. However, both of these results may break down if $k > 0$, because occupational prestige can be non-monotonic in academia's size. In that case, an increase in the size of academia can itself provide the increase in reward which is needed to sustain an equilibrium in which academia is larger. Unfortunately, this means that multiplicity of equilibria will be a concern in the general case.

4.3.2 Compensated Equilibria

My aim is to establish the extent to which local status can influence sorting. Much of that goal can be accomplished by focusing on the compensated equilibria of this model, which are easier to study than the equilibrium itself: If certain selection patterns cannot be sustained in any compensated equilibrium, then they cannot hold in equilibrium either.

Lemma 1. For every $M_A \in (0, 1)$ there exists a unique compensated equilibrium in which academia is of size M_A ; this compensated equilibrium will be denoted by $\psi^c(\cdot; M_A)$. The compensated equilibrium $\psi^c(\cdot; M_A)$ is continuous in M_A .

For every non-degenerate size of academia $M_A \in (0, 1)$, there exists a unique compensated equilibrium. This is consistent with the interpretation of c as a compensating differential: Academics need to be paid this much more to ensure that M_A academic jobs will be filled. This property forms the cornerstone of my analysis, because it allows me to study how the compensating differential and the distribution of skill in each occupation change with taste parameters for a given M_A .

Lemma 2. If a change in the taste parameters (l_F, δ) or the wage function w_F causes a strict increase in $F(x_A, x_F)$ for all $(x_A, x_F) \in (a_x, b_x]^2$, then it also causes an increase in $G_F(x_F)$ for all $x_F \in [a_x, b_x]$ (and strictly for some) and a decrease in $G_A(x_A)$ for all $x_A \in [a_x, b_x]$ (and strictly for some), in any $\psi^c(\cdot; M_A)$.

The function $F(x_A, x_F)$ captures the extent to which rewards differ with skill in academia *relative* to the extent to which rewards differ with skill in finance. If rewards become steeper in academia (relative to finance), then the distribution of skill improves in academia and worsens in finance, both in the sense of first-order stochastic dominance. Intuitively, more differentiation in rewards increases the rewards of high-skilled workers and punishes low-skilled workers; the size of academia is kept constant by adjustments to the compensating differential.

Lemma 2 can be used to examine the impact of both an increase in the importance of local status in academia relative to finance (an increase in δ

that keeps l_F constant) and an increase in the overall importance of local status (an increase in l_F that keeps δ constant). In particular, we have that

$$\begin{aligned} \frac{\partial}{\partial \delta} F(x_A, x_F) &> 0, \\ \frac{\partial}{\partial l_F} F(x_A, x_F) &> 0 \iff \delta \leq 0.5 \frac{(1-M)w'_F(x_F)}{\frac{\partial}{\partial x_F} H(x_A, x_F)}. \end{aligned} \tag{9}$$

Thus an increase in δ always improves the distribution of skill in academia, whereas an increase in l_F improves the distribution of skill in academia as long as the importance of local status in academia relative to finance is not too high. For instance, if local status rewards are symmetric across occupations ($\delta = 0$), then an increase in the overall local status intensity improves the distribution of skill in academia.

Because we are interested in how strongly social status can affect occupational sorting, it is going to be useful to understand what happens in each compensated equilibrium in the limit, as local status becomes infinitely more important than wages.

Lemma 3. Fix $M_A \in (0, 1)$ and $\delta \in \mathbf{R}$, and consider the limit of $\psi^c(\cdot; M_A)$ as $l_F \rightarrow \infty$. (i) If $M_A \geq (\leq) 0.5$, then $\lim_{l_F \rightarrow \infty} G_A(x) \geq (\leq) \lim_{l_F \rightarrow \infty} G_F(x)$ for all $x \in [a_x, b_x]$. Accordingly, (ii) for any $M_A \in (0, 0.5)$ and $\delta \in \mathbf{R}$ there exists some $l_F^* > 0$ such that if $l_F \geq l_F^*$ then $\bar{x}_A^A - \bar{x}_F^F > 0$ in $\psi^c(\cdot; M_A)$.

Lemma 3 states that—keeping academia’s size and the importance of local status in academia relative to finance constant—if local status becomes infinitely more important than wages in each occupation, then the distribution of the academic skill among academics dominates the distribution of the financial skill among bankers if and only if academia is the smaller occupation ($M_A \leq 0.5$). To understand the intuition behind this result, divide the workers into four groups: (a) good at both types of jobs; (b) bad at both types of jobs; (c) good at research, bad at finance; and (d) bad at research, good at finance. If local status becomes infinitely important in both occupations, then rewards become symmetric across occupations. Therefore, if academia is a small occupation, it will predominantly attract workers from the good at research, bad at finance group, whereas finance will attract most of the workers from the three remaining groups. Conse-

quently, finance will employ many workers who are bad at finance, whereas academia will employ only good academics.

Lemma 4. (i) There exists some $y > 0$ such that if $\delta \leq \min\{y, l_F\}$ and $M_A \in [0.5, 1)$, then $\bar{x}_A^A - \bar{x}_F^F < 0$ in $\psi^c(\cdot; M_A)$. (ii) For any $M'_A \in (0, 1)$ and $l_F \geq 0$, there exist some $\delta^*, d > 0$ such that if $\delta \geq \delta^*$ and $M_A \leq M'_A$ then $\bar{x}_A^A - \bar{x}_F^F > d$ in $\psi^c(\cdot, M_A)$.

Lemma 4(i) states that finance continues to attract workers of (on average) higher skill than academia in compensated equilibria in which academia is large ($M_A \geq 0.5$) as long as the importance of local status in academia relative to finance remains sufficiently small. To understand the intuition, first suppose that local status rewards are symmetric across occupations ($\delta = 0$). In that case, finance attracts better workers than academia as long as it is the smaller occupation, regardless of how much workers care about wages relative to local status (by the results in Section 4.1, Lemma 2, Equation (9), and Lemma 3). Naturally then, finance continues to attract workers of (on average) higher skill if local status is just slightly more important in academia than in finance.

Lemma 4(ii) states that if local status becomes sufficiently important in academia relative to finance, then academia attracts workers of (on average) higher skill than finance does. If the importance of local status in academia is very high, then academia attracts all workers who are highly skilled at research. Because the two skills are interdependent, this means that academia also attracts most of the workers who are highly skilled at finance, so that majority of the remaining bankers have low skill.

4.3.3 Local Status and Sorting

As a compensated equilibrium is an equilibrium if $c = 0$, I can use Lemmas 1 to 4 to examine how much of an impact the taste for local status can have on equilibrium sorting.

Proposition 2. Suppose $k = 0$. If $\delta \geq w_A - w_F(a_x)$ ($\delta \leq w_A - w_F(b_x)$) then academia (finance) unravels in the unique equilibrium, so that $\psi^e(x_A) = a_x$ ($\psi^e(x_A) = b_x$).

If local status becomes very important in academia relative to finance ($\delta \geq w_A - w_F(a_x)$), then the lowest-ranked academic receives a lower reward than a banker of skill a_x , regardless of that academic's skill. Hence no equilibrium with a positive size of academia can be supported: If the lowest-ranked worker leaves academia, the previously second-lowest-ranked worker becomes lowest-ranked and leaves too. This leads to a complete unraveling of the academic sector.¹³ Therefore, the relative nature of local status imposes a bound on the importance of local status in academia relative to finance. In particular, if academic wages are low, then local status can be only slightly more important in academia than in finance.

Theorem 2. Suppose that $k = 0$ and let I_W denote the set $(w_A - w_F(b_x), w_A - w_F(a_x))$. (i) There exists some $(\delta, l_F) \in I_W \times \mathbf{R}_{\geq 0}$ such that $\bar{x}_A^A - \bar{x}_F^F > 0$ and $M_A < 0.5$ in the unique equilibrium. (ii) However, if w_A is sufficiently close to $w_F(a_x)$, then there exists no $(\delta, l_F) \in I_W \times \mathbf{R}_{\geq 0}$ such that $\bar{x}_A^A - \bar{x}_F^F > 0$ and $M_A \geq 0.5$ in the unique equilibrium.

Proof. (i) Temporarily set $\delta = 0$, choose any $M'_A \in (0, \min\{H_M(\psi^b), 0.5\})$ and set $l_F > l_F^*$, where l_F^* is as in Lemma 3(ii). Setting $\delta = 0$ implies that $w_F(x_F^m) = w_A$ in equilibrium and thus $x_F^m = \psi^b$. Therefore, we have that $M_A(\psi^e) \geq H_M(\psi^b)$ and thus $M_A(\psi^e) > M'_A$. Denote the level of academic wages for which academia's size is M'_A in equilibrium by w'_A ; clearly, the equilibrium under w'_A is the same as the compensated equilibrium $\psi^c(\cdot; M'_A)$ under wage level w_A , with $c = w'_A - w_A$. Thus, as by Proposition 1(ii) the size of academia in a compensated equilibrium is increasing in c , the compensating differential for which $\psi^c(\cdot; M_A)$ is a compensated equilibrium (denoted by $c(M_A)$) is increasing in M_A ; as $M'_A < M_A(\psi^e)$ and $c(M_A(\psi^e)) = 0$ it follows that $c(M'_A) < 0$. Second, by Lemmas 2 and 3(ii), if $l_F > l_F^*$ and $\delta \geq 0$ then $\bar{x}_A^A > \bar{x}_F^F$. Finally, if $\delta \approx w_A - w_F(a_x)$ then $c(M'_A) > 0$ by Equation 7, because $c = w'_A - w_A$, for any alternative wage level w'_A . As $\psi^c(\cdot; M_A, \delta)$ is continuous in δ , there exists some $\delta' \in (0, w_A - w_F(a_x))$ for which $c(M'_A) = 0$. Thus if $\delta = \delta'$, then $M_A = M'_A$ and $\bar{x}_A^A > \bar{x}_F^F$ in equilibrium.

¹³The unraveling result *does not* depend on the assumption that academic wages are constant, or even on the assumption that wages are an exogenous function of skill. It requires only that the marginal product of every worker in academia is finite.

(ii) Suppose that $w_A - w_F(a_x) \leq \min\{y, 0.25(w_F(H_M^{-1}(0.5)) - w_F(a_x))\}$, so that $\delta < w_A - w_F(a_x)$ only if $\delta < \min\{y, 0.25(w_F(H_M^{-1}(0.5)) - w_F(a_x))\}$. Thus by Lemma 4(i) it suffices to show that if $l_F < \delta$, then $M_A(\psi^e) < 0.5$ in any equilibrium. Suppose, by way of contradiction, that $l_F < \delta$ and $M_A(\psi^e) \geq 0.5$. The latter implies that $x_F^s > H_M^{-1}(0.5)$, which yields

$$w_F(H_M^{-1}(0.5)) - l_F < w_F(x_F^s) + l_F(2G_F(x_F^s) - 1) = w_A + l_A(2G_A(x_A^s) - 1) \leq w_A + l_A.$$

As $\delta = l_A - l_F < w_A - w_F(a_x)$, it follows that

$$\begin{aligned} w_F(H_M^{-1}(0.5)) - w_F(a_x) &< w_A - w_F(a_x) + l_A - l_F + 2l_F \\ &< 2(w_A - w_F(a_x)) + 2l_F \\ &< \frac{w_F(H_M^{-1}(0.5)) - w_F(a_x)}{2} + 2l_F, \end{aligned}$$

which immediately implies that $l_F > 0.25(w_F(H_M^{-1}(0.5)) - w_F(a_x)) > \delta$. Contradiction! \square

The main take-away from Theorem 2 is that the relative nature of local status introduces a trade-off between the equilibrium size and quality of academia's workforce, and that this trade-off is particularly stark if academic wages are low. The intuition for this result builds on Proposition 2, Lemma 3 and Lemma 4. In particular, we know by now that (a) if local status is much more important in academia than in finance, then academia is small in equilibrium and (b) that if local status matters sufficiently strongly in both occupations, then the smaller occupation attracts better workers on average. It follows that local status concerns can, on their own, cause academia to attract workers of higher skill than finance (if both δ and l_F are sufficiently high). Crucially, if the academic wage is small, then this can be the case only if academia is the smaller occupation. For academics to be both more skilled (on average) and more plentiful than bankers, local status must be sufficiently more important in academia than in finance. However, this scenario is impossible if the academic wage is low, as then academia will become small as soon as local status is even slightly more important in academia than in finance!

4.4 The Interaction between Prestige and Local Status

In this section, I consider what happens if workers care about both occupational prestige and local status, in which case rewards are given by $t_A(x_A) = w_A + (l_F + \delta)(2G_A(x_A) - 1) + k o_A$ and $t_F(x_F) = w_F(x_F) + l_F(2G_F(x_F) - 1) + k o_F$. Crucially, because the set of compensated equilibria does not depend on k , the results from Section 4.3.2 remain relevant in this section.

Theorem 3. For any $M'_A \in (0, 1)$ and any $l_F \geq 0$, there exists some $\bar{\delta} \in \mathbf{R}_{\geq 0}$ such that if $\delta > \bar{\delta}$ and k is sufficiently high given δ then (i) academia is large ($M_A(\psi^e) > M'_A$) and attracts higher-quality talent than finance ($\bar{x}_A^A - \bar{x}_F^F > 0$) in all equilibria; and (ii) the set of equilibria is non-empty.

Proof. (i) Fix M'_A and l_F , and choose some $\delta > \max\{\delta^*, w_A - w_F(a_x)\} \equiv \bar{\delta}$, where δ^* is as in Lemma 4(ii). This immediately implies that $o_A > o_F$ in any equilibrium, provided such an equilibrium exists, as otherwise $w_A - \delta + k(o_A - o_F) < w_F(a_x)$ and academia unravels. Denote $\min\{4d, \frac{d}{M'_A(1-M'_A)}\}$ by d_o . The fact that $\delta > \delta^*$ implies that $o_A - o_F > d_o > 0$ in any compensated equilibrium in which $M_A \leq M'_A$ (by Lemma 4(ii) and the fact that $o_A - o_F = \frac{\bar{x}_A^A - \bar{x}_F^F}{M_A(1-M_A)}$). Consider any $k' \geq \frac{w_F(b_x) - w_A + \delta}{d_o} \equiv \bar{k}$. Consider an alternative academic wage w'_A for which there is an equilibrium of size M_A ; then the compensating differential corresponding to $\psi^e(\cdot; M_A)$ is equal to $c = w'_A - w_A$ and it follows from Equation (5) that

$$c = w_F(x_F^m) - w_A + \delta - k'(o_A - o_F),$$

which is strictly negative for any $M_A \leq M'_A$. Hence there exists no equilibrium in which $M_A(\psi^e) \leq M'_A$, and thus $M_A(\psi^e) > M'_A$ in any equilibrium.

(ii) We are left to show that there exists at least one equilibrium. Let us start by temporarily setting k to 0 and denoting the (clearly unique) x that solves $\bar{x}/H_M(x) = x$ as \tilde{x} . Consider some $w'_A \in (w_F(\tilde{x}) + \delta, w_F(b_x) + \delta)$, and note that Proposition 1 implies that there exists a unique ψ^c that corresponds to $c = w'_A - w_A$; let $M''_A = M_A(\psi^c)$ denote the size of academia in that compensated equilibrium. Because $w_F(x_F^m) = w_A - \delta + c$, it follows

that $x_F^m > \tilde{x}$ in this compensated equilibrium, and hence $M_A'' > H_M(\tilde{x})$ and $\bar{x}_F^F > \tilde{x}$. Finally, because $M_A \bar{x}_A^A + (1 - M_A) \bar{x}_A^F = \bar{x}$, where \bar{x}_A^j denotes the average academic skill among members of occupation j , we have that $\bar{x}_A^A < \bar{x}/H_M(\tilde{x})$, and thus $\bar{x}_A^A < \tilde{x}$. It follows, therefore, that $o_F - o_A > 0$, which implies that $M_A'' > M_A'$. Finally, as the set of compensated equilibria does not depend on k , this compensated equilibrium exists if $k = k'$, where $c(k') > 0$ because of academia's negative prestige. As $c(k', M_A') < 0$ and $c(k', M_A'') > 0$, the continuity of ψ^c with respect to M_A implies that there has to exist some $M_A > M_A'$ such that $c(k', M_A) = 0$, which concludes the proof. \square

Theorem 3 states that if the importance of local status in academia relative to finance is sufficiently high and workers care about occupational prestige sufficiently strongly, then there must exist an equilibrium, and academia must be large and attract workers of (on average) higher skill than finance in any equilibrium. This result is quite remarkable: Given that, on its own, occupational prestige decreases the size of academia, one might expect that Theorem 2 captures the absolute limit of what social status can accomplish. And yet it turns out that the interaction between the two status components can have an arbitrarily strong impact on sorting.¹⁴

How is this possible? As the joint impact of occupational prestige and local status is much greater than the sum of their individual impacts, it stands to reason that there exists some complementarity between the two components of social status. Specifically, occupational prestige and local status act as complements in regard to the compensation $w_A - (l_F + \delta) + k o_A(M_A)$ received by the lowest-ranked academic in the compensated equilibrium $\psi^c(\cdot; M_A)$:

$$\frac{\partial^2}{\partial k \partial \delta} (w_A - (l_F + \delta) + k o_A(M_A)) = \frac{\partial}{\partial \delta} o_A(M_A) > 0,$$

where the inequality follows from Lemma 2 and Equation (9). Intuitively, in any compensated equilibrium, high δ provides the differentiation of rewards needed for academia to attract workers of high skill, which increases the prestige of academia. Once the average skill of academics is high

¹⁴It is also worth noting that Theorem 3 holds l_F fixed, just as Theorem 2 kept k fixed. Thus the two results allow for the same number of degrees of freedom.

enough, the taste for occupational prestige increases the level of rewards in academia, instead of decreasing it as in the prestige-only case. This in turn relaxes the bound on the importance of local status in academia relative to finance, which prevents the unraveling of academia when δ is high.

5 Concluding Remarks

To conclude, I will discuss (a) the policy implications of my results (Section 5.1), (b) two alternative mechanisms that could explain the puzzle of selection into academia (Section 5.2) and (c) two other occupations in which social status likely plays an important role in determining selection (Section 5.3).

5.1 Policy Implications

The results in this paper have significant policy implications, mostly because they suggest a novel relationship between the level of income taxation and selection patterns. To see this, note that the strength of the desire for status depends on the extent to which workers' *real* wages depend on their choice of occupation and their occupation-specific skill. If income taxes were very high, then the choice of occupation would result in very small differences in the real wage, which would make social status a very important aspect of occupational choice.

To be more specific, suppose that taxes are linear and denote the tax rate by τ . Equation (3) and Definition 1 imply that a model with tax rate τ and social status parameters (l_A, l_F, k) is equivalent to a model with no tax and social status parameters $(\bar{l}_A, \bar{l}_F, \bar{k}) \equiv (\frac{l_A}{1-\tau}, \frac{l_F}{1-\tau}, \frac{k}{1-\tau})$. Therefore, an increase in the tax rate is equivalent to a proportional increase in δ , k and l_F .

If local status is more important in academia than finance ($l_A > l_F$) and workers care about occupational prestige at least a little ($k > 0$) then a sufficiently high tax rate guarantees that academia attracts workers of higher skill than finance.¹⁵ Guaranteeing that academia attracts more workers

¹⁵If $\tau = 1$ and $l_A > l_F$ then there can be no non-degenerate equilibrium in which academia is less prestigious than finance, as the lowest ranked academic would always

than finance is trickier. In fact, it can be shown that if k is small enough compared to l_A , then academia must be smaller than finance even if $\tau = 1$.

Thus, in some cases the tax rate may be too blunt a tool to ensure that large numbers of highly skilled workers become academics. Luckily, the government can also plausibly manipulate l_A and l_F directly. For example, the government could introduce (or eliminate) awards for the best research and thus increase (decrease) l_A . In finance, a significant portion of the information about rank is likely to be signalled through conspicuous consumption, and thus an increase in excise taxation on luxury goods is likely to decrease l_F .

Of course, if the government is able to set l_A, l_F and τ at will, then they can implement any combination of $(\bar{l}_A, \bar{l}_F, \bar{k})$ and hence sustain essentially any selection patterns they wish. Interestingly, if k is small compared to the initial l_A , a policy that would both increase the size of academia and improve selection into it may, somewhat counter-intuitively, require putting less emphasis on local status rewards in academia. This is because the high value of \bar{l}_A will be achieved by setting a high tax rate—in which case a low value of l_A is needed to ensure that occupational prestige features heavily in the workers' occupational choice.

5.2 Alternative Explanations

The puzzle of positive selection into academia can be explained by mechanisms other than the interaction of local status and occupational prestige. In this section I discuss the two most natural alternative explanations—preference heterogeneity and capacity constraints in academia.

5.2.1 Preferences

The simplest framework that allows the study of the impact of preferences on selection is a standard normal Roy's model. Specifically, suppose that every agent is characterised by a three-dimensional vector (x_A, x_F, x_P) ,

prefer to work in finance otherwise. A non-degenerate equilibrium does exist, however, because (a) Lemma 3 and the proof of Lemma 4 ensure that if $\tau = 1$, then $\bar{x}_A^A > \bar{x}_F^F$ in any compensated equilibrium in which academia is smaller than finance, and thus (b) $o_A - o_F$ goes to infinity as M_A goes to zero. Hence, no matter how small k is, there will exist some small value of M_A for which the economy will be an equilibrium.

distributed according to a standard tri-dimensional normal distribution, with ρ_{ij} denoting the correlation between x_i and x_j . The new random variable, x_P captures the worker's relative preference for working in finance.

Without status concerns, the payoff the worker receives in academia is equal to their academic wage, whereas the payoff in finance is a product of the wage and the relative preference for finance

$$t_A(x_A) = w(x_A), \quad t_F(x_F) = w_F(x_F) + x_P + \mu_P,$$

where $w_i(x_i) = w_i + \sigma_i x_i$. Therefore, a worker is willing to work in finance for a lower wage than in academia if and only if $\sigma_P x_P + \mu_P > 1$.

Using standard properties of the joint normal distribution, one can show that

$$\bar{x}_A^A = \frac{\phi(z) (\sigma_A - \rho_{FA}\sigma_F - \sigma_P\rho_{PA})}{\sigma_V\Phi(z)}, \quad \bar{x}_F^F = \frac{\phi(z) (\sigma_F - \rho_{FA}\sigma_A + \sigma_P\rho_{PF})}{\sigma_V(1 - \Phi(z))},$$

where $\sigma_V = \sqrt{\text{Var}(t_F(X_F) - t_A(X_A))}$, $z = (w_A - w_F - \mu_P)/\sigma_V$, and $\phi(\cdot)$ and $\Phi(\cdot)$ denote the pdf and cdf of the standard normal distribution, respectively.

The size of academia is equal to $\Phi(z)$ and the size of finance is equal to $1 - \Phi(z)$. For simplicity, let us restrict attention to the case where $\bar{x}_F^F \geq 0$, so that selection into finance is positive. If this is the case, then academia can be both larger than finance ($\Phi(z) \geq 0.5$) and attract more skilled workers than finance ($\bar{x}_A^A > \bar{x}_F^F$) only if

$$\sigma_A - \rho_{FA}\sigma_F - \sigma_P\rho_{PA} > \sigma_F - \rho_{FA}\sigma_A + \sigma_P\rho_{PF}$$

which reduces to

$$\frac{\sigma_F - \sigma_A}{\sigma_P} < -\frac{\rho_{PA} + \rho_{PF}}{1 + \rho_{PA}}. \quad (10)$$

Of course, workers' preferences can explain the observed patterns of selection: Indeed, sufficiently rich preferences can explain virtually all phenomena. However, the conditions needed for these selection patterns to emerge are fairly strong. First, if wages are more differentiated in finance than academia ($\sigma_F - \sigma_A > 0$), then it is not at all sufficient that workers prefer academia over finance: In fact, the average preference for academia,

$-\mu_P$, does not appear in Equation (10). What is necessary is that workers' preferences for finance and academia are sufficiently *heterogenous*, that is, that σ_F is sufficiently large. However, and second, preference heterogeneity is also not sufficient: On top of that, it must be the case that the relative preference for academia ($-x_P$) is correlated more strongly with the academic skill than the relative preference for finance (x_P) is correlated with financial skill ($\rho_{PA} + \rho_{PF} < 0$). In other words, it is not enough that workers with high academic skill like working in academia much more than workers with low academic skill: It must also be the case that skilled bankers enjoy working in finance not that much more than low-skilled bankers.

An empirical researcher interested in determining whether preferences or social status are the main reason why selection into academia is positive should turn their attention to the response of selection to exogenous changes in wages and the distribution of skill: A change in the composition academia's workforce induced by the exogenous change will not alter anyone's enjoyment from being an academic, but it will affect their social status. To be more specific, consider an increase in w_A . In the model with preference heterogeneity and normally distributed skills, under the assumption that $z > 0$ and $\bar{x}_A^A > \bar{x}_F^F > 0$, this would result in a worsening of the distribution of skill in academia and an improvement in the distribution of skill in finance, both in the monotone likelihood ratio (MLR) sense.¹⁶ In other words, the influx of low skilled academics would be proportionally greater than the influx of high- and medium-skilled academics: All workers benefit equally from an increase in w_A , but most high and medium-skilled workers have already joined academia.

In the model with local status, however, an increase in w_A benefits medium-skilled workers more than low-skilled workers, and thus is unlikely to result in an MLR worsening of the skill distribution. The reason is that lowest-skilled academics are always lowest-ranked as well ($G_A(a_X) = 0$), so that the change in skill distribution induced by the change in wages would not affect their local status. Medium-skilled workers would, however, enjoy an increase in their local status—as the distribution of skill worsens in academia, workers of the same skill would end up having a higher rank.

¹⁶This follows from the formula for conditional probabilities for bivariate normal variables and the log-concavity of the univariate normal distribution.

5.2.2 Capacity Constraints in Academia

A second alternative explanation of the puzzle is that there is a fixed (but possibly quite high) number of jobs in academia. If working in academia is extremely pleasant and universities are able to screen for academic ability, then most people would want to work in academia. However, due to the limited number of academic jobs only the highest skilled would be actually hired by universities. As a result, academia could end up with highly-skilled workers despite paying low and flat wages.

This seemingly plausible explanation has a major flaw. Namely, it is very hard to see how could a situation arise in which academia can screen for ability and yet pays higher-than-market-clearing wages. In particular, this could possibly happen only if universities have some degree of monopsony power: Otherwise, some university would profitably deviate by offering a much lower wage to workers. Under the assumption that all workers receive the same wage, a monopsonist may find it optimal to offer higher-than-market-clearing wages: Offering too high a wage allows the monopsonist to have their pick of workers. Critically, however, given that the monopsonist is able to screen for ability they should be able to pay wages that depend on ability; and skill-dependent wages allow the monopsonist to attract high skilled workers without leaving them any rents. But of course, if there are no rents then the market clears and the puzzle remains unexplained.

5.3 Other Relevant Occupations

In this section I will discuss two further examples of occupations in which social status concerns are important for selection: The civil service and the military.

The puzzle of selection into civil service seems to be as stark as the puzzle of selection into academia. [Lucifora and Meurs \(2006\)](#) find that while low-skilled workers are paid more in the civil service than in the private sector, the opposite is true for highest-skilled workers; in other words, the wage schedule is flatter in the civil service. In addition, they document that while the service sector pays significantly more on average than the private sector, around half of that difference is accounted for by differences in observable characteristics, suggesting that the selection into

the civil service is better than into the private sector.

Social status, and particularly local status, likely plays an important role in determining the selection into the British civil service, at the very least. First of all, the civil service has an inherently hierarchical structure, with well-defined and easily observed grades; this should translate into high l_A . Second, the famous Whitehall I and II studies have shown (Singh-Manoux, Adler, and Marmot, 2003; Singh-Manoux, Marmot, and Adler, 2005) that civil servants with lower subjective social status have significantly worse health outcomes even after controlling for education and income. As health is an important component of ones well-being, this suggests a very strong, direct link between relative position and utility. In addition, it is also well-known that a disproportionately large proportion of the British New Year’s honours (a pure status reward) is awarded to civil servants (Phillips, 2004). In other words, if one dreams of becoming a Dame or a Lord, becoming a top-ranked civil servant is likely their best bet.

Another example of an occupation in which local status plays an important role and wages are flat is the military. Asch and Warner (2001) report that the wage schedule in U.S. army is much flatter than in the private sector. At the same time, the military is a famously hierarchical occupation, with well-established status rewards, such as medals and orders. It is, therefore, quite plausible that the military also uses local status to substitute for steeper wage schedules.

A Omitted Proofs

Proof of Proposition 1

(i) Because $F(x_A, x_F)$ is continuously differentiable in x_F on $[a_x, b_x]$, it is also Lipschitz continuous in x_F . Therefore, the IVP defined by Equation (7) has a unique solution on D for a given $M_A \in (0, 1)$. The resulting separation function, denoted by $\psi(\cdot; M_A)$, is an equilibrium if and only if it also satisfies Equation (8).

As $\frac{\partial}{\partial M_A} F(x_A, x_F) < 0$ it follows by Theorem 6 in Birkhoff and Rota (1969) and the Comparison Theorem (Theorem OA.1 in the Online Ap-

pendix C of this paper), that $\psi(\cdot; M_A)$ (and thus also the RHS of Equation (8)) is continuous and strictly decreasing in M_A ; thus, if a compensated equilibrium exists, it must be unique. To show existence, first note that because $\psi(x_A; M_A)$ is increasing in x_A it must be the case that

$$\int_{a_x}^{b_x} \frac{\partial}{\partial x_A} H(r, \psi(r)) dr \geq \int_{a_x}^{b_x} \frac{\partial}{\partial x_A} H(r, x_F^m) dr = x_F^m > 0.$$

Pick an arbitrary $M'_A \in (0, 1)$ and denote $\int_{a_x}^{b_x} \frac{\partial}{\partial x_A} H(r, \psi(r; M'_A)) dr$ by M''_A . If $M'_A > M''_A$ then existence follows from the intermediate value theorem because the RHS of Equation (8) is greater than M_A for any $M_A < x_F^m$. If $M'_A < M''_A$ then existence again follows from the intermediate value theorem, because

$$M''_A = \int_{a_x}^{b_x} \frac{\partial}{\partial x_A} H(r, \psi(r; M'_A)) dr > \int_{a_x}^{b_x} \frac{\partial}{\partial x_A} H(r, \psi(r; M''_A)) dr.$$

To see that M_A is increasing in w_A , consider any $w''_A > w'_A$ and suppose that $M_A(w''_A) \leq M_A(w'_A)$. Then by Equation (8) and the Comparison Theorem, we have that $M_A(w''_A) > M_A(w'_A)$. Contradiction!

(ii) For $w_A < w_F(a_x) + \delta$ ($w_A > w_F(b_x) + \delta$) we have that $t_A(x_A^m) + c < (>)t_F(x_F^m)$, which contradicts Equation (5). For $w_A = w_F(b_x) + \delta$ we have that $x_F^m = b_x$ and thus $M_A = 1$. Finally, if $w_A = w_F(a_x) + \delta$ then $x_A^m \geq 0$ and $x_F^m = a_x$, and the boundary condition in Equation (7) changes to $\psi^e(x_A^m) = a_x$. But then the unique solution to the IVP problem is $\psi^e(x_A) = a_x$ and thus $M_A = 0$.

Proof of Lemma 1

First, suppose that $k = 0$ and consider an alternative level of academic wages w'_A . Clearly, the unique equilibrium under w'_A is a compensated equilibrium under w_A , with $c = w'_A - w_A$. Therefore, any compensated equilibrium must correspond to a compensating differential $c \in (w_F(a_x) - w_A + \delta, w_F(b_x) - w_A + \delta)$ (by Proposition 1 (ii)) and must be characterized by the IVP defined in Equation (7) (with the initial condition $\psi^c(a_x) = w_F^{-1}(w_A + c - \delta)$). The solution to the IVP defined in Equation (7) can be equally well expressed as a function of the compensating differential for a

given $M_A: \psi(x_A; c)$. It follows by a reasoning analogous to that in the proof of Proposition 1 that (a) $\psi(x_A; c)$ is continuous and increasing in c ; (b) for $c = w_F(a_x) - w_A + \delta$ we have $\psi(x_A; c) = a_x$ and thus $M_A(\psi(x_A; c)) = 0 < M_A$, and (c) for $c = w_F(b_x) - w_A + \delta$ we have $\psi(x_A; c) = b_x$ and thus $M_A(\psi(x_A; c)) = 1 > M_A$. Existence follows the intermediate value theorem, uniqueness follows from the monotonicity of $\psi(x_A; c)$ in c , whereas continuity of $\psi^c(\cdot, M_A)$ wrt M_A is a consequence of the continuity of $\psi(x_A; c)$ wrt c . Second, ψ is a compensated equilibrium for $k = 0$ if and only if it is a compensated equilibrium for any $k > 0$, as k affects only the value of the corresponding c ; thus, the results hold for any value of $k \geq 0$.

Proof of Lemma 2

(i) It follows immediately from the Comparison Theorem that if $F(x_A, x_F)$ increases strictly for all $(x_A, x_F) \in (a_x, b_x]^2$ and $\psi^c(x'_A; M_A)$ increases for some $x'_A \in (a_x, b_x)$, then $\psi^c(x'_A; M_A)$ must strictly increase for all $x_A \in [x'_A, b_x]$. As a strict increase in $\psi^c(\cdot, M_A)$ for all x_A violates Equation (8), there must exist some x''_A such that $\psi^c(x_A; M_A)$ falls strictly for all $x_A < x''_A$ and increases otherwise, which implies that $G_A(x_A)$ falls for all $x_A \in [a_x, b_x]$. To characterize the density of skill in finance, we can use (an extension of) a right inverse of $\psi^c(\cdot; M_A)$, denoted by $\phi^c(\cdot; M_A) : [a_x, b_x] \rightarrow [a_x, b_x]$:

$$\phi^c(x_F) = \sup\{x_A \in [a_x, b_x] : \psi^c(x_A) < x_F\}. \quad (11)$$

Clearly then

$$g_F(x_F) = \frac{h_M(x_F)\Pr(X_A < \phi^c(x_F)|X_F = x_F)}{1 - M_A}.$$

The results for ψ^c imply that there must also exist such x''_F that $\phi^c(\cdot; M_A)$ strictly increases for all $x_F < x''_F$ and falls otherwise. This concludes the proof.

Proof of Lemma 3

(i) Consider a separation function ψ^a which solves:

$$\frac{\partial}{\partial x_A} \psi^a(x_A) = \frac{1}{M_A} \frac{\frac{\partial}{\partial x_A} H(x_A, \psi^a(x_A))}{0.5\beta w'_F(\psi^a(x_A)) + \frac{1-\beta\delta}{1-M_A} \frac{\partial}{\partial x_F} H(x_A, \psi^a(x_A))} \quad (12)$$

for $x_A \in (x_A^{ma}, x_A^{sa})$ and

$$M_A = \int_{a_x}^{b_x} \frac{\partial}{\partial x_A} H(\psi^a(r, x_F^{ma}, x_A^{ma}), r) dr. \quad (13)$$

Because $x_A^{ma} > a_x$ implies $\psi^a(x_A^{ma}) = a_x$, it follows by the same logic as in the proof of Lemma 1 that $x_A^{sa} = a_x$.¹⁷ Further, $H_M(x_A^{sa}) > 1 - M_A$, as otherwise the RHS of Equation (13) must be larger than M_A . Choose some $\bar{x}_A \in (0, H_M^{-1}(1 - M_A)) \subset (x_A^{ma}, x_A^{sa})$ and suppose that $\psi^a(\bar{x}_A) = \alpha$. Because the RHS of Equation (12) is Lipschitz-continuous on $(a_x, b_x)^2$, it follows from Theorem 4.32 in Precup (2018) that there exists a unique $\psi^a(\cdot; \alpha)$ which solves the initial value problem given by Equation (12) and $\psi^a(\bar{x}_A) = \alpha$. Thus, it follows by a reasoning analogous to that in points (b) and (c) in the proof of Lemma 1 that for any β there exists a unique ψ^c that satisfies Equations (12) and (13).

Clearly, the problem defined by Equations (12) and (13) is equivalent to the problem defined by Equations (7)-(8) for $\beta = 1/(l_F + \delta)$, as long as $\beta > 0$. As $\psi^a(\cdot; \beta)$ is continuous in β , it follows that $\lim_{l_F \rightarrow \infty} \psi^c(\cdot; M_A; l_F) = \psi^a(\cdot; M_A, \beta = 0)$, so that the limiting compensating equilibrium exists and is unique.

Set $\beta = 0$. Because $G_A^a(x_A^{ma}) = G_F^a(x_F^{ma}) = 0$, integrating Equation (12) reveals that $G_F^a(\psi^a(x_A)) = G_A^a(x_A)$ for any M_A , which implies that $x_A^{sa} = x_F^{sa} = b_x$. Note that if $M_A = 0.5$, then $\psi^a(x_A) = x_A$. As the RHS of Equation (12) is decreasing in M_A , it follows from the Comparison Theorem that if $M_A \geq (\leq) 0.5$ and $\psi^a(x_A) < (>) x_A$ for any $x_A \in (a_x, b_x)$, then $x_F^{sa} \neq b_x$ ($x_A^{sa} \neq b_x$); contradiction!¹⁸ Therefore, if $M_A \geq (\leq) 0.5$ then $\psi^a(x_A) \geq (\leq) x_A$ for all x_A , which implies that $G_F^a(x) \leq (\geq) G_A^a(x)$.

¹⁷Note that the RHS in Equation (12) is Lipschitz-continuous on $(a_x, b_x) \times [a_x, b_x]$ for any $\beta \geq 0$.

¹⁸The RHS of Equation (12) is Lipschitz-continuous on $x_A \in [x, 1]$ for any $x > 0$, so the Comparison Theorem applies.

(ii) The result follows from Lemma 2, Lemma 3 (i) and the fact that $\psi^c(\cdot; M_A; l_F)$ is continuous in l_F .

Proof of Lemma 4

(i) Define $w_{F\min'} \equiv \min_{x_F \in [a_x, b_x]} \frac{w'_F(x_F)}{h_M(x_F)}$. I will prove this result in three steps.

STEP 1: If $M_A \geq 0.5$ and $\delta \leq 0.5(1 - M_A)w'_{F\min}$ then $\bar{x}_F^F > \bar{x}_A^A$ in $\psi^c(\cdot; M_A)$.

Under these conditions $\bar{x}_F^F - \bar{x}_A^A$ is decreasing in l_F by Equation (9) and Lemma 2. The result follows because $\lim_{l_F \rightarrow \infty} \bar{x}_F^F - \bar{x}_A^A > 0$ by Lemma 3.

Stating the second step requires new notation. First, denote $\Pr(X_A < x_A | X_F = x_F) = \frac{\partial}{\partial x_F} H(\phi^c(x_F), x_F) / h_M(x_F)$ as $P(x_A | x_F)$. Second, for any $M_A \in [0.5, 1]$ denote the $x_A \in (0, 1)$ for which $\min_{x_F \in [a_x, b_x]} \frac{\partial}{\partial x_F} P(x_A | x_F) = 2(1 - M_A)$ by $x'_A(M_A)$.¹⁹ Third, define

$$t(M_A) \equiv \max_{x_F \in [a_x, b_x]} P(x'_A(M_A) | x_F) \in [2(1 - M_A), 1].$$

Lastly, define $z \equiv \min_{x_F \in [a_x, b_x]} 1 / [(b_x - a_x)h_M(x_F)]$, with $z \in (0, 1]$.²⁰

STEP 2: If $\delta \leq l_F \left(\frac{M_A}{t(M_A)} \frac{4M_A - 4 + z}{2M_A - 2 + z} - 1 \right)$ and $M_A > 1 - z/2$, then $\bar{x}_F^F > \bar{x}_A^A$ in $\psi^c(\cdot; M_A)$.

Define $\alpha \equiv \frac{t(M_A)l_A}{M_A l_F}$; note that $l_A - l_F = \delta \leq l_F \left(\frac{M_A}{t(M_A)} \frac{4M_A - 4 + z}{2M_A - 2 + z} - 1 \right)$ implies that $1 - \alpha \geq 0$.

First, I will show that, for any x_F , either $G_F(x_F) \leq \alpha H_M(x_F)$ or $G_F(x_F) \leq 2H_M(x_F) - 1$, which would imply that $H_M(x_F) - G_F(x_F) \geq K(H_M(x_F))$, where

$$K(s) \equiv \begin{cases} (1 - \alpha)s & \text{if } s \in [0, 1/(2 - \alpha)] \\ 1 - s & \text{if } s \in (1/(2 - \alpha), 1]. \end{cases} \quad (14)$$

Case 1: $x_F \leq x_F^m$. In that case, $G_F(x_F) = 0 \leq \alpha x_F$ and the result follows immediately.

Case 2: $x_F \in (x_F^m, \psi^c(x'_A(M_A)))$. First, note that in this case we

¹⁹ $x'_A(M_A)$ exists and is unique, because $\min_{x_F \in [a_x, b_x]} P(x_A | x_F)$ is continuous and strictly increasing in x_A by the Envelope Theorem, and $P(0 | x_F) = 0$, $P(1 | x_F) = 1$.

²⁰ $z > 0$ because H is twice continuously differentiable on its support, and $z \leq 1$ because $1 = H_M(b_x) \leq 1/z$.

have $x_F \leq x_F^s$.²¹ It follows, therefore, that $\phi^c(x_F) \leq x'_A(M_A)$ and thus $\frac{\partial}{\partial x_F} H(\phi^c(x_F), x_F) \leq \frac{\partial}{\partial x_F} H(x'_A(M_A), x_F) \leq h_M(x_F)t(M_A)$, where $\phi^c(\cdot)$ is as defined in Equation (11). Because all workers join some occupation, it follows that

$$M_A G_A(\phi^c(x_F)) + (1 - M_A) G_F(x_F) = H(\phi^c(x_F), x_F), \quad (15)$$

which implies that $G_A(\phi^c(x_F)) \leq \frac{H(\phi^c(x_F), x_F)}{M_A}$. Note that Equation (3) implies that

$$t_F(x_F) - t_F(x_F^m) = t_A(\phi^c(x_F)) - t_A(\phi^c(x_F^m)) \quad (16)$$

so that

$$\begin{aligned} G_F(x_F) &\leq G_F(x_F) + \frac{w_F(x_F) - w_F(x_F^m)}{2l_F} \\ \text{[by Equation (16)]} &= \frac{l_A}{l_F} G_A(\phi^c(x_F)) \\ \text{[by Equation (15)]} &\leq \frac{l_A}{M_A l_F} H(\phi^c(x_F), x_F) \leq \frac{t(M_A) l_A}{M_A l_F} H_M(x_F) = \alpha H_M(x_F). \end{aligned}$$

Case 3: $x_F \in [\psi^c(x'_A), 1]$. In that case, $g_F(x_F) \geq \frac{\partial}{\partial x_F} H(x'_A(M_A), x_F)/(1 - M_A) \geq 2h_M(x_F)$. As $1 - G_F(x_F) = \int_{x_F}^1 g_F(r) dr$, it follows that $G_F(x_F) \leq 1 - 2(1 - H_M(x_F)) = 2H_M(x_F) - 1$.

Second, denote by $G_A^F(\cdot)$ the cdf of the distribution of X_A among bankers and notice that $H_M(x_A) - G_A(x) < 1 - M_A$, because $H_M(x_A) - M_A G_A(x_A) = (1 - M_A) G_A^F(x_A)$.

Finally, because

$$\bar{x}_i^i - \bar{x} = \int_{a_x}^{b_x} H_M(x) - G_i(x) dx,$$

it follows that

$$\bar{x}_F^F - \bar{x} \geq \int_0^1 \frac{K(r)}{h_M(H_M^{-1}(r))} dr \geq z(b_x - a_x) \int_0^1 K(r) dr = \frac{z(b_x - a_x)(1 - \alpha)}{2(2 - \alpha)}$$

²¹From the definition of set D and the increasingness of ψ^c follows that $\max_{x_A \in [a_x, b_x]} \psi^c(x_A) = x_F^s$.

and $\bar{x}_A^A - \bar{x} < (b_x - a_x)(1 - M_A)$, so that

$$\bar{x}_F^F - \bar{x}_A^A > (b_x - a_x) \left(\frac{z(1 - \alpha)}{2(2 - \alpha)} + M_A - 1 \right).$$

A little algebra reveals that if $\delta \leq l_F \left(\frac{M_A}{t(M_A)} \frac{4M_A - 4 + z}{2M_A - 2 + z} - 1 \right)$ and $M_A > 1 - z/2$ then $\bar{x}_F^F - \bar{x}_A^A > 0$.

STEP 3: There exists an $\bar{M}_A \in (1 - z/4, 1)$, such that $\frac{M_A}{t(M_A)} \frac{4M_A - 4 + z}{2M_A - 2 + z} \geq 2$ for all $M_A \in (\bar{M}_A, 1)$.

As $P(x_A|x_F) = 0$ if and only if $x_A = a_x$, it follows that $x'_A(1) = a_x$ and thus $t(b_x) = 0$. Because $t(M_A)$ is differentiable by the Envelope Theorem, it is also continuous and the result follows.

Together, Steps 1 to 3 prove the result, with $y = 0.5(1 - \bar{M}_A)w'_{F\min}$.

(ii) First, observe that in compensated equilibrium $\psi^c(\cdot; M_A)$, the average occupation- $j \in \{A, F\}$ specific skill among academics is:

$$\bar{x}_j^A \equiv G_A(x_A^{sc})E(X_j|X_F < \psi^c(X_A; M_A), X_A < x_A^{sc}) + (1 - G_A(x_A^{sc}))E(X_j|X_A \geq x_A^{sc}).$$

Second, note that $M_A x_F^A + (1 - M_A)x_F^F = \bar{x}$, and thus if $x_F^A \geq \bar{x}$ then $x_F^F \leq \bar{x}$. Therefore, it is sufficient to show that $E(X_j|X_A \geq x_A^{sc})$ is bounded above \bar{x} and that $G_A(x_A^{sc}) \approx 0$ for all j and sufficiently high δ .

First, I will bound $G_A(x_A^{sc})$ from above. By Equation (5), we have that

$$w_F(b_x) - w_F(a_x) + 2l_F \geq w_F(x_F^s) - w_F(x_F^m) + 2l_F G_F(\phi(x_A)) = 2l_A G_A(x_A^{sc}),$$

which implies that $G_A(x_A^{sc}) \leq \frac{0.5(w_F(b_x) - w_F(a_x)) + l_F}{l_A}$.

Second, I will bound $E(X_j|X_A \geq x_A^{sc})$ from below. Because all workers with $x_A > x_A^{sc}$ join academia, it must be the case that $H_M(x_A^{sc}) > 1 - M_A \geq 1 - M'_A$ in any compensated equilibrium. Clearly, thus

$$E(X_A|X_A \geq x_A^{sc}) \geq E(X_F|X_A \geq x_A^{sc}) \geq b \equiv \min_{x_A \in [H_M^{-1}(1 - M'_A), b_x]} E(X_F|X_A \geq x_A).$$

Denote $\Pr(X_F \leq x_F|X_A \geq x_A)$ by $F(x_F|x_A)$; because $(1 - F(x_F|x_A))(1 - H_M(x_A)) = 1 - H_M(x_A) - H_M(x_F) + H(x_A, x_F)$ it follows from $H(x_A, x_F) > H_M(x_A)H_M(x_F)$ that $F(x_F|x_A) < H_M(x_F)$ for all $(x_A, x_F) \in (a_x, b_x)^2$. We have, thus, that $b > \bar{x}$.

To conclude the proof, set $d = 0.5(b - \bar{x})$. Then it follows that if $\delta \geq \frac{b(w_F(b_x) - w_F(a_x) + 2l_F) + \bar{x}l_F}{b - \bar{x}}$ then $\bar{x}_A^A, \bar{x}_F^A > \bar{x} + d$ and thus $\bar{x}_A^A - \bar{x}_F^F > \bar{x}_A^A - \bar{x} > d$.

Proof of Proposition 2

If $\delta \geq w_A - w_F(a_x)$ ($\delta \leq w_A - w_F(b_x)$) then $w_F(a_x) - w_A + \delta \geq 0$ ($w_F(b_x) - w_A + \delta \leq 0$). Thus, by Proposition 1 and Lemma 1 it follows that $c > 0$ ($c < 0$), in any $\psi^c(\cdot, M_A)$, and thus we have that $\psi^e(x_A) = a_x$ ($\psi^e(x_A) = b_x$) in the unique equilibrium.

References

- Ager, P., L. Bursztyn, L. Leucht, and H.-J. Voth (2021). Killer Incentives: Relative Position, Performance and Risk-Taking among German Fighter Pilots, 1939-45 . *Review of Economic Studies*, forthcoming.
- Albornoz, F., A. Cabrales, and E. Hauk (2020). Occupational choice with endogenous spillovers. *The Economic Journal Accepted Author Manuscript*(n/a).
- Amir, R. and M. Knauff (2008, August). Ranking Economics Departments Worldwide on the Basis of PhD Placement. *The Review of Economics and Statistics* 90(1), 185–190.
- Asch, B. J. and J. T. Warner (2001). A theory of compensation and personnel policy in hierarchical organizations with application to the united states military. *Journal of Labor Economics* 19(3), 523–562.
- Becker, G. S. and K. M. Murphy (2000). *Social Economic: Market Behavior in a Social Environment*. Harvard University Press.
- Birkhoff, G. and G.-C. Rota (1969). *Ordinary Differential Equations*. Blaisdell Publishing Company.
- Bó, E. D., F. Finan, O. Folke, T. Persson, and J. Rickne (2017). Who Becomes A Politician? *The Quarterly Journal of Economics* 132(4), 1877–1914.
- Borjas, G. J. (1987, September). Self-Selection and the Earnings of Immigrants. *American Economic Review* 77(4), 531–53.
- Bottan, N. L. and R. Perez-Truglia (2020). Choosing your pond: Location choices and relative income. *Review of Economics and Statistics*, forthcoming.
- Bursztyn, L., B. Ferman, S. Fiorin, M. Kanz, and G. Rao (2018, 8). Status Goods: Experimental Evidence from Platinum Credit Cards*. *The Quarterly Journal of Economics* 133(3), 1561–1595.
- Card, D., A. Mas, E. Moretti, and E. Saez (2012, October). Inequality

- at Work: The Effect of Peer Salaries on Job Satisfaction. *American Economic Review* 102(6), 2981–3003.
- Cardoso, A. R. (2012). Money and rank in the labor market. *Economics Letters* 115(2), 325–328.
- Damiano, E., H. Li, and W. Suen (2010, February). First in Village or Second in Rome? *International Economic Journal* 51(1), 263–288.
- Damiano, E., H. Li, and W. Suen (2012). Competing for Talents. *Journal of Economic Theory* 147(6), 2190–2219.
- Dasgupta, P. and P. A. David (1987). *Information Disclosure and the Economics of Science and Technology*, pp. 519–542. London: Palgrave Macmillan UK.
- Dasgupta, P. and P. A. David (1994, September). Toward a New Economics of Science. *Research Policy* 23(5), 487–521.
- de Bartolome, C. A. M. (1990, February). Equilibrium and inefficiency in a community model with peer group effects. *Journal of Political Economy* 98(1), 110–133.
- Fershtman, C., K. Murphy, and Y. Weiss (1996, February). Social Status, Education and Growth. *Journal of Political Economy* 104(1), 108–132.
- Fershtman, C. and Y. Weiss (1993, July). Social Status, Culture and Economic Performance. *Economic Journal* 103(419), 946–59.
- Frank, R. (1984). Interdependent Preferences and the Competitive Wage Structure. *The RAND Journal of Economics* 15(4), 510–520.
- Frank, R. (1986). *Choosing the Right Pond: Human Behavior and the Quest for Status*. Oxford University Press.
- Gola, P. (2015). Occupational Sorting and the Structure of Status. University of Oxford. PDF file.
- Gola, P. (2021). Supply and Demand in a Two-Sector Matching Model. *Journal of Political Economy* 129(3), 940–978.
- Heckman, J. and G. Sedlacek (1985). Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-selection in the Labor Market. *Journal of Political Economy* 93(6), 1077–1125.
- Heckman, J. J. and B. E. Honore (1990). The empirical content of the roymodel. *Econometrica* 58(5), 1121–1149.
- Hill, R. and C. Stein (2021). Race to the Bottom: Competition and Quality in Science. Working Paper.
- Hopkins, E. and T. Kornienko (2004, September). Running to Keep in the Same Place: Consumer Choice as a Game of Status. *American Economic Review* 94(4), 1085–1107.
- Huberman, B. A., C. H. Loch, and A. Onculer (2004). Status as a Valued Resource. *Social Psychology Quarterly* 67(1), pp. 103–114.
- Ioannidis, J. P., K. W. Boyack, and J. Baas (2020). Updated science-

- wide author databases of standardized citation indicators. *PLoS Biology* 18(10), e3000918.
- Jeon, D.-S. and D. Menicucci (2008, June). Money, fame and the allocation of talent: Brain drain and the institution of science. *Journal of Economic Behavior & Organization* 66(3-4), 558–581.
- Lucifora, C. and D. Meurs (2006). The public sector pay gap in france, great britain and italy. *Review of Income and wealth* 52(1), 43–59.
- Luttmer, E. F. P. (2005). Neighbors as Negatives: Relative Earnings and Well-Being. *The Quarterly Journal of Economics* 120(3), pp. 963–1002.
- Machin, S. and A. Oswald (2000). UK Economics and the Future Supply of Academic Economists. *The Economic Journal* 110(464), 334–349.
- Mani, A. and C. H. Mullin (2004, October). Choosing the Right Pond: Social Approval and Occupational Choice. *Journal of Labor Economics* 22(4), 835–862.
- Merton, R. (1973). The Normative Structure of Science. In *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press.
- Morgan, J., D. Sisak, and F. Várdy (2018). The ponds dilemma. *The Economic Journal* 128(611), 1634–1682.
- Perez-Truglia, R. (2020). The effects of income transparency on well-being: Evidence from a natural experiment. *American Economic Review* 110(4), 1019–54.
- Phillips, S. H. (2004). *Review of the honours system*. Cabinet Office Downing Street.
- Precup, R. (2018). *Ordinary Differential Equations: Example-driven, Including Maple Code*. Walter de Gruyter GmbH & Co KG.
- Roy, A. D. (1951). Some Thoughts On The Distribution Of Earnings. *Oxford Economic Papers* 3(2), 135–146.
- Singh-Manoux, A., N. E. Adler, and M. G. Marmot (2003). Subjective social status: its determinants and its association with measures of ill-health in the whitehall ii study. *Social Science and Medicine* 56(6), 1321–1333.
- Singh-Manoux, A., M. G. Marmot, and N. E. Adler (2005). Does subjective social status predict health and change in health status better than objective status? *Psychosomatic medicine* 67(6), 855–861.
- Smith, A. (1776). *The Wealth of Nations*. Mobile Reference, 2010. Kindle file.
- Statistics Norway (1997-2018). Microdata: Wage statistics module (Lønn). (accessed January 18th 2020).
- Stephan, P. E. (1996). The economics of science. *Journal of Economic Literature* 34(3), 1199–1235.

Stern, S. (2004). Do scientists pay to be scientists? *Management science* 50(6), 835–853.

Weiss, Y. and C. Fershtman (1992, April). On the Stability of Occupational Rankings. *Rationality and Society* 4(2), 221–233.

B Robustness Checks

In this section, I relax the various simplifying assumptions from Section 3 one by one, in order to examine how critical each of them is for the results. I find that the main message of the paper (“the impact of each component of status on sorting is limited, but their joint impact is not”) is robust.

B.1 Asymmetric Taste for Prestige

In Section 3, I assumed that the taste for occupational prestige is the same in the two occupations. Suppose, instead, that the taste for prestige differs between academia and finance and, possibly, depends on the size of each sector: $k_A(M_A) \neq k_F(M_F)$. By Equation (1), the difference in occupational prestige rewards between academia and finance would become $(k_F(M_F)M_A + k_A(M_A)M_F)\frac{\alpha_A}{M_F}$. Thus a non-degenerate equilibrium of the altered model, can be supported in my model with $k = k_F(M_F)M_A + k_A(M_A)M_F$, with all other primitives of the model unchanged.

B.2 Alternative Specification of Occupational Prestige

In this section, I explore an alternative specification of occupational prestige, while retaining the assumption that the average prestige reward in the population is equal to 0. Specifically, consider any random variable $X_R \in [a_x, b_x]$ that has a strictly increasing and continuously differentiable distribution $Z : [a_x, b_x] \rightarrow [0, 1]$; for example, X_R could be a combination of characteristics that the society finds commendable: intelligence, creativity, courage, honesty, etc. Denote the joint distribution of the financial skill, the academic skill and the prestige characteristic by J , with $J(x_A, x_F, b_x) = H(x_A, x_F) > H_M(x_F)H_M(x_A)$. I impose no restrictions on J other than those inherited from H and Z .

Define the occupational prestige in academia and finance as follows:

$$o_F \equiv \frac{\bar{x}_R^F}{E(x_R)} - 1 = \frac{E(x_R|X_F > \psi(X_A))}{E(x_R)} - 1 \quad (\text{OA.1})$$

$$o_A \equiv \frac{\bar{x}_R^A}{E(x_R)} - 1 = \frac{E(x_R|X_F < \psi(X_A))}{E(x_R)} - 1. \quad (\text{OA.2})$$

This functional form ensures that, as in the baseline, the average occupational prestige reward in the population is equal to 0, and hence an increase in the taste for prestige affects welfare only through sorting.

The main message of this article holds as long as

$$J(x_A, b_x, x_R) > H_M(x_A)Z(x_R)$$

for all $(x_A, x_R) \in (a_x, b_x)^2$, that is, as long as the academic skill is positively interdependent with the prestige characteristics. To be more specific, let me discuss each of the main results separately. Trivially, if only occupational prestige matters, then there still must exist a single cutoff of financial skill such that all workers with $x_F > \psi^p$ join finance; this, together with $H(x_A, x_F) > H_M(x_F)H_M(x_A)$, ensures that $\bar{x}_A^A < \bar{x}_F^F$ in any equilibrium. It thus follows that occupational prestige cannot, on its own, cause academia to attract workers of (on average) higher skill than finance does.²² Theorem 2 is obviously completely unaffected, as it describes what happens if workers do not care about occupational prestige.

While I was unable to prove that Theorem 3 carries over unchanged in general, it is very easy to show a result with the same message. Namely, for any $M'_A \in (0, 1)$ and any $l_F \geq 0$, there must exist some $(\delta, k) \in \mathbf{R}_{\geq 0}^2$ for which *there exists* an equilibrium in which academia is large ($M_A \geq M'_A$), is more prestigious than finance ($o_A > o_F$) and attracts workers of higher skill than finance ($\bar{x}_A^A > \bar{x}_F^F$).²³ This result is weaker than Theorem 3 in two ways. First, as it does not guarantee that *all* non-degenerate equilibria

²²If we were to further assume that $J(b_x, x_F, x_R) > H_M(x_F)Z(x_R)$, then an increase in k would decrease the size of academia and Theorem 1 would carry over in its entirety.

²³It should be clear from the proof of Lemma 4(iii), that an analogous result holds also in this alternative specification. That is, for any fixed M , if we set δ high enough, then $o_A > o_F$ and $\bar{x}_A^A > \bar{x}_F^F$ in $\psi^c(\cdot; M_A)$. If, in addition, we set δ to some value greater than $w_A - w_F(0)$, then for $k = 0$ it must be the case that $c(M_A) > 0$. However, $c(M_A)$ increases linearly in k because $o_A > o_F$, and thus we can always find some $k > 0$ for which $c(M'_A) = 0$.

will have the property that $M_A \geq M'_A$, $o_A > o_F$, and $\bar{x}_A^A > \bar{x}_F^F$.

Finally, it is worth explaining why the assumption $J(x_A, b_x, x_R) > H_M(x_A)Z(x_R)$ plays a critical role. The positive interdependence between the academic skill and the prestige characteristics *guarantees* that if academia attracts mostly workers who are highly skilled academics, then academia is prestigious. If this assumption is violated, it might be impossible for academia to both be prestigious and attract workers of (on average) higher skill than finance; and if academia is not prestigious, then it might be impossible for academia to both be larger than finance and attract workers of higher skill than finance (by Theorem 2).

B.3 Non-Flat Wages in Academia

Suppose that $w_A(x_A) = w_A + g * f(x_A)$, where $f'(x_A) > 0$. In the main body, I assume that $g = 0$, which plays a role similar to the requirement that $w_A < w_A^*$ in Theorem 2(ii): My results imply that, on its own, neither component of social status is able to counter the impact of flat academic wages *if wages in academia are sufficiently flat compared to finance*. This conclusion is continuous in g : There exists some $g^* > 0$ such that Theorems 1 and 2 hold for all $g < g^*$. Theorem 3 holds, of course, for any finite g .

B.4 Endogenous Wages

In Section 3, I have effectively assumed that the marginal product of worker (x_A, x_F) is an exogenous function of her occupation-specific skill. Alternatively, one could follow Heckman and Sedlacek (1985) and assume that the marginal product depends on sorting. In this section, I briefly explain why allowing for endogenous marginal product (and thus also wages) would leave Theorems 1, 2 and 3 unchanged.

To be specific, define the functions

$$T_i(\psi) \equiv M_i(\psi) \int_{a_x}^{b_x} m_i(x) dG_i(x; \psi),$$

where $i \in \{A, F\}$. Suppose that the marginal product of worker (x_A, x_F) in occupation i under sorting ψ is equal to $p_i(T_i(\psi))m_i(x_i)$, where $p_i : [0, \bar{T}_i] \rightarrow \mathbf{R}_{>0}$ is decreasing and continuous and $\bar{T}_i = \int_{a_x}^{b_x} m_i(x) dx$. As a result, the

wage of worker (x_A, x_F) in finance is $w_F(x_F; T_F) = p_F(T_F)m_F(x_F)$, and her wage in academia is $w_A(x_A; T_A) = p_A(T_A)w_A$. Finally, to ensure that there exists an equilibrium in the no-status benchmark, let us assume that $w_A \in (\frac{p_F(\bar{T}_F)}{p_A(0)}m_F(a_x), \frac{p_F(0)}{p_A(\bar{T}_A)}m_F(b_x))$.

Let us define a new concept, of a *twice-compensated equilibrium*, and redefine the concepts of a compensated equilibrium and an equilibrium for the context of the model with endogenous wages.

Definition 2. A sorting ψ^{cp} constitutes a *twice-compensated equilibrium* if and only if (a) ψ^{cp} is non-degenerate and (b) there exist some *compensating differential* $c_d \in \mathbf{R}$ and a *compensating price* $c_p \in \mathbf{R}_{>0}$ such that for all $(x_A, x_F) \in [a_x, b_x]^2$,

$$\begin{aligned} \psi^{cp}(x_A) > x_F &\Rightarrow s_A(x_A; \psi^c) + c_d > c_p m_F(x_F) + s_F(x_F; \psi^c), \\ \psi^{cp}(x_A) < x_F &\Rightarrow s_A(x_A; \psi^c) + c_d < c_p m_F(x_F) + s_F(x_F; \psi^c), \end{aligned} \quad (\text{OA.3})$$

where s_i denotes the local status function (defined in Equation (2)). A sorting ψ^c constitutes a compensated equilibrium if it constitutes a twice-compensated equilibrium with $c_p = p_F(T_F(\psi^c))$. A sorting ψ^e constitutes an equilibrium if it constitutes a compensated equilibrium with $c_d = w_A p_A(T_A(\psi^e)) + k(o_A(\psi^e) - o_F(\psi^e))$.

A sorting ψ^c can constitute a compensated equilibrium only if it constitutes a twice-compensated equilibrium for some $c_p \in [p_F(0), p_F(\bar{T}_i)]$. The crucial insight is that the set of twice-compensated equilibria that correspond to $c_p \in [p_F(0), p_F(\bar{T}_i)]$ is the same as the union over $c_p \in [p_F(0), p_F(\bar{T}_i)]$ of the sets of compensated equilibria of the *baseline model* that correspond to specifications in which $w_F(x_F) = c_p m_F(x_F)$. Because the compensated equilibrium of the baseline model is continuous in c_p , the crucial results derived for the compensated equilibria of the baseline model (specifically, the discussion in Section 4.1, Lemma 1, and Lemma 4) have exact analogues if wages are endogenous. To understand the intuition behind this, consider Lemma 4(ii) as an example. The result from Section 4.3 implies that regardless of the extent to which wages in finance differ with skill, we can always make local status so important in academia that academia is more prestigious than finance. As wages in finance differ with skill the most if $c_p = p_F(\bar{T}_i)$, it follows that if $\delta > \delta^*(p_F(\bar{T}_i))$ then academia

must be more prestigious than finance in the compensated equilibrium of the model with endogenous wages. Given this insight, it is very easy to show that Theorem 1, Theorem 2 and Theorem 3 remain unchanged if wages are endogenous.²⁴

B.5 Skill Interdependence

The assumption that $H(x_A, x_F) > H_M(x_A)H_M(x_F)$ is natural in the context of sorting into academia and finance, as both occupations rely heavily on cognitive skills. If this assumption is violated, then it may well be impossible to any occupation to be both larger and attract workers of higher skill on average, no matter how differentiated the rewards are. To see why, consider the no-status baseline and assume that $x_A = b_x - x_F$, that is, that financial and academic skills are perfectly negatively correlated. In that case, the fact that only workers with financial skill $\geq \psi^b$ join finance implies that only workers with academic skill $\geq b_x - \psi^b$ join academia. This implies that $\bar{x}_F^F > \bar{x}_A^A$ if and only if $M_A > 0.5$. Thus if skills are perfectly negatively interdependent, then *even if* one occupation offers infinitely more-differentiated wages than the other occupation, it can attract workers of higher skill only if it is smaller than the other occupation.²⁵ It follows, therefore, that, at the very best, we can have $\bar{x}_A^A > \bar{x}_F^F$ only if $M_A < 0.5$.

C The Comparison Theorem

The following, well-known result plays a key role in many of the proofs in the paper.

Theorem OA.1 (Comparison Theorem). Let h and k be solutions of the differential equations

$$h'(x) = A(x, h(x)), \quad k'(x) = B(x, k(x))$$

²⁴In the case of Theorem 2(ii), the impossibility holds for $w_A \in \left(\frac{p_F(\bar{T}_F)}{p_A(0)}m_F(0), w_A^*\right)$.

²⁵If skills are imperfectly negatively interdependent, then it is possible for finance to attract workers of (on average) higher skill than academia and be the larger occupation, but not arbitrarily large.

respectively, where $A(x, y) \leq B(x, y)$ for $x \in [a, b]$ and A and B are Lipschitz-continuous in h and k , respectively. Let also $h(a) \leq k(a)$. Then $h(x) \leq k(x)$ for all $x \in (a, b]$. If, further, $A(x, h(x)) < B(x, h(x))$ or $h(a) < k(a)$, then $h(x) < k(x)$ for all $x \in (a, b]$.

Proof. It follows immediately from Theorem 8, Corollary 1 and Corollary 2 in [Birkhoff and Rota \(1969\)](#). \square

D Wage Dispersion in Academia and Finance

In this Appendix, I compare the empirical distributions of wages in academia and finance in Norway.

D.1 Data

I use the registry database on salaries (*lonnstatistikk*) maintained by Statistics Norway ([Statistics Norway, 2018](#)). I use the data from 2015 to 2018. For each individual, I see each of the jobs they held in that year (and its occupation code), their average monthly salary, and the number of hours they worked in that job (as a percentage of a full-time job). To compute the wage distributions in academia and finance, I sum the salaries that each individual received in all jobs that fall within academia or finance, and then calculate the full time equivalent of the sum of these salaries. I also remove the top and bottom promile of earners.²⁶ Table [OA.1](#) lists the occupation codes and titles included in my definition of academia and finance.

D.2 Results

The first two columns of Table [OA.2](#) report summary statistics for the distribution of wages in academia and finance in Norway in 2018. The distribution in finance is much more dispersed than in academia: The 90/10 wage ratio is around 25% higher in finance than in academia (2.41 in finance compared to 1.957 in academia). The standard deviation of wages of

²⁶The results are unchanged qualitatively if those are included, but the variance of wages in finance explodes to implausible levels in that case.

Occupation	ISCO-08 Code	Description
Finance	2412	Financial and investment advisers
	2413	Financial analysts
	3311	Securities and finance dealers and brokers
	3312	Credit and loans officers
Academia	2310	University and higher education teachers

Table OA.1: Classification into Finance and Academia

Note: This table lists the ISCO-08 4-digits codes that I used to define finance and academia, respectively.

wages is about 1.8 as high in academia as in finance, whereas the standard deviation of log wages is over 30 % higher in finance than in academia.

Of course, one may worry that the distribution of wages for any particular year provides only limited information about the steepness of the wage schedule in that occupation. For example, wages in finance could be very similar across individuals over long periods of time, but be subjected to idiosyncratic shocks that blow-up the within-year variance. To alleviate these concerns, I also compute a multi-year average salary for all workers who worked in academia/finance in each year from 2015 to 2018.²⁷ The results for this exercise are reported the last two columns of Table OA.2. Reassuringly, the difference in wage dispersion between finance and academia increases: the 90/10 ratio is now over 35% larger in finance, and the standard deviation of wages is more than twice as high in finance than in academia.

D.3 Wage Distribution vs. Wage Schedules

One obvious problem with the results above is that they compare wage dispersion *conditional* on selection, rather than the exogenous wage schedules which are the primitive of my model. However, if we are willing to assume a standard log-normal Roy's model, then our information on conditional wage dispersion and the size of each occupation suffices to conclude that the unconditional wage schedule must be steeper in finance than in academia, and that—in the absence of social status concerns—selection into finance must be more positive than into academia.

²⁷I stop at 2015, as the occupational codes have changed in that year.

	Finance 2018	Academia 2018	Finance 2015-18	Academia 2015-18
1%	19282.03	19900	27881	30924.81
10%	34790.67	34156.88	35215.06	39233.75
50%	50000	48080.86	48477.26	50223.52
90%	83961.8	66850	81805.25	67321.42
99%	162375	100491.8	163541.7	92800.5
90/50	2.413	1.957	2.323	1.716
Mean	56642.27	49523.64	55783.98	52329.7
Mean(log)	10.856	10.761	10.851	10.84
Sd	28696.83	16119.89	29393.44	12879.44
Sd(log)	.424	.325	.362	.221
N	41432	31251	23621	17170

Table OA.2: Summary Statistics for the Wage Distributions in Finance and Academia

Note: Own calculations using the registry database maintained by Statistics Norway; the amounts reported are the average monthly salaries in norwegian kroners. The second column summarizes the distribution in finance in 2018, the third summarizes the distribution in academia in 2018, the fourth summarizes the distribution in finance for years 2015-2018, and the fifth summarizes the distribution in academia for years 2015-2018.

To see this, consider a Roy's model with normally distributed wages, in which each worker draws a vector (w_A, w_F) of academic and financial wages, respectively, and then joins the occupation that offers them a higher wage. The joint distribution of (w_A, w_F) is normal, with means μ_A, μ_F , variances σ_A, σ_F and covariance σ_{AF} . In this model, the measure of workers joining occupation i is simply:

$$M_i = \Phi(c_i)$$

where Φ denotes the cdf of a standard normal distribution, $c_i = (\mu_i - \mu_j) / \sigma$ and $\sigma = \sqrt{\text{Var}(w_A - w_F)}$.

As academia and finance employ similar number of workers (e.g., between 2015 and 2018, 23621 workers remained in finance and 17170 workers remained in academia, see Table OA.2, which would translate into $M_A = 0.43$ in the notation from the model section), I will assume that $c_i \approx 0$. Denote the conditional variance of log-wages in occupation i by

$$\tilde{\sigma}_i \equiv \text{Var}(w_i | w_i > w_j).$$

It follows then from Equation (5) in Heckman and Sedlacek (1985) (after

some rearranging) that if $c \approx 0$, then

$$\sigma_A - \sigma_F \approx \frac{\tilde{\sigma}_A - \tilde{\sigma}_F}{1 - \frac{2}{\pi}} < 0.$$

Finally, from Equation (4) in [Heckman and Sedlacek \(1985\)](#) it follows that

$$\tilde{\mu}_i \equiv \mathbb{E}(w_i | w_i > w_j) = \mu_i + \frac{(\sigma_i - \sigma_{AF})}{\sigma \Phi(c_i)} \frac{1}{\sqrt{2\pi}} e^{-\frac{c_i^2}{2}},$$

which implies that if $c_i \approx 0$ then

$$(\tilde{\mu}_A - \mu_A) - (\tilde{\mu}_F - \mu_F) \approx \frac{2}{\sqrt{2\pi}} (\sigma_A - \sigma_F) < 0.$$

Therefore, if academia is of similar size as finance and the observed wage dispersion is lower in academia, then selection into academia must be worse than selection into finance in a Roy's model with normally distributed wages.