On the Importance of Social Status for Occupational Sorting*

Paweł Gola†

7th September 2023

Abstract

Models of self-selection predict that occupations with flat wage schedules attract workers of lower average ability. However, in certain prominent occupations such as academia and the civil service, wages are flat yet 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 allows flat-wage occupations to attract predominantly 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 occupations paying flat wages 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 the editor Amanda Friedenberg and three anonymous referees for their insightful comments, as well as Charles Brendon, Peter Dolton, Peter Eso, Dawei Fang, Sanjeev Goyal, Claudia Herresthal, Ed Hopkins, Ian Jewitt, Coen Teulings, Adrien Vigier, Yuejun Zhao, and seminar participants in Brighton, Cambridge, Gothenburg, Helsinki, Huddersfield, Luxembourg, Prague, Oslo, Oxford, Trondheim, Vienna, Warsaw, and Xiamen, for helpful discussions 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 relative position among their peers (Huberman, Loch, and Onculer, 2004; Luttmer, 2005; Card, Mas, Moretti, and Saez, 2012; Perez-Truglia, 2020). In particular, individuals are prepared to forego substantial pecuniary benefits (Cardoso, 2012; Bursztyn, Ferman, Fiorin, Kanz, and Rao, 2018; Bottan and Perez-Truglia, 2020) and undertake significant risks (Ager, Bursztyn, Leucht, and Voth, 2021) to achieve a higher rank. These findings suggest that relative concerns may influence occupational choices as well; after all, it is considerably easier to attain high relative standing if one joins a profession where the competitors are less accomplished. How profoundly, however, can this pursuit of rank affect equilibrium sorting patterns?

Consider the phenomenon of highly talented workers sorting into occupations that provide significantly flatter wage schedules than competing professions. Examples include academia (Hamermesh, 2018; Machin and Oswald, 2000), the civil service (Lucifora and Meurs, 2006), and the officer corps (Asch and Warner, 2001). The literature on self-selection famously posits that jobs with flatter wage schedules should attract workers of lower skill than competing occupations offering steeper schedules (Roy, 1951; Heckman and Sedlacek, 1985; Borjas, 1987; Heckman and Honore, 1990). However, academia (Stern, 2004; Bó, Finan, Folke, Persson, and Rickne, 2017) and the civil service (Lucifora and Meurs, 2006) are typically chosen by individuals who are at least as talented as workers in other professional occupations.[1] This discrepancy creates a puzzle: why are highly skilled individuals willing to join occupations with flatter wage schedules?

Note that the answer to this puzzle remains of interest irrespective of whether one believes these flat wage schedules accurately reflect differences in marginal product. If they do, then the positive selection into academia and civil service diverts talent from industries where it could be more productive. If they don't—for reasons ranging from the political costs of spending public funds on high salaries to the public good problems inherent to some of these professions—then these occupations might warrant even

[1] A comprehensive review of wage schedules, quality of selection, and the role of rank in academia, civil service and the officer corps is provided in Section 5.2.

stronger positive selection, given that each of them provides services that are of critical importance for the society: Academia creates the foundations for future technological progress, the civil service makes governance possible, and the army ensures security. In either case, understanding the driving force behind this positive selection is essential.

A potential clue lies in the high importance of relative standings and the ease of assessing an individual's rank within each of academia, the civil service and the military. For instance, most academics readily disclose their publication and citation records, high-ranked British civil servants are honored with aristocratic titles, and in the military one's rank is literally worn on one's sleeve. Given that more accessible information about relative standings leads to greater differences in happiness between high- and low-rank individuals (Perez-Truglia, 2020), the prominence of rank-concerns in these professions could indeed provide the additional differentiation of rewards that attracts high-quality workers, despite the seemingly discouraging wage structure.

In this paper, I examine the circumstances under which relative concerns can explain the positive selection into occupations offering flat wage schedules. 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 two occupations. Each worker is endowed with a two-dimensional vector of occupation-specific skills. The first of the two occupations pays all workers the same, flat wage; the wage in the other occupation is an increasing function of the occupation-specific skill. I will label the flat-wage occupation as 'academia' and the varying-wage occupation as 'finance'—however, they could equally well represent any two occupations or sectors that substantially differ in how much wages vary with skill (e.g., public and private sectors). As academia pays wages that do not vary with skill, in the no-status benchmark (i.e., if wages were the workers' only reward) all workers with high enough financial skill join finance, and academia attracts exclusively workers with low financial skill. This implies that academics on average have low academic skill as well, because I assume that the academic and financial skills are positively correlated.

I depart from Roy's setting by assuming that, apart from wages, workers also care about their social status, which consists of two components: *occu-*

pational prestige (i.e., between-occupation relative concerns) and *local status* (i.e., within-occupation relative concerns), each determined endogenously.[2] 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 assume that occupational prestige is determined by the average skill in the two professions and thus enters the payoff function in each occupation as an endogenous constant.[3] Because of that, the payoff in academia remains constant with respect to skill, and thus academia continues to attract workers with low financial and academic skill, exactly as in the no-status benchmark. On top of that, if workers care only about wages and prestige, then academia attracts fewer workers than in the benchmark: The flat wage schedule in academia forces most high-skilled workers to become bankers, making finance the prestigious occupation—this, in turn, attracts a larger number of workers to finance than the wage level itself would warrant.[4]

Next, I consider the impact of local status only, which is modeled as an increasing 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. This increasing 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' payoffs 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.

[2]The evidence that within-occupation relative concerns influence people's behaviour has already been discussed. A separate strand of the literature (see, for example, Dolton, Makepeace, and van der Klaauw, 1989; Zhan, 2015) shows that workers are drawn to occupations with high prestige.

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

[4]This 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).

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 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 payoffs 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 several topics: the policy implications of my results; the evidence on wage structures and selection into academia, civil service, and the officer corps; the appeal (or lack thereof) of two alternative explanations for the motivating puzzle; and the case of exogenously given occupational prestige. Appendix A contains the proofs of all propositions and lemmas. Online Appendix B discusses why none of the simplifying assumptions are critical.

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.[5] 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 payoff 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 payoff 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

[5] Albornoz, 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.

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.[6] In those papers workers choose between 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 payoff 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 payoff 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

[6] The 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.

fame payoff 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 incentivises 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 $(x_A, x_F) \in [x_l, x_h]^2$, where x_A and x_F are the skills used in academia and finance, respectively. The joint distribution of (x_A, x_F) in the population is denoted by $H : [x_l, x_h]^2 \to [0, 1]$. H is (1) twice continuously differentiable, (2) has a strictly positive, finite density in its support and (3) is symmetric, that is, $H(x, y) = H(y, x)$. Symmetry implies, among other, that both skills have the same marginal distribution; the cdf of the marginal distribution is denoted as H_M and its pdf as h_M. Finally, (4) the joint distribution of skill is more concordant than an independent distribution, that is $H(x_A, x_F) \geq H_M(x_A)H_M(x_F)$ for all $(x_A, x_F) \in [x_l, x_h]^2$. Together with the other assumptions this implies that there is positive (but not perfect) correlation between the two skills.

A *sorting* $\sigma : [x_l, x_h]^2 \to \{A, F\}$ describes the occupational choice of all workers. Every sorting induces the set of types $\sigma^{-1}(\{i\})$ that are sorted into occupation $i \in \{A, F\}$. A sorting is non-degenerate if $\sigma^{-1}(\{A\}), \sigma^{-1}(\{F\})$ both have a positive measure. Given a non-degenerate sorting δ, the size of occupation i is denoted by

$$M_i(\sigma) \equiv \int_{\sigma^{-1}(\{i\})} \frac{\partial^2}{\partial x_A \partial x_F} H(r, s)\, \mathrm{d}s\, \mathrm{d}r, \tag{1}$$

the distribution $G_i : [x_l, x_h] \to [0, 1]$ of skill x_i among the workers sorted into occupation i is denoted by

$$G_i(x; \sigma) \equiv \frac{1}{M_i(\sigma)} \int_{\{\mathbf{x} \in \sigma^{-1}(\{i\}):\, x_i \leq x\}} \frac{\partial^2}{\partial x_A \partial x_F} H(r, s)\, \mathrm{d}s\, \mathrm{d}r,$$

the lowest level of skill x_i among workers sorted into occupation i is denoted by $x_i^m(\sigma) \equiv \inf_{x_i}\{x_i : \exists_{x_j}(x_i, x_j) \in \sigma^{-1}(\{i\})\}$ (where $j \neq i$) and the highest level of skill x_i for which nevertheless some workers join occupation $j \neq i$ is denoted by $x_i^s(\sigma) \equiv \sup_{x_i}\{x_i : \exists_{x_j}(x_i, x_j) \in \sigma^{-1}(\{j\})\}$. Workers sorted into academia will be called academics, and workers sorted into finance will be called bankers.

I will now introduce the three components of payoffs, and then define the total payoff 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

$$o_A(\sigma) = \frac{\bar{x}_A^A(\sigma) - \bar{x}_F^F(\sigma)}{M_A(\sigma)} \quad \text{and} \quad o_F(\sigma) = \frac{\bar{x}_F^F(\sigma) - \bar{x}_A^A(\sigma)}{M_F(\sigma)}, \tag{2}$$

where $\bar{x}_i^i(\sigma) \equiv \int_{x_l}^{x_h} x \mathrm{d}G_i(x;\sigma)$ denotes the average x_i among workers sorted into 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 under sorting σ is

$$s(G_i(x_i;\sigma)). \tag{3}$$

The function $s : [0,1] \to \mathbb{R}$ is (a) strictly increasing, with $s' > 0$ and (b) satisfies $\int_0^1 s(r)\,\mathrm{d}r = 0$ so that local status is zero-sum.[7] Note that these

[7]That is, the average local status within an occupation always equals zero: $\int_{x_l}^{x_h} s\left(G_i(x_i;\sigma)\right) \mathrm{d}G_i(x_i;\sigma) = 0$.

two assumptions imply that $s(0) < 0 < s(1)$.

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

Payoffs and (Compensated) Equilibrium Given a sorting σ, the *payoff* of an agent (x_A, x_F) from joining occupation $i \in \{A, F\}$ is a weighted sum of her wage, the prestige of occupation i, and her local status within it:

$$\pi_i(x_i; \sigma) = w_i(x_i) + l_i s(G_i(x_i; \sigma)) + k o_i(\sigma), \tag{4}$$

where $l_i \geq 0$ is the importance of local status 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) $\delta \equiv l_A - l_F$.

Entry into each occupation is free, and every worker joins the occupation which maximizes her payoff. However, 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 σ_c constitutes a *compensated equilibrium* if and only if (a) σ_c is non-degenerate and (b) there exists some *compensating differential* $c \in \mathbf{R}$ such that for all $(x_A, x_F) \in [x_l, x_h]^2$

$$\begin{aligned} (x_A, x_F) \in \sigma_c^{-1}(\{A\}) &\Rightarrow \pi_A(x_A; \sigma_c) + c > \pi_F(x_F; \sigma_c), \\ (x_A, x_F) \in \sigma_c^{-1}(\{F\}) &\Rightarrow \pi_A(x_A; \sigma_c) + c < \pi_F(x_F; \sigma_c). \end{aligned} \tag{5}$$

Condition (5) 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 payoff in academia as the payoff she would receive in finance (and *vice versa*). Of course, compensated equilibria are closely related to the equilibria of this model.

Definition 2. A non-degenerate sorting σ_e constitutes an equilibrium if and only if it constitutes a compensated equilibrium for $c = 0$. The degenerate sorting $\sigma(x_A, x_F) = F$ $(\sigma(x_A, x_F) = M)$ for all (x_A, x_F) constitutes an equilibrium if there exists an $\epsilon > 0$ such that for all $m \in (0, \epsilon)$ there exists a compensated equilibrium σ_c such that $M_A(\sigma_c) = m$ $(M_F(\sigma_c) = m)$ and $c \geq 0$ $(c \leq 0)$.

A non-degenerate sorting constitutes an equilibrium if all workers join the occupation that maximises their payoff, taking the sorting decisions of all other workers as given. As social status payoffs are only defined for non-degenerate sortings, we say that a degenerate sorting constitutes an equilibrium if, and only if, in all compensated equilibria in which a small number of workers joins the currently empty occupation i, workers in that occupation receive a positive compensating differential.

Note that 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 σ_c 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 payoffs as a constant, and thus acts 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 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 payoffs to zero, that is, they ensure that $M_A o_A + M_F o_F = 0$.[8] This implies that

[8]They 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 $(\bar{x}_i^i - \bar{x}_j^j)k$. Clearly, the sum of the two workers payoffs is 0 on a meeting level; however, workers from the smaller occupation will participate in a larger number of

(a) any change in sorting leaves the sum of occupational prestige payoffs 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.[9] Online Appendix B.3 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. It is natural to assume that within-occupation ranking is based on the occupation-specific skill, as the esteem received from peers is likely to be strongly related to how well the agent performs her job. A common alternative is to assume that the ranking depends on income (Hopkins and Kornienko, 2004). This is equivalent to my assumption if $w'(x_A) > 0$; my main results are robust to setting $w'(x_A)$ to be strictly positive (but small).

The assumption that local status is zero-sum is both plausible and common in the literature (see, for example, the discussion preceeding Proposition 5 in Damiano et al., 2010), but it is critical for the results.

between-occupation meetings, and thus occupational prestige enters their payoff function with weight k/M_i.

[9] The fact that occupational prestige depends on the difference between average skill levels emerges naturally in a micro-foundation where social status is based on purely ordinal rank comparisons (see Appendix A in Gola, 2015). Hence, the functional forms imposed on occupational prestige and on local status are completely consistent with each other, even though one appears to be cardinal and the other ordinal.

Specifically, in order for Theorems 2(i) and 3 to hold, one needs $s(0) < 0$, which is implied by local status being zero-sum. A detailed discussion of the importance of $s(0)$ being negative can be found in Online Appendix B.2.

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.[10] 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.4) and (b) endogenous wage functions in both occupations (OA B.5). I also explain that if skills are strongly negatively interdependent, then the smaller occupation always attracts better workers, regardless of how steep the payoff schedules are (OA B.6).

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 payoffs, that is, if $\pi_i(x_i) = w_i(x_i)$. The equilibrium is trivial: There exists a single cutoff value $\psi^b = w_F^{-1}(w_A)$ such that all workers with $x_F > \psi^b$

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

join finance and all workers with $x_F < \psi^b$ join academia. This is because payoffs 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 $H(x_A, x_F) \geq H_M(x_A)H_M(x_F)$, it follows that $G_A(x_A; \sigma_e) = H(x_A, \psi^b)/H_M(\psi^b) \geq H_M(x_A)$. This implies that $\bar{x}_A^A(\sigma_e)$ is weakly lower than the population-wide average skill, $\bar{x} \equiv \int_{x_l}^{x_h} x h_M(x) \mathrm{d}x$. As finance attracts only workers with $x_F > \psi^b$, it must be that $\bar{x} < \bar{x}_F^F(\sigma_e)$, which means that, on average, academia attracts less skilled workers than finance does.

4.2 The Prestige-Only Case

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 $\pi_i(x_i; \sigma) = w_i(x_i) + k o_i(\sigma)$.

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

Proof. Because $k o_i$ does not depend on the worker's type, any compensated equilibrium must be characterised by a single cutoff ψ^p, such that all workers with $x_F > \psi^p$ join finance and all workers with $x_F < \psi^p$ join academia. Again, because $H(x_A, x_F) \geq H_A(x_A)H_F(x_F)$ in any compensated equilibrium σ_c we have that $\bar{x}_A^A(\sigma_c) \leq \bar{x} < \bar{x}_F^F(\sigma_c)$, which implies that the same holds for any non-degenerate equilibrium σ_e. From this follows immediately that $c = w_F(\psi^p) - w_A + k\frac{\bar{x}_F^F(\sigma_c) - \bar{x}_A^A(\sigma_c)}{H_M(\psi^p)(1 - H_M(\psi^p))}$, so that $c > 0$ if $\psi^p \geq \psi^b$, which implies that $M_A(\sigma_e) < H_M(\psi^b)$ 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 $\psi^p = 0$. □

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, payoffs 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 payoffs 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 The Local-Status-Only Case

In this section, I consider what happens if workers care about local status but not about occupational prestige, in which case payoffs are given by $\pi_A(x_A; \sigma) = w_A + (l_F + \delta)s(G_A(x_A; \sigma))$ and $\pi_F(x_F; \sigma) = w_F(x_F) + l_F s(G_F(x_F; \sigma))$. 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 equilibrium sorting.

4.3.1 Characterizing the Equilibrium

If local status matters, then the payoff in academia increases in x_A in any sorting σ: Workers with higher academic skill are more willing to join academia than the less skilled ones, and the cutoff value $\psi(x_A)$ of the financial skill is non-decreasing in the academic skill. Thus, any compensated equilibrium sorting σ_c can be characterised by a *separation function* $\psi_{\sigma_c} : [x_l, x_h] \to [x_l, x_h]$ such that (1) $(x_A, x_F) \in {\sigma_c}^{-1}(\{A\})$ if and only if $\psi_{\sigma_c}(x_A) > x_F$ and (2) $(x_A, x_F) \in \sigma_c^{-1}(\{F\})$ if and only if $\psi_{\sigma_c}(x_A) < x_F$.[11] Figure 1 depicts how a separation function determines the sorting of workers into occupations.

Suppose that $w_A \in (w_F(x_l) - \delta s(0), w_F(x_h) - \delta s(0))$.[12] Because the payoff functions $\pi_F(\cdot), \pi_A(\cdot)$ are continuous in skill, it follows from Condition (5) and the definition of an equilibrium that a sorting σ_e constitutes an

[11] Workers for whom $x_F = \psi_{\sigma_c}(x_A)$ are of measure zero, and can thus be ignored without loss of generality.

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

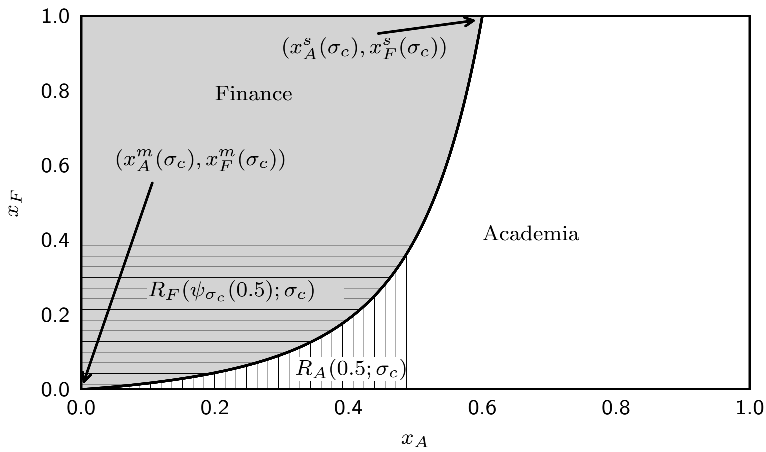

Figure 1: The Separation Function and Sorting

Notes: The solid black curve depicts the separation function ψ_{σ_c}. The white (light gray) area below (above) ψ_{σ_c} depicts the space of workers that join academia (finance). The vertically (horizontally) hatched area represents the space of workers with skill $x_A \leq 0.5$ ($x_F \leq \psi_{\sigma_c}(0.5)$) who join academia (finance); their number depends on the number of workers who reside in this space, with $R_A(x_A; \sigma_c) + R_F(\psi_{\sigma_c}(x_A); \sigma_c) = H(x_A, \psi_{\sigma_c}(x_A))$.

equilibrium if and only if, for all $x_A \in [x^m_A(\sigma_e), x^s_A(\sigma_e)]$,

$$\pi_F(\psi_{\sigma_e}(x_A); \sigma_e) = \pi_A(x_A; \sigma_e). \tag{6}$$

Note that because $G_i(x^m_i(\sigma_e); \sigma_e) = 0$, Equation (6) implies $x^m_F(\sigma_e) = w_F^{-1}(w_A + \delta s(0)) > x_l$. Hence, as $\min\{x^m_A(\sigma), x^m_F(\sigma)\} = x_l$ it follows that $x^m_A(\sigma_e) = x_l$.

Equation (6) can be rewritten as a system of one differential and one algebraic equation. To see that, we need to make two observations. First, denote the product of a worker's rank in occupation i and the size of occupation i under sorting σ by $R_i(x_i; \sigma) \equiv M_i(\sigma) G_i(x_i; \sigma)$ and call this object *the cumulative mass function*. The derivative of the mass function will be denoted by r_i; in academia, the equilibrium r_i is given by:

$$r_A(x_A; \sigma_e) = h_M(x_A) \Pr(X_F < \psi_{\sigma_e}(x_A) | X_A = x_A) = \frac{\partial}{\partial x_A} H(x_A, \psi_{\sigma_e}(x_A)).$$

Second, because all workers join some occupation and the separation function is increasing, it follows that for any $(x_A, x_F) \in [x^m_A(\sigma_c), x^s_A(\sigma_c)]$

and any equilibrium separation function ψ_{σ_e}, we get

$$R_A(x_A;\sigma_e) + R_F(\psi_{\sigma_e}(x_A);\sigma_e) = H(x_A, \psi_{\sigma_e}(x_A)). \tag{7}$$

Jointly, these two observations and Equation (6) imply that on $[x_l, x_A^s(\sigma_e)]$:

$$r_A(x_A;\sigma_e) = F(x_A, R_A(x_A;\sigma_e); M_A(\sigma_e)) \quad \text{and} \quad R_A(x_l;\sigma_e) = 0, \tag{8}$$

where

$$F(x_A, R_A; M_A) \equiv \frac{\partial}{\partial x_A} H\left(x_A, Z(\frac{w_A}{l_F+\delta} + s(R_A/M_A), R_A, x_A; M_A)\right),$$

$Z(y, R_A, x_A; M_A)$ is the inverse with respect to x_F of the function

$$L(x_A, x_F, R_A; M_A) \equiv \frac{1}{l_F+\delta}\left(w_F(x_F) + l_F s\left(\frac{H(x_A,x_F) - R_A}{1-M_A}\right)\right)$$

and L is defined for $x_A, x_F \in [x_l, x_h]^2$, $\frac{H(x_A,x_F)-R_A}{1-M_A} \in [0,1]$.[13] However, $R_A(\cdot;\sigma_e)$ must clearly meet one further condition, which is

$$R_A^e(x_h;\sigma_e) = M_A(\sigma_e). \tag{9}$$

Note that for a given M_A, Equation (8) is an initial-value problem. Overall, to find the equilibrium, we need to solve the IVP defined by Equation (8) for a every $M_A \in (0,1)$, and then solve for $M_A(\sigma_e)$ using Equation (9).

Proposition 1. Suppose $k = 0$. (i) If $w_A \in (w_F(x_l) - \delta s(0), w_F(x_h) - \delta s(0))$, then there exists a unique non-degenerate equilibrium σ_e and the size of academia $M_A(\sigma_e)$ is increasing and continuous in w_A. (ii) If $w_A \notin (w_F(x_l) - \delta s(0), w_F(x_h) - \delta s(0))$ 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 is academia's size. In that case, an increase in the size of academia can itself provide the increase in payoff which is needed to sustain an equilibrium in which academia is

[13] Of course, $r_A(x_A;\sigma_e) = h_M(x_A)$ for $x_A > x_A^s(\sigma_e)$, by the definition of $x_A^s(\sigma_e)$.

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 then 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 $\sigma_c^{M_A}$. The separation function $\psi_{\sigma_c^{M_A}}$ 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 decrease in $\frac{\partial}{\partial x_F} L(x_A, x_F, R_A; M_A)$ for all admissible x_A, x_F, R_A, then it also causes an increase in $G_F(x_F; \sigma_c^{M_A})$ for all $x_F \in [x_l, x_h]$ (and strictly for some) and a decrease in $G_A(x_A; \sigma_c^{M_A})$ for all $x_A \in [x_l, x_h]$ (and strictly for some).

The function $\frac{\partial}{\partial x_F} L(x_A, x_F, R_A; M_A)$ captures the extent to which payoffs differ with skill in finance *relative* to the extent to which payoffs differ with skill in academia. If payoffs become less steep in finance (relative to academia), then the distribution of skill improves in academia and worsens in finance, both in the sense of first-order stochastic dominance. Intuitively, less differentiation in payoffs decreases the payoff of high-skilled workers and rewards low-skilled workers; the sizes of academia and finance 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

$$\begin{aligned} &\frac{\partial^2 L(x_A, x_F, R_A)}{\partial \delta \partial x_F} < 0, \\ &\frac{\partial^2 L(x_A, x_F, R_A)}{\partial l_F \partial x_F} < 0 \iff \delta < \frac{(1-M_A) w_F'(x_F)}{\frac{\partial}{\partial x_F} H(x_A, x_F) s'\left(\frac{H(x_A,x_F)-R_A}{1-M_A}\right)}. \end{aligned} \tag{10}$$

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 payoffs 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 $\sigma_c^{M_A}$ as $l_F \to \infty$. (i) If $M_A \geq (\leq)\ 0.5$, then $\lim_{l_F \to \infty} G_A(x; \sigma_c^{M_A}) \geq (\leq) \lim_{l_F \to \infty} G_F(x; \sigma_c^{M_A})$ for all $x \in [x_l, x_h]$. 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 $\bar{x}_A^A(\sigma_c^{M_A}) - \bar{x}_F^F(\sigma_c^{M_A}) > 0$.

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 payoffs 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. Consequently, 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 $\bar{x}_A^A(\sigma_c^{M_A}) - \bar{x}_F^F(\sigma_c^{M_A}) < 0$. (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 $\bar{x}_A^A(\sigma_c^{M_A}) - \bar{x}_F^F(\sigma_c^{M_A}) > d$.

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 payoffs 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 (10), 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 $\delta \geq \frac{w_A - w_F(x_l)}{-s(0)}$ ($\delta \leq \frac{w_A - w_F(x_h)}{-s(0)}$) then academia (finance) unravels in the unique equilibrium, so that $\psi_{\sigma_e}(x_A) = x_l$ ($\psi_{\sigma_e}(x_A) = x_h$).

If local status becomes very important in academia relative to finance ($\delta \geq w_A - w_F(x_l)$), then the lowest-ranked academic receives a lower payoff than a banker of skill x_l, 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.[14] 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 denote the set $(\frac{w_A - w_F(x_h)}{-s(0)}, \frac{w_A - w_F(x_l)}{-s(0)})$ by I_W. (i) There exists some $(\delta, l_F) \in I_W \times \mathbf{R}_{\geq 0}$ such that $\bar{x}_A^A(\sigma_e) - \bar{x}_F^F(\sigma_e) > 0$ and $M_A < 0.5$ in the unique equilibrium. (ii) However, if w_A is sufficiently close to $w_F(x_l)$, then there exists no $(\delta, l_F) \in I_W \times \mathbf{R}_{\geq 0}$ such that $\bar{x}_A^A(\sigma_e) > \bar{x}_F^F(\sigma_e)$ and $M_A(\sigma_e) \geq 0.5$ 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 $w_F(x_F^m(\sigma_e)) = w_A$ and thus $x_F^m(\sigma_e) = \psi^b$. Therefore, we have that $M_A(\sigma_e) \geq H_M(\psi^b)$ and thus $M_A(\sigma_e) > M'_A$. 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 $\sigma_c(M'_A)$ under wage level w_A, with $c = w'_A - w_A$. Because the size of academia in a compensated equilibrium is increasing in c by Proposition 1(i), the compensating differential for which $\sigma_c^{M_A}$ is a compensated equilibrium (denoted by $c(\sigma_c^{M_A})$) must increase in M_A. Thus, $M'_A < M_A(\sigma_e)$ and $c(\sigma_c^{M_A(\sigma_e)}) = 0$ imply that $c(\sigma_c^{M'_A}) < 0$. Second, by Lemmas 2 and 3(ii), if $l_F > l_F^*$ then $\bar{x}_A^A(\sigma_c^{M'_A})) > \bar{x}_F^F(\sigma_c^{M'_A})$. Finally, note that if $\delta = \frac{w_A - w_F(x_l)}{-s(0)}$ then $c(\sigma_c^{M'_A}) = w'_A - w_A > 0$. As $\psi_{\sigma_c^{M_A}}$ is continuous in δ, so must be $c(\sigma_c^{M'_A})$,

[14]The 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.

and thus there exists some $\delta' \in (0, \frac{w_A - w_F(x_l)}{-s(0)})$ for which $c(\sigma_c^{M'_A}) = 0$. Thus if $\delta = \delta'$, then $M_A(\sigma_e) = M'_A$ and $\bar{x}_A^A(\sigma_e) > \bar{x}_F^F(\sigma_e)$ (by Lemma 2).

(ii) Suppose that $\frac{w_A - w_F(x_l)}{-s(0)} \leq \min\{y, \frac{w_F(H_M^{-1}(0.5)) - w_F(x_l)}{2(s(1)-s(0))}\}$, so that $\delta < \frac{w_A - w_F(x_l)}{-s(0)}$ only if $\delta < \min\{y, \frac{w_F(H_M^{-1}(0.5)) - w_F(x_l)}{2(s(1)-s(0))}\}$. Suppose that $M_A(\sigma_e) \geq 0.5$. This implies that $x_F^s(\sigma_e) > H_M^{-1}(0.5)$, which yields

$$\begin{aligned} w_F(H_M^{-1}(0.5)) + s(0)l_F &< w_F(x_F^s(\sigma_e)) + l_F s(G_F(x_F^s(\sigma_e); \sigma_e)) \\ &= w_A + l_A s(G_A(x_A^s(\sigma_e); \sigma_e)) \leq w_A + l_A s(1). \end{aligned}$$

As $\frac{w_A - w_F(x_l)}{-s(0)} > \delta = l_A - l_F$, it follows that

$$\begin{aligned} \frac{w_F(H_M^{-1}(0.5)) - w_F(x_l)}{s(1) - s(0)} &< \frac{w_A - w_F(x_l) + (l_A - l_F)s(1)}{s(1) - s(0)} + l_F \\ &< \frac{(w_A - w_F(x_l))}{-s(0)} + l_F \\ &< \frac{w_F(H_M^{-1}(0.5)) - w_F(x_l)}{2(s(1) - s(0))} + l_F, \end{aligned}$$

which immediately implies that $l_F > \frac{w_F(H_M^{-1}(0.5)) - w_F(x_l)}{2(s(1)-s(0))} > \delta$. Hence, if $M_A(\sigma_e) \geq 0.5$ then $\delta < \min\{y, l_F\}$ and the result follows by Lemma 4(i). $\square$

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!

Let me stress that the punitive aspect of local status, that is the punishment inflicted on low-ranked academics, not only pushes academia to be small, but also plays a critical role in ensuring that academia attracts higher quality workers than finance. Intuitively, the punishment discourages workers who are low-skilled both in research and in banking from joining academia in equilibrium, and hence brings down the average skill in finance.[15] If, instead, local status only provided payoff differentiation but did not punish low-rank workers—that is, if $s(0) = 0$—then all workers with financial skill lower than $w_F^{-1}(w_A)$ would always join academia and none of them would join finance. If, in addition, wages in academia were high enough, then this would set the average skill in finance at a level that can never be surpassed by (a necessarily large, because of the high w_A!) academia. Hence, if being lowest-ranked was not associated with any stigma, then it may not be possible to ensure that academia attracts higher quality workers, no matter how high δ, l_F and even k were. See Online Appendix B.2 for a detailed discussion.

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 payoffs are given by $\pi_A(x_A;\sigma) = w_A + (l_F + \delta)s(G_A(x_A;\sigma) + ko_A(\sigma)$ and $\pi_F(x_F;\sigma) = w_F(x_F) + l_F s(G_F(x_F;\sigma)) + ko_F(\sigma)$. 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 ($M_A(\sigma_e) > M'_A$) and attracts higher-quality talent than finance ($\bar{x}_A^A(\sigma_e) > \bar{x}_F^F(\sigma_e)$) in all equilibria; and (ii) the set of equilibria is

[15] This did not matter in the compensated equilibrium, as punishment necessary for sustaining academia of a given size could always be inflicted through the compensating differential.

non-empty.

Proof. (i) Fix M_A' and l_F, and choose some $\delta > \max\{\delta^*, \frac{w_A - w_F(x_l)}{-s(0)}\} \equiv \bar{\delta}$, where δ^* is as in Lemma 4(ii). This immediately implies that $o_A(\sigma_e) > o_F(\sigma_e)$ in any equilibrium, provided such an equilibrium exists, as otherwise $w_A + \delta s(0) + k(o_A(\sigma_e) - o_F(\sigma_e)) < w_F(x_l)$ and academia unravels. Denote $\min\{4d, \frac{d}{M_A'(1-M_A')}\}$ by d_o. The fact that $\delta > \delta^*$ implies that $o_A(\sigma_c^{M_A}) - o_F(\sigma_c^{M_A}) > d_o > 0$ for any $M_A \leq M_A'$ (by Lemma 4(ii) and the fact that $o_A(\sigma) - o_F(\sigma) = \frac{\bar{x}_A^A(\sigma) - \bar{x}_F^F(\sigma)}{M_A(1-M_A)}$). Consider any $k' \geq \frac{w_F(x_h) - w_A - \delta s(0)}{d_o} \equiv \bar{k}$, as well as an alternative academic wage w_A' for which there is an equilibrium of size M_A; then the compensating differential corresponding to $\sigma_c^{M_A}$ is equal to $c(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}) = w_A' - w_A$ and it follows from Equation (6) that

$$c(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}) = w_F\left(x_F^m(\sigma_c^{M_A})\right) - w_A - \delta s(0) - k'\left(o_A(\sigma_c^{M_A}) - o_F(\sigma_c^{M_A})\right),$$

which is strictly negative for any $M_A \leq M_A'$. Hence there exists no equilibrium in which $M_A(\sigma_e) \leq M_A'$, and thus $M_A(\sigma_e) > M_A'$ 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 $w_A' \in (w_F(\tilde{x}) - \delta s(0), w_F(x_h) - \delta s(0))$, and note that Proposition 1 implies that there exists a unique σ_c that corresponds to $c = w_A' - w_A$; let M_A'' denote the size of academia in that compensated equilibrium. Because $w_F(x_F^m(\sigma_c^{M_A''}) = w_A' + \delta s(0)$, it follows that $x_F^m(\sigma_c^{M_A''}) > \tilde{x}$ in this compensated equilibrium, and hence $M_A'' > H_M(\tilde{x})$ and $\bar{x}_F^F(\sigma_c^{M_A''}) > \tilde{x}$. Finally, because $M_A(\sigma)\bar{x}_A^A(\sigma) + (1 - M_A(\sigma))\bar{x}_A^F(\sigma) = \bar{x}$, where $\bar{x}_A^j(\sigma)$ denotes the average academic skill among members of occupation j, we have that $\bar{x}_A^A(\sigma_c^{M_A''}) < \bar{x}/H_M(\tilde{x}) = \tilde{x}$. It follows, therefore, that $o_F(\sigma_c^{M_A''}) - o_A(\sigma_c^{M_A''}) > 0$, 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 $c(k', \sigma_c^{M_A''}) > 0$ because of academia's negative prestige. As $c(k', \sigma_c^{M_A'}) < 0$ and $c(k', \sigma_c^{M_A''}) > 0$, the continuity of $\psi_{\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}}$ with respect to M_A implies that there has to exist some $M_A > M_A'$ such that $c(k', \sigma_c^{M_A}) = 0$, which concludes the proof. □

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.[16]

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 $w_A + (l_F + \delta)s(0) + ko_A(\sigma_c^{M_A})$ received by the lowest-ranked academic in the compensated equilibrium $\sigma_c^{M_A}$:

$$\frac{\partial^2}{\partial k \partial \delta} \left(w_A + (l_F + \delta)s(0) + ko_A(\sigma_c^{M_A}) \right) = \frac{\partial}{\partial \delta} o_A(\sigma_c^{M_A}) > 0,$$

where the inequality follows from Lemma 2 and Equation (10). Intuitively, in any compensated equilibrium, high δ provides the differentiation of payoffs needed for academia to attract workers of high skill, which increases the prestige of academia. Once the average skill of academics is high enough, the taste for occupational prestige increases the level of payoffs 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 (a) discuss the policy implications of my results (Section 5.1), (b) review the evidence that academia, civil service and officer corps offer flat wage schedules, attract high-quality workers and put great emphasis on rank (Section 5.2), (c) address two alternative mechanisms that could explain the puzzle of selection into academia (Section 5.3), and (d) consider

[16] It 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.

what would happen if occupational prestige was exogenous (Section 5.4).

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 (4) 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.[17] Guaranteeing that academia attracts more workers 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

[17] If $\tau = 1$ 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 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 $\bar{x}^A_A(\sigma_c^{M_A}) > \bar{x}^F_F(\sigma_c^{M_A})$ for any $M_A < 0.5$, and thus (b) $o_A(\sigma_c^{M_A}) - o_F(\sigma_c^{M_A})$ 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.

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 payoffs 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 Evidence on Wage Schedules, Selection and Rank

In this section I will review in more detail the evidence regarding the wage schedules, selection and the importance of rank in academia, civil service and the military.

Academia Hamermesh (2018) provides recent evidence that even in the US—the country most famous for rewarding successful academics generously—academia pays both lower and much less differentiated wages to *doctorate-holders* than other employers. Specifically, at the 5th percentile of the distribution academic doctorate-holders are paid 10% *more* than non-academic doctorate-holders, but at the 95th percentile, academic doctorate-holders are paid 50% less! The average wage is 24% lower in academia. Tables 2, 3 and A.3 in Bakija, Cole, and Heim (2010) indicate that while in 2004 1.58% of academics and 2.05% of finance workers reported an income (excluding capital gains) placing them among 1% top earners, only 0.08% of academics and as much as 0.27% of finance workers reported an income placing them among 0.1% top earners, which suggests much larger upper-tail wage inequality in finance than in academia. Machin and Oswald (2000) discuss the remuneration of workers holding a postgraduate degree in economics, and also find evidence of greater differentiation of wages in the private sector than in academia. Hence, wages are relatively flat in academia both when comparing occupations as a whole, and within relatively narrow categories of workers.

Regarding selection, the results in Stern (2004) suggest that academically-oriented research jobs (i.e., those that incentivise the publication of results) attract better researchers than commercially-focused research jobs. Specifically, Stern (2004) finds that research-oriented jobs pay higher wages than commercially-oriented jobs when not controlling for ability, but lower wages once ability is accounted for. This implies positive selection into academic-oriented research jobs. Direct evidence on selection into academia is available for Sweden. Using data on actual intelligence tests taken by all males in Sweden, Bó et al. (2017) show in their Table II that the average cognitive score among academic economists and political scientists is higher than among parliamentarians, CEOs, and lawyers and judges.

Rank plays an important role in academia. The vast majority of academics make their entire publication and citation records publicly available, and there exist dedicated websites that rank academics both globally and within their countries departments.[18] The highest-ranked academics receive widely publicised and highly prestigious awards, such as the Nobel prize.

Civil Service 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 civil service 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. 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,

[18] For 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).

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.

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. Unfortunately, I was unable to find any evidence on the quality of selection in the officer corps.

5.3 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.

5.3.1 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), 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

$$\pi_A(x_A) = w(x_A), \qquad \pi_F(x_F) = w_F(x_F) + \gamma_P x_P + \mu_P,$$

where $w_i(x_i) = w_i + \gamma_i x_i$. Therefore, a worker is willing to work in finance for a lower wage than in academia if and only if $\gamma_P x_P + \mu_P > 0$.

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

$$\bar{x}_A^A(\sigma_e) = \frac{\phi(z)\left(\gamma_A - \rho_{FA}\gamma_F - \gamma_P\rho_{PA}\right)}{\gamma_V \Phi(z)}, \quad \bar{x}_F^F(\sigma_e) = \frac{\phi(z)\left(\gamma_F - \rho_{FA}\gamma_A + \gamma_P\rho_{PF}\right)}{\gamma_V(1-\Phi(z))},$$

where $\gamma_V = \sqrt{\mathrm{Var}\left(\pi_F(X_F) - \pi_A(X_A)\right)}$, $z = (w_A - w_F - \mu_P)/\gamma_V$, 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

$$\gamma_A - \rho_{FA}\gamma_F - \gamma_P\rho_{PA} > \gamma_F - \rho_{FA}\gamma_A + \gamma_P\rho_{PF}$$

which reduces to

$$\frac{\gamma_F - \gamma_A}{\gamma_P} < -\frac{\rho_{PA} + \rho_{PF}}{1 + \rho_{PA}}. \tag{11}$$

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 ($\gamma_F - \gamma_A > 0$), 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 (11). What is necessary is that workers' preferences for finance and academia are sufficiently *heterogenous*, that is, that γ_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 of 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 $\bar{x}_A^A(\sigma_e) > \bar{x}_F^F(\sigma_e) > 0$, 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.[19] 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 ($G_A(x_l) = 0$), 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.

5.3.2 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,

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

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.

5.4 Exogenous Occupational Prestige

So far, I have assumed that occupational prestige depends endogenously on the talent-pool within an occupation. While this assumption is plausible, it is also possible that occupational prestige depends at least partly on some exogenous characteristics of the job (it's difficulty, or perhaps social usefulness). The model can be very easily reinterpreted to allow for this possibility: positive (negative) occupational prestige would then act exactly as an increase (decrease) in wages by a constant. In other words, a model with exogenously given occupational prestige is isomorphic to the model from Section 4.3, in which only wages and local status are present. This implies, in particular, that exogenous occupational prestige also cannot explain the puzzle on its own.

Similarly, it remains true that the interaction between occupational prestige and local status can explain why academia is both large and attracts highly skilled people. The only difference is that with exogenous prestige, academia needs to either pay relatively high wages on average or be exogenously much more prestigious than finance. In other words, if academia were inherently prestigious then the academic payoff w_A (which now includes wages and prestige) would not be close to $w_F(x_l)$, and thus Theorem 2 (ii) would have no bite.

In order to establish empirically whether (some part of) occupational prestige is endogenous, the best strategy would be to study the impact of exogenous wage or tax shocks on selection. This is similar to the strategy proposed in Section 5.3.1, and for a good reason: if occupational prestige is exogenous, it acts as a positive amenity that is equally valued by all members of the profession. One specific type of a shock that would work well, is a change to the wage function that decreases its gradient but increases the level. This could happen, for example, due to the introduction of occupation-specific minimum wages or collective bargaining. If occupational prestige was exogenous, this would cause an increase in the number of people joining the occupation. However, if a substantial part of prestige was endogenous (and workers cared about prestige sufficiently strongly), then the most likely outcome would be a decrease in the number of people joining the occupation—as the worsening of selection caused by lower wage differentiation would substantially decrease the occupation's prestige, which would lead to a fall in the overall compensation for most workers.

A Omitted Proofs

Proof of Proposition 1

(i) Define extensions (a) $\bar{s} : \mathbb{R} \to \mathbb{R}$ of the local status function $s : [0, 1] \to \mathbb{R}$, such that $\bar{s}(r) = s(0) + \int_0^r s'(\max\{0, \min\{1, t\}\})\mathrm{d}t$, (b) $\bar{w}_F : \mathbb{R} \to \mathbb{R}$ of the finance wage function $w_F : [x_l, x_h] \to \mathbb{R}$, such that $\bar{w}_F(x_F) = w_F(x_l) + \int_{x_l}^{x_F} w_F'(\max\{0, \min\{1, t\}\})\mathrm{d}t$, (c) an extension $\bar{H} : \mathbb{R}^2 \to [0, 1]$ of distribution $H : [x_l, x_h]^2 \to [0, 1]$ such that

$$\bar{H}(x_A, x_F) = \int_{x_l}^{x_F} \frac{\partial}{\partial x_F} H(x_A, \max\{x_l, \min\{x_h, t\}\})\mathrm{d}t,$$

(d) an extension $\bar{L} : [x_l, x_h]^2 \times [0, 1] \to \mathbb{R}$ of L such that $\bar{L}(x_A, x_F, R_A) = \frac{\bar{w}_F(x_F)}{l_F+\delta} + \frac{l_F}{l_F+\delta}\bar{s}(\frac{\bar{H}(x_A,x_F)-R_A}{1-M_A})$, (e) an extension $\bar{Z}$ of Z such that $\bar{Z}$ is an inverse of $\bar{L}$ wrt y and (f) an extension $\bar{F} : [x_l, x_h] \times [0,1] \to \mathbb{R}$ of F such that

$$\bar{F}(x_A, R_A; M_A) \equiv \frac{\partial}{\partial x_A} \bar{H}\left(x_A, \bar{Z}(\frac{w_A}{l_F+\delta} + \bar{s}(R_A/M_A), R_A, x_A; M_A) \right).$$

Let us then define the following IVP:

$$r_A^{\mathrm{IVP}}(x_A; M_A) = \bar{F}(x_A, R_A^{\mathrm{IVP}}(x_A; M_A); M_A) \quad \text{and} \quad R_A^{\mathrm{IVP}}(x_l; M_A) = 0. \tag{12}$$

Note that $R_A^{\mathrm{IVP}}(\cdot; M_A) = R_A(\cdot; \sigma_e)$ if and only if $R_A^{\mathrm{IVP}}(x_h; M_A) = M_A$.[20]

Differentiating the function

$$\bar{L}(x_A, \bar{Z}(y, R_A, x_A; M_A), R_A; M_A) = y$$

wrt to y, R_A, and x_A yields

$$\frac{\partial}{\partial y}\bar{Z} = \frac{1}{\frac{\partial}{\partial x_F}\bar{L}} > 0, \quad \frac{\partial}{\partial R_A}\bar{Z} = -\frac{\frac{\partial}{\partial R_A}\bar{L}}{\frac{\partial}{\partial x_F}\bar{L}} > 0$$
$$\frac{\partial}{\partial x_A}\bar{Z} = -\frac{\frac{\partial}{\partial x_A}\bar{L}}{\frac{\partial}{\partial x_F}\bar{L}} < 0$$

We have, therefore, that

$$\frac{\partial}{\partial R_A}\bar{F} = \frac{\partial^2 \bar{H}(x_A, \bar{Z})}{\partial x_A \partial x_F} \frac{\frac{l_F+\delta}{M_A}\bar{s}'\left(R_A/M_A\right) + \frac{\bar{s}'\left(\frac{\bar{H}(x_A,\bar{Z})-R_A}{1-M_A})\right)}{1-M_A}}{\left(\bar{w}_F'(\bar{Z}) + \frac{\partial}{\partial x_F}\bar{H}(x_A, \bar{Z})\bar{s}(\frac{\bar{H}(x_A,\bar{Z})-R_A}{1-M_A})\right)}. \tag{13}$$

Note that $\bar{s}(\cdot)$, $\bar{w}_F(\cdot)$, and $\bar{H}(x_A, \cdot)$ are continuously differentiable, and hence so is $\bar{Z}(\bullet; x_A)$ (in R_A and y) on $[x_l, x_h]^2 \times [0,1]$, which implies that it is also Lipschitz continuous. For $x_A \in [x_l, x_h]$, $\frac{\partial}{\partial x_A}\bar{H}(x_A, \cdot)$ is also continuously differentiable; thus, $\bar{F}(x_A; \cdot)$ (which is a composition of $\frac{\partial}{\partial x_A}\bar{H}(x_A, \cdot)$ and $\bar{Z}(\bullet, x_A)$) is Lipschitz continuous on $[0,1]$. Therefore, the IVP defined by Equation (12) has a unique solution on $[x_l, x_h]$.

[20]Clearly, $F(x_A, R(x_A; \sigma^e); M_A(\sigma^e))$ coincides with $\bar{F}(x_A, R(x_A; \sigma^e); M_A(\sigma^e))$, which means that $R(x_A; \sigma^e)$ must satisfy Equation (12). Similarly, for any $R_A^{\mathrm{IVP}}(\cdot; M_A)$ such that $R_A^{\mathrm{IVP}}(x_h; M_A) = M_A$ it is clearly the case that $F(x_A, R^{\mathrm{IVP}}(x_A; M_A); M_A)$ coincides with $\bar{F}(x_A, R^{\mathrm{IVP}}(x_A; M_A); M_A)$ for such x_A that $r_A^{IVP}(x_A; M_A) < h_M(x_A)$; hence, any $R_A^{\mathrm{IVP}}(\cdot; M_A)$ such that $R_A^{\mathrm{IVP}}(x_h; M_A) = M_A$ solves Equations (8) and (9).

Furthermore, notice that

$$\frac{\partial}{\partial M_A}\bar{F} = -\frac{\partial^2 \bar{H}(x_A,\bar{Z})}{\partial x_A \partial x_F} \frac{\frac{l_F+\delta}{M_A^2}\bar{s}'\left(R_A/M_A\right) + \frac{\bar{H}(x_A,\bar{Z})-R_A}{(1-M_A)^2}\bar{s}'\left(\frac{\bar{H}(x_A,\bar{Z})-R_A}{(1-M_A)}\right)}{\left(\bar{w}_F'(\bar{Z}) + \frac{\partial}{\partial x_F}\bar{H}(x_A,\bar{Z})\bar{s}(\frac{\bar{H}(x_A,\bar{Z})-R_A}{1-M_A})\right)} < 0.$$

It follows, therefore, by Theorem 6 in Birkhoff and Rota (1969) and the Comparison Theorem (Theorem OA.1 in the Online Appendix C of this paper), that $R_A^{\text{IVP}}(\cdot; M_A)$ is continuous and strictly decreasing in M_A; thus, if a compensated equilibrium exists, it must be unique.

To show existence, let us start by defining

$$\psi^{\text{IVP}}(x_A; M_A) = \bar{Z}(\frac{w_A}{l_F+\delta} + \bar{s}(R_A^{\text{IVP}}(x_A; M_A)/M_A), R_A^{\text{IVP}}(x_A; M_A), x_A; M_A)$$

. It follows that

$$\begin{aligned}
\frac{d}{dx_A}\psi^{\text{IVP}}(x_A; M_A) &= \frac{r_A^{\text{IVP}}(x_A; M_A)}{M_A}\bar{s}'\left(\frac{R_A^{\text{IVP}}(x_A; M_A)}{M_A}\right)\frac{\partial}{\partial y}\bar{Z} \\
&\quad + r_A^{\text{IVP}}(x_A; M_A)\frac{\partial}{\partial R_A}\bar{Z} + \frac{\partial}{\partial x_A}\bar{Z} \\
&= \frac{r_A^{\text{IVP}}(x_A; M_A)}{M_A}\bar{s}'\left(\frac{R_A^{\text{IVP}}(x_A; M_A)}{M_A}\right)\frac{\partial}{\partial y}\bar{Z} > 0.
\end{aligned}$$

Since $r_A^{\text{IVP}}(x_A; M_A) = \frac{\partial}{\partial x_A}\bar{H}(x_A, \psi^{\text{IVP}}(x_A; M_A))$, it must be the case that $r_A^{\text{IVP}}(x_A; M_A) \geq \frac{\partial}{\partial x_A}\bar{H}(x_A, \psi^{\text{IVP}}(x_l; M_A))$, so that

$$\begin{aligned}
R_A^{\text{IVP}}(x_h; M_A) \geq \int_{x_l}^{x_h} \frac{\partial}{\partial x_A}\bar{H}(x_A, \psi^{\text{IVP}}(x_l; M_A))\mathrm{d}x_A &= \bar{H}_M(\psi^{\text{IVP}}(x_l; M_A)) \\
= \bar{H}_M\left(\bar{w}_F^{-1}(w_A + \delta s(0))\right) > 0.
\end{aligned}$$

Pick an arbitrary $M_A' \in (0,1)$ and denote $R_A^{\text{IVP}}(x_h; M_A')$ by M_A''. If $M_A' > M_A''$, then existence follows from the intermediate value theorem because the RHS of Equation (9) is greater than M_A for any $M_A < H_M(\bar{w}_F^{-1}(w_A+\delta s(0)))$. If $M_A' < M_A''$, then existence again follows from the intermediate value theorem because $M_A'' = R_A^{\text{IVP}}(x_h; M_A') > R_A^{\text{IVP}}(x_h; M_A'')$.

Finally, as

$$\frac{\partial}{\partial w_A}\bar{F} = \frac{\partial^2 \bar{H}(x_A, \bar{Z})}{\partial x_A \partial x_F} \frac{1}{\left(\bar{w}'_F(\bar{Z}) + \frac{\partial}{\partial x_F}\bar{H}(x_A, \bar{Z})\bar{s}(\frac{\bar{H}(x_A,\bar{Z})-R_A}{1-M_A})\right)} > 0$$

it follows from the Comparison Theorem and Corollary 1 in Milgrom and Roberts (1994) that $M_A(\sigma_e)$ increases in w_A.

(ii) For $w_A < w_F(x_l) - \delta s(0)$ ($w_A \geq w_F(x_h) - \delta s(0)$) we have that $\pi_A(x_A^m(\sigma_e)) < (>) \pi_F(x_F^m(\sigma_e))$ as long as $x_F^m(\sigma_e) < x_h$, which implies that there are no non-degenerate equilibria. The existence of a degenerate equilibrium follows from the fact an equilibrium exists for any $w_A \in (w_F(x_l) - \delta s(0), w_F(x_h) - \delta s(0))$ (by Definition 2).

If $w_A = w_F(x_l) - \delta s(0)$ then the equilibrium is still characterised by Equations (8) and (9), but the initial condition changes to $R_A(x_A^m(\sigma_e); \sigma_e) = 0$, as our reasoning as to why ($x_A^m(\sigma_e) = x_l$ does not apply anymore (because $x_F^m(\sigma_e) = x_l$). Clearly, $F(x_A^m(\sigma_e), R_A(x_A^m(\sigma_e); M_A); M_A) = \frac{\partial}{\partial x_A}H(x_A, w_F^{-1}(w_A + \delta s(0)))$, which means that for any x_A^m and any M_A the unique solution to the initial-value problem is $R_A^{\text{IVP}}(x_A; M_A) = 0$; and hence $M_A(\sigma_e) = 0$.

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 $c \in (w_F(x_l) - w_A - \delta s(0), w_F(x_h) - w_A - \delta s(0))$ (by Proposition 1 (ii)) and must be characterized by the IVP defined by Equation (12) (with w_A replaced by $w_A + c$). The solution to this IVP can be now expressed as a function of the compensating differential for a given M_A: $R_A^{\text{INV}}(x_A; c)$. It follows by a reasoning analogous to that in the proof of Proposition 1 that (a) $R_A^{\text{INV}}(x_A; c)$ is continuous and increasing in c (because $\frac{\partial}{\partial w_A}F(x_A, R_A) > 0$; (b) from the proof of Proposition1 (ii) follows that for $c = w_F(x_l) - w_A + \delta$ we have $R_A^{\text{INV}}(x_h; c) = 0 < M_A$, and (c) for $c = w_F(x_h) - w_A + \delta$ we have $R_A^{\text{INV}}(x_h; c) = 1 > M_A$. Existence follows from the intermediate value theorem, uniqueness follows from the monotonicity of $R_A^{\text{INV}}(x_A; c)$ in c, whereas the continuity of ψ_{σ_c} wrt M_A is a consequence of the continuity of $R_A^{\text{INV}}(x_A; c)$ wrt c. Second, σ_c 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 will compare the compensated equilibria of two specifications of the model: the *old* one and the *new* one. The old specification is denoted by $\Theta_1 \equiv \{\delta_1, l_{F1}\}$ and the new one by $\Theta_2 \equiv \{\delta_2, l_{F2}\}$. For notational simplicity, I will study small changes to convex combinations of the old and new specifications, with

$$\Theta = \theta\Theta_2 + (1-\theta)\Theta_2.$$

I will denote the mass function holding in compensated equilibrium $\sigma_c^{M_A}$ under parameters Θ by $R_A(\cdot; \sigma_c^{M_A}, \Theta)$. However, as I will consider small changes in θ around Θ, I will generally suppress Θ from notation, so that

$$\frac{\partial}{\partial\theta} R_A(x_A; \sigma_c(M_A)) \equiv \frac{\mathrm{d}}{\mathrm{d}\theta} R_A(x_A; \sigma_c^{M_A}, \Theta) = (\Theta_2 - \Theta_1)\nabla_\Theta R_A(x_A; \sigma_c^{M_A}, \Theta).$$

Note that because $R_A(x_A; \sigma_c(M_A))$ must solve Equation (8), Equation (8) satisfies the conditions from Gronwall (1919), and $\frac{\partial}{\partial M_A} R_A^{IVP}(x_h; M_A) < 0$ (by the proof of Proposition 1), $\frac{\partial}{\partial\theta} R_A(x_A; \sigma_c(M_A))$ exists and is continuous in x_A and θ by the implicit function theorem.

(i) Define the sets

$$\Xi_0 \equiv \{x \in [x_A^m(\sigma_c^{M_A}), x_A^s(\sigma_c^{M_A})] : \frac{\partial}{\partial\theta} R_A(x_A; \sigma_c^{M_A}) = 0\},$$
$$\Xi_1 \equiv \{x \in [x_A^m(\sigma_c^{M_A}), x_A^s(\sigma_c^{M_A})] : \frac{\partial}{\partial\theta} R_A(x_A; \sigma_c^{M_A}) > 0\},$$
$$\Xi_2 \equiv \{x \in (x_1, x_F^m(\sigma_c^{M_A})] : \frac{\partial}{\partial\theta} R_A(x_A; \sigma_c^{M_A}) \leq 0\},$$

where x_1 denotes the infimum of Ξ_1. Similarly, denote the infimum of Ξ_2 by x_2.

I will first show that

$$x', x'' \in \Xi_0, x'' > x', \frac{\partial}{\partial\theta} r_A(x'; \sigma_c^{M_A}) \geq 0, \Rightarrow \frac{\partial}{\partial\theta} r_A(x''; \sigma_c^{M_A}) > 0. \quad (14)$$

To see why, let us start by defining $\bar{c}(\sigma^c_{M_A}) \equiv \frac{w_A + c(\sigma^c_{M_A})}{l_A}$. Next, note that for any $x \in \Xi_0$, $\frac{\partial}{\partial\theta} r(x; \sigma_c^{M_A}) \geq (>) 0$ iff

$$\frac{\partial}{\partial\theta}\,\bar{c}(\sigma_c^{M_A}) \geq (>)\, \frac{\partial}{\partial\theta}\, L\left(x_A, q(x_A), R_A(x_A; \sigma_c^{M_A}; M_A)\right) \tag{15}$$

for all $q(x_A) \equiv Z\left(c + s\left(\frac{R_A(x_A;\sigma_c^{M_A})}{M_A}\right), R_A(x_A; \sigma_c^{M_A}), x_A; M_A\right)$. Differentiating $L(\bullet)$ wrt θ yields

$$\frac{\partial}{\partial\theta}\, L(x_A, x_F, R_A) = w_F(x_F)\,\frac{\partial}{\partial\theta}\left(\frac{1}{l_F+\delta}\right) + s\left(\frac{H(x_A, x_F) - R_A}{1 - M_A}\right)\frac{\partial}{\partial\theta}\,\frac{l_F}{l_F+\delta}.$$

whereas differentiating q wrt x_A yields

$$\frac{\mathrm{d}}{\mathrm{d}x_A} q = \underbrace{\frac{r_A(x_A;\sigma_c^{M_A})}{M_A} s'\left(\frac{R_A(x_A;\sigma_c^{M_A})}{M_A}\right)\frac{\partial}{\partial x_F} Z}_{>0} + \underbrace{\frac{\partial}{\partial x_A} Z + r_A(x_A;\sigma_c^{M_A})\frac{\partial}{\partial R_A} Z}_{=0} > 0.$$

Using Equation (8) one can then find that

$$\frac{\mathrm{d}}{\mathrm{d}x_A}\frac{\partial}{\partial\theta} L(x_A, q(x_A), R_A(x_A;\sigma_c^{M_A})) = \underbrace{\frac{\partial^2 L}{\partial\theta\partial x_A} + r_A(x_A;\sigma_c^{M_A})\frac{\partial^2 L}{\partial\theta\partial R_A}}_{=0} + \underbrace{\frac{\mathrm{d}q}{\mathrm{d}x_A}}_{>0}\underbrace{\frac{\partial^2 L}{\partial\theta\partial x_F}}_{<0} < 0,$$

and implication (14) follows from Inequality (15).

Because $\frac{\partial}{\partial\theta} R(x_l; \sigma_c^{M_A}) = 0$, it must be the case that $x_1 \in \Xi_0$ and $\frac{\partial}{\partial\theta} r_A(x_1) \geq 0$; clearly, $x_2 \in \Xi_0$ as well. Thus, by implication (14) $\frac{\partial}{\partial\theta} r_A(x_2) > 0$ which contradicts the definition of x_2. Hence, if Ξ_1 is non-empty, then Ξ_2 must be empty. But this implies $\frac{\partial}{\partial\theta} R_A(x_h; \sigma_c^{M_A}) > 0$, which contradicts Equation (9). Thus, Ξ_1 is empty as well.

Finally, suppose there exists some $x_3 \in \Xi_0$ such that $x_3 \in (x_l, x_F^s(\sigma^c_{M_A}))$. Then as $x_l \in \Xi_0$ and Ξ_1 is empty, it must be that $\frac{\partial}{\partial\theta} r_A(x_3) \leq 0$, which means, by implication (14), that $\frac{\partial}{\partial\theta} r_A(x_l) < 0$. Hence, $\frac{\partial}{\partial\theta} R_A(x_A) < 0$ for x_A close to x_l, which completes the proof.

(ii) It follows immediately from Equation (7) that $\frac{\partial}{\partial\theta}R_F(\psi(x_A;\sigma_c^{M_A});\sigma_c^{M_A}) = -\frac{\partial}{\partial\theta}R_A(x_A;\sigma_c^{M_A})$ for all $x_A \in [x_A^m(\sigma_e), x_A^s(\sigma_e)]$.

Proof of Lemma 3

(i) Define $\beta \equiv 1/(l_F+\delta)$; we can then express the function L with l_F replaced by β:

$$L(x_A, x_F, R_A) = \beta w_F(x_F) + (1-\beta\delta)s\left(\frac{H(x_A,x_F)-R_A}{1-M_A}\right).$$

Denote the solution to the set of Equations (8)-(9) with respect to R_A and w_A for a given $M_A \in (0,1)$ and $\beta \geq 0$ by $R_A^a(\cdot;M_A,\beta)$. My first goal is to show that $\lim_{l_F\to\infty} R_F(\cdot;\sigma_c^{M_A})$ exists and corresponds to $R_A^a(\cdot;M_A,0)$. It suffices to show that (a) there exists a unique $R_A^a(\cdot;M_A,0)$ and (b) that $R_A^a(\cdot;M_A,\beta)$ is continuous in β at $\beta=0$ (as we have already shown existence, uniqueness and continuity for $\beta>0$).

If $\beta=0$, then the function L becomes $s\left(\frac{H(x_A,x_F)-R_A}{1-M_A}\right)$ and hence its inverse is $Z(y,R_A,x_A) = T((1-M_A)s^{-1}(y)+R_A, x_A)$, where $T(H(x_A,x_F),x_A) = x_F$ for all $x_A \in (x_l, x_h]$. It follows that

$$r_A^a(x_A;M_A,0) = \frac{\partial}{\partial x_A}H\left(x_A, T\left(\frac{R_A^a(x_A;M_A,0)}{M_A}, x_A\right)\right). \tag{16}$$

The issue, of course, is that $H(z_l,x_F)=0$ and hence T is undefined for $x_A=x_l$. Hence, we cannot use $R_A^a(x_l;M_A,0)=0$ as the initial condition. Instead, choose some $\bar{x}_A \in (0, H_M^{-1}(1-M_A))$ and suppose that $R_A(\bar{x}_A)=\alpha$.[21] Because the RHS of Equation (8) is Lipschitz-continuous on $(x_l,x_h)^2$, it follows from Theorem 4.32 in Precup (2018) that there exists a unique solution, denoted by $R_A^{a\mathrm{IVP}}(\cdot;M_A,\beta,\alpha)$, which solves the initial value problem given by Equation (8) and satisfies $R_A^{a\mathrm{IVP}}(\bar{x}_A;M_A,\beta,\alpha)=\alpha$. As a corollary, $R_A^{a\mathrm{IVP}}(\cdot;M_A,\beta,\alpha)$ is differentiable in both α and β by Gronwall (1919). Clearly, $R_A^{a\mathrm{IVP}}(x_A;M_A,\beta,0)=0$ and hence $R_A^{a\mathrm{IVP}}(1;M_A,\beta,0)<M_A$. Similarly, it follows trivially that $R_A^{a\mathrm{IVP}}(1;M_A,\beta,M_A)>M_A$. The existence of a solution to Equations (8)-(9) follows then from the continuity of $R_A^{a\mathrm{IVP}}(\cdot;M_A,\beta,\alpha)$ in α. The Comparison Theorem implies that

[21] $H_M(x_A^{sa}) > 1-M_A$, as otherwise the RHS of Equation (9) must be larger than M_A.

$\frac{\partial}{\partial\alpha}R_A^{a\mathrm{IVP}}(\bar{x};M_A,0,\alpha)>0$, which proves uniqueness.

It remains to show that the compensated equilibrium is continuous in β at $\beta=0$. Consider the function $R_A^{a\mathrm{IVP}}(x_h,M_A,\beta,\alpha)-M_A$; clearly, for $\beta=0$ and $\alpha=R_A^a(\bar{x}_A;M_A,0)$ this function is equal to 0. As $\frac{\partial}{\partial\alpha}R_A^{a\mathrm{IVP}}(x_h;M_A,0,\alpha)>0$, it follows from the implicit function theorem that $R_A^a(\bar{x}_A;M_A,\beta)=\alpha$ is continuously differentiable in β at $\beta=0$. Finally, because

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\beta}R_A^a(x_A;M_A,\beta)=&\frac{\partial}{\partial\beta}R_A^{a\mathrm{IVP}}(x_A;M_A,\beta,R_A(\bar{x}_A,M_A,\beta))\\&+\frac{\partial R_A^a(\bar{x}_A;M_A,\beta)}{\partial\beta}\frac{\partial}{\partial\alpha}R_A^{a\mathrm{IVP}}(x_A;M_A,\beta,R_A^a(\bar{x}_A,M_A,\beta))\end{aligned}$$

it follows that the solution to compensated equilibrium is differentiable (and hence continuous) in β at $\beta=0$.Hence, $R_A^a(\cdot;M_A,0)=\lim_{l_f\to\infty}R_A(x_A;\sigma_c^{M_A})$.

Next, note that Equation 16 can be rewritten as

$$g_A^a(x_A;M_A,0)=\frac{\frac{\partial}{\partial x_A}H(x_A,T((G_A^a(x_A;M_A,0),x_A))}{M_A}. \tag{17}$$

where $G_A^a(x_A;M_A,0)\equiv\frac{R_A(x_A;M_A,0)}{M_A}$ and $g_A^a(x_A;M_A,0)\equiv\frac{\partial}{\partial x_A}G_A^a(x_A;M_A,0)$. By the same logic as above, $G_A^a(x_A;M_A,0)$ is continuously differentiable in M_A; it then follows immediately from the Comparison Theorem that if $\frac{\partial}{\partial M_A}G_A^a(x_A;M_A,0)\leq 0$ for any $x_A\in(x_l,x_h)$, then $\frac{\partial}{\partial M_A}G_A^a(x_h;M_A,0)<0$; contradiction![22] It follows that $\frac{\partial}{\partial M_A}\lim_{l_f\to\infty}G_A(x;\sigma_c^{M_A})>0$ for all $x_A\in(x_l,x_h)$. It is easy to see from Equation (6) that if $l_F\to\infty$ then $G_F(\psi(x_A;\sigma_c^{M_A}),\sigma_c^{M_A})=G_A(x_A;\sigma_c^{M_A})$. From this and Equation (7) follows that if $\lim_{l_f\to\infty}$ then $\frac{\partial}{\partial M_A}G_A(x_A;\sigma_c^{M_A})=-\frac{\partial}{\partial M_A}G_F(\psi(x_A;\sigma_c^{M_A}),\sigma_c^{M_A})$, so that $\frac{\partial}{\partial M_A}G_F(x_F,\sigma_c^{M_A})<0$ for any $x_F\in(x_l,x_h)$. Finally, notice that if $M_A=0.5$ then $G_A^a(x_A;M_A,0)=H(x_A,x_A)$ satisfies Equation (17) and hence $\lim_{l_f\to\infty}G_A(x;\sigma_c^{M_A})=\lim_{l_f\to\infty}G_F(x;\sigma_c^{M_A})$ by Equation (7). The result follows readily.

(ii) The result follows from Lemma 2, Lemma 3 (i) and the fact that $\sigma_c^{M_A}$ is continuous in l_F.

[22]The RHS of Equation (16) is Lipschitz-continuous on $(x_l,x_h]$ so the Comparison Theorem applies.

Proof of Lemma 4

Let us define four constants: $\underline{s'} \equiv \min_{r\in[0,1]} s(r)$, $\bar{s'} \equiv \max_{r\in[0,1]} s(r)$, $w \equiv \frac{\min_{x_F\in[x_l,x_h]} \frac{w_F'(x_F)}{h_M(x_F)}}{\bar{s'}}$ as well as $\kappa \equiv \frac{\underline{s'}}{\bar{s'}}$.

(i) I will prove this part in three steps.

STEP 1: If $M_A \geq 0.5$ and $\delta \leq (1-M_A)w$ then $\bar{x}_F^F(\sigma_c^{M_A}) > \bar{x}_A^A(\sigma_c^{M_A})$.

Under these conditions $\bar{x}_F^F(\sigma_c^{M_A}) - \bar{x}_A^A(\sigma_c^{M_A})$ is decreasing in l_F by Equation (10) and Lemma 2. The result follows because $\lim_{l_F\to\infty} \bar{x}_F^F(\sigma_c^{M_A}) - \bar{x}_A^A(\sigma_c^{M_A}) > 0$ by Lemma 3.

In order to state step 2, we will first need to define $\Pr(X_A < x_A | X_F = x_F) = \frac{\partial}{\partial x_F} H(x_A, x_F)/h_M(x_F)$ as $P(x_A|x_F)$. Second, for any $M_A \in [0.5, 1]$ denote the $x_A \in [x_l, x_h]$ for which $\min_{x_F\in[x_l,x_h]} P(x_A|x_F) = 2(1-M_A)$ by $x_A'(M_A)$.[23] Third, define

$$t(M_A) \equiv \max_{x_F\in[x_l,x_h]} P(x_A'(M_A)|x_F) \in [2(1-M_A), 1].$$

Lastly, define $z \equiv \min_{x_F\in[x_l,x_h]} 1/\left[(x_h - x_l)h_M(x_F)\right]$, with $z \in (0, 1]$.[24]

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

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

STEP 3: If $\delta \leq l_F(\frac{M_A}{\kappa t(M_A)} \frac{4M_A-4+z}{2M_A-2+z} - 1)$ and $M_A > 1 - z/2$, then $\bar{x}_F^F(\sigma_c^{M_A}) > \bar{x}_A^A(\sigma_c^{M_A})$.

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

Lemma 5. For any x_F, either $G_F(x_F; \sigma_{\mathrm{c}}^{M_A}) \leq \alpha H_M(x_F)$ or $G_F(x_F; \sigma_{\mathrm{c}}^{M_A}) \leq 2H_M(x_F) - 1$, which implies that $H_M(x_F) - G_F(x_F; \sigma_{\mathrm{c}}^{M_A}) \geq K(H_M(x_F))$, where

$$K(s) \equiv \begin{cases} (1-\alpha)s & \text{if } s \in [0, 1/(2-\alpha)] \\ 1 - s & \text{if } s \in (1/(2-\alpha), 1]. \end{cases} \tag{18}$$

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

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

Proof. *Case 1:* $x_F \leq x_F^m(\sigma_c^{M_A})$. In that case, $G_F(x_F;\sigma_c^{M_A}) = 0 \leq \alpha H(x_F)$ and the result follows immediately.

Case 2: $x_F \in (x_F^m(\sigma_c^{M_A}), \psi_{\sigma_c^{M_A}}(x'_A(M_A))]$. To make progress, let us use (an extension of) a right inverse of $\psi_{\sigma_c^{M_A}}$, denoted by $\phi_{\sigma_c^{M_A}} : [x_l, x_h] \to [x_l, x_h]$:

$$\phi_{\sigma_c^{M_A}}(x_F) = \sup\{x_A \in [x_l, x_h] : \psi_{\sigma_c^{M_A}}(x_A) < x_F\}. \tag{19}$$

First, note that in this case we have $x_F \leq x_F^s(\sigma_c^{M_A})$.[25] It follows that $\phi_{\sigma_c^{M_A}}(x_F) \leq x'_A(M_A)$ and thus $\frac{\partial}{\partial x_F}H(\phi_{\sigma_c^{M_A}}(x_F), x_F) \leq \frac{\partial}{\partial x_F}H(x'_A(M_A), x_F) \leq h_M(x_F)t(M_A)$, with the last inequality following from the definition of $t(M_A)$,

Of course, Equations (7) and (6) can be equivalently stated as

$$\pi_F(x_F;\sigma_c^{M_A}) - \pi_F(x_F^m(\sigma_c^{M_A})) = \pi_A(\phi_{\sigma_c^{M_A}}(x_F)) - \pi_A(\phi_{\sigma_c^{M_A}}(x_F^m(\sigma_c^{M_A}))) \tag{20}$$
$$H(\phi_{\sigma_c^{M_A}}(x_F), x_F) = M_A G_A(\phi_{\sigma_c^{M_A}}(x_F);\sigma_c^{M_A}) + (1-M_A)G_F(x_F;\sigma_c^{M_A}) \tag{21}$$

Finally, note that $s(r) - s(0) \in [\underline{s}'r, \bar{s}'r]$. Therefore, we have that

$$\begin{aligned}
G_F(x_F;\sigma_c^{M_A}) &\leq \frac{s(G_F(x_F;\sigma_c^{M_A})) - s(0)}{\underline{s}'} \\
&\leq \frac{1}{\underline{s}'}\left(s(G_F(x_F;\sigma_c^{M_A})) - s(0) + \frac{w_F(x_F) - w_F(x_F^m(\sigma_c^{M_A}))}{l_F}\right) \\
\text{[by Equation (20)]} \quad &= \frac{l_A}{\underline{s}' l_F}\left(s(G_A(\phi_{\sigma_c^{M_A}}(x_F);\sigma_c^{M_A})) - s(0)\right) \\
&\leq \frac{l_A \bar{s}'}{\underline{s}' l_F} G_A(\phi_{\sigma_c^{M_A}}(x_F);\sigma_c^{M_A}) \\
\text{[by Equation (21)]} \quad &\leq \frac{l_A \kappa}{M_A l_F} H(\phi_{\sigma_c^{M_A}}(x_F), x_F) \leq \frac{\kappa t(M_A) l_A}{M_A l_F} H_M(x_F) = \alpha H_M(x_F).
\end{aligned}$$

Case 3: $x_F \in [\psi_{\sigma_c^{M_A}}(x'_A(M_A)), 1]$. In that case

$$g_F(x_F;\sigma_c^{M_A}) \geq \frac{\frac{\partial}{\partial x_F}H(x'_A(M_A), x_F)}{1-M_A} = \frac{P(x'_A(M_A)|x_F)h_M(x_F)}{1-M_A} = 2h_M(x_F).$$

As $1 - G_F(x_F;\sigma_c^{M_A}) = \int_{x_F}^1 g_F(r;\sigma_c^{M_A})\mathrm{d}r$, it follows that $G_F(x_F;\sigma_c^{M_A}) \leq 1 - 2(1 - H_M(x_F)) = 2H_M(x_F) - 1$. ☐

[25] From the definition of $x_F^s(\sigma)$ and the increasingness of $\psi_{\sigma_c^{M_A}}$ follows that either $\psi_\sigma(x_F^s(\sigma)) = x_h$ or $x_F^s(\sigma)\} = x_h$. In either case, $\max_{x_A \in [x_l, x_h]} \psi_\sigma(x_A) = x_F^s(\sigma)$.

Next, denote by $G_A^F(\cdot;\sigma)$ distribution of X_A among bankers under sorting σ. Notice that $H_M(x_A) - G_A(x;\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}) < 1 - M_A$, because $H_M(x_A) - M_A G_A(x_A;\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}) = (1-M_A)G_A^F(x_A;\sigma_{\mathrm{c}}^{M_{\mathrm{A}}})$.

Finally, because

$$\bar{x}_i^i(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}) - \bar{x} = \int_{x_l}^{x_h} H_M(x) - G_i(x;\sigma_{\mathrm{c}}^{M_{\mathrm{A}}})\,\mathrm{d}x,$$

it follows that

$$\bar{x}_F^F(\sigma_c^{M_A}) - \bar{x} \geq \int_0^1 \frac{K(r)}{h_M(H_M^{-1}(r))}\mathrm{d}r \geq z(x_h - x_l)\int_0^1 K(r)\mathrm{d}r = \frac{z(x_h - x_l)(1-\alpha)}{2(2-\alpha)}$$

and $\bar{x}_A^A(\sigma_c^{M_A}) - \bar{x} < (x_h - x_l)(1 - M_A)$, so that

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

A little algebra reveals that if $\delta \leq l_F(\frac{M_A}{\kappa t(M_A)}\frac{4M_A-4+z}{2M_A-2+z} - 1)$ and $M_A > 1 - z/2$ then $\bar{x}_F^F(\sigma_c^{M_A}) - \bar{x}_A^A(\sigma_c^{M_A}) > 0$, which proves Step 3.

Set $y = (1-\bar{M}_A)w$. By Step 1, if $\delta < y$, then $x_A^A(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}) - x_F^F(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}) < 0$ for any $M_A \leq \bar{M}_A$. By Step 2, if $M_A > \bar{M}_A$ and $\delta < l_F$, then $\delta \leq l_F(\frac{M_A}{\kappa t(M_A)}\frac{4M_A-4+z}{2M_A-2+z} - 1)$, which by Step 3 implies that $x_A^A(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}) - x_F^F(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}) < 0$.

(ii) Observe that in any compensated equilibrium σ_c, the average occupation-$j \in \{A, F\}$ specific skill among academics is:

$$\begin{aligned}\bar{x}_j^A(\sigma_c) &\equiv G_A(x_A^s(\sigma_c);\sigma_c)E\left(X_j | X_F < \psi_{\sigma_c}(X_A), X_A < x_A^s(\sigma_c)\right) \\ &= (1 - G_A(x_A^s(\sigma_c);\sigma_c))E(X_j | X_A \geq x_A^s(\sigma_c)).\end{aligned} \tag{22}$$

First, I will bound $G_A(x_A^s(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}});\sigma_{\mathrm{c}}^{M_{\mathrm{A}}})$ from above. By Equation (6), we have that

$$\begin{aligned}&w_F(x_h) - w_F(x_l) + l_F s(1) \geq \\ &w_F(x_F^s(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}})) - w_F(x_F^m(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}})) + l_F\left(s(G_F(x_F^s(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}))) - s(0)\right) = \\ &l_A\left(s(G_A(x_A^s(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}))) - s(0)\right) \geq l_A \underline{s}' G_A(x_A^s(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}})),\end{aligned}$$

which can be rewritten as $G_A(x_A^s(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}})) \leq v(\delta, l_F) \equiv \frac{w_F(x_h) - w_F(x_l) + l_F s(1))}{(l_F + \delta)\underline{s}'}$.

Next, let me bound $E(X_j | X_A \geq x_A^s(\sigma_{\mathrm{c}}^{M_{\mathrm{A}}}))$ from below. Because all workers with $x_A > x^s_A(\sigma_{\rm c}^{M_{\rm A}})$ join academia, it must be the case that $H_M(x^s_A(\sigma_{\rm c}^{M_{\rm A}})) > 1-M_A \geq 1-M'_A$. Clearly, then $E(X_j|X_A \geq x^s_A(\sigma_{\rm c}^{M_{\rm A}}) \geq b_j$ where where

$$b_j \equiv \min_{x_A \in [H_M^{-1}(1-M'_A), x_h]} E(X_j|X_A \geq x_A).$$

Trivially, $b_A > \bar{x}$. 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 $H(x_A,x_F) \geq H_M(x_A)H_M(x_F)$ that $F(x_F|x_A) \leq H_M(x_F)$ for all $(x_A,x_F) \in (x_l,x_h)^2$. We have, thus, that $b_F \geq \bar{x}$.

Note that $M_A x^A_F(\sigma_{\rm c}^{M_{\rm A}}) + (1-M_A)x^F_F(\sigma_{\rm c}^{M_{\rm A}}) = \bar{x}$, and hence $x^F_F(\sigma_{\rm c}^{M_{\rm A}}) = \frac{\bar{x}-M_A x^A_F(\sigma_{\rm c}^{M_{\rm A}})}{1-M_A}$. Set $d = 0.5(b_A - \bar{x})$. From Equation (22) follows that

$$\bar{x}^A_j(\sigma_{\rm c}^{M_{\rm A}}) \geq (1-v(\delta,l_F))\, b_j.$$

Hence, we have that

$$\begin{aligned}
\bar{x}^A_A(\sigma_{\rm c}^{M_{\rm A}}) - \bar{x}^F_F(\sigma_{\rm c}^{M_{\rm A}}) &= \bar{x}^A_A(\sigma_{\rm c}^{M_{\rm A}}) - \frac{\bar{x}-M_A x^A_F(\sigma_{\rm c}^{M_{\rm A}})}{1-M_A} \\
&\geq (1-v(\delta,l_F))\, b_A - \bar{x} - \frac{M_A}{1-M_A} v(\delta,l_F)\bar{x} \\
&\geq (b_A-\bar{x}) - v(\delta,l_F)\left(b_A + \frac{M'_A}{1-M'_A}\bar{x}\right).
\end{aligned}$$

It follows that if $v(\delta,l_F) \leq 0.5\frac{b_A-\bar{x}}{b_A+\frac{M'_A}{1-M'_A}\bar{x}}$ then $\bar{x}^A_A(\sigma_{\rm c}^{M_{\rm A}}) - \bar{x}^F_F(\sigma_{\rm c}^{M_{\rm A}}) > d$. For any given l_F, this inequality must be satisfied for sufficiently high δ because $v(\delta,l_F)$ is decreasing in δ and tends to 0 as δ tends to infinity.

Proof of Proposition 2

If $\delta \geq \frac{w_A-w_F(x_l)}{-s(0)}$ ($\delta \leq \frac{w_A-w_F(x_h)}{-s(0)}$), then $w_F(x_l)-w_A+\delta s(0) \geq 0$ ($w_F(x_h)-w_A+\delta s(0) \leq 0$). Thus, by Proposition 1 and Lemma 1 and their proofs it follows that $c(\sigma_{\rm c}^{M_{\rm A}}) > 0$ ($c(\sigma_{\rm c}^{M_{\rm A}}) < 0$), and thus we have that $\psi_{\sigma_{\rm e}}(x_A) = x_l$ ($\psi_{\sigma_{\rm e}}(x_A) = x_h$) 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.

Bakija, J., A. Cole, and B. Heim (2010, March). Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from U.S. Tax Return Data. Department of Economics Working Papers 2010-22, Department of Economics, Williams College.

Becker, G. S. and K. M. Murphy (2000). *Social EEconomic: 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.

Dolton, P., G. Makepeace, and W. van der Klaauw (1989). Occupational Choice and Earnings Determination: The Role of Sample Selection and Non-pecuniary Factors. *Oxford Economic Papers 41*(1), 573–594.

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.

Gronwall, T. H. (1919). Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Annals of Mathematics 20*(4), 292–296.

Hamermesh, D. S. (2018). Why are professors ¡poorly paid¿? *Economics of Education Review 66*, 137–141.

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 roy model. *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 Empricial Investigations*. University of Chicago Press.

Milgrom, P. and J. Roberts (1994). Comparing Equilibria. *American Economic Review 84*(3), 441–59.

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.

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.

Zhan, C. (2015). Money v.s. prestige: Cultural attitudes and occupational choices. *Labour Economics 32*, 44–56.

FOR ONLINE PUBLICATION ONLY

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, with the taste for prestige in occupation i given by $k_i(M_i)$. By Equation (2), the difference in occupational prestige rewards between academia and finance would become $\frac{M_A k_F(1-M_A)+(1-M_A)k_A(M_A)}{(1-M_A)M_A}(\bar{x}_A^A(\sigma)-\bar{x}_F^F(\sigma))$. Thus any non-degenerate equilibrium σ_e of the altered model can be supported in my model with $k = M_A(\sigma_e)k_F(1-M_A(\sigma_e))+(1-M_A(\sigma_e))k_A(M_A(\sigma_e))$, with all other primitives of the model unchanged.

It is easy to see that the only main result that is at any risk from changes to assumptions about k is Theorem 3, and specifically the fact that *all equilibria* must exhibit large M_A and $\bar{x}_A^A(\sigma_e) > \bar{x}_F^F(\sigma_e)$. For this result to hold, we require that $\bar{x}_A^A(\sigma_c^{M_A}) - \bar{x}_F^F(\sigma_c^{M_A}) > d$ for all $M_A < M_A'$ implies that there exists some $d_o > 0$, such that $o_A(\sigma_c^{M_A}) - o_F(\sigma_c^{M_A}) > d_o$ for all $M_A < M_A'$. This must be true as long as $\frac{k_F(1-M_A)M_A+k_A(M_A)(1-M_A)}{(1-M_A)M_A}$ is finite for any $M_A \in [0,1]$; this condition is, of example, trivially satisfied if $k_A(M_A) = M_A k$ and $k_A(1-M_A) = (1-M_A)k$.

B.2 Non Zero-Sum Local Status

The zero-sum nature of local status (captured by the assumption that $\int_0^1 s(r)\mathrm{d}r = 0$) has an impact on the results, but only because it implies $s(0) < 0$. If local status were not zero-sum but $s(0)$ would remain negative, then all results would remain unchanged. The importance of $s(0) < 0$ is that it implies that the skill of the lowest ranked worker in finance, x_F^m, can be *lowered* by increasing δ, which is necessary for ensuring that finance does not attract exclusively high-skilled workers in equilibrium.

In contrast, if $s(0)$ was equal to zero (or positive, although this seems implausible), then it may not be possible to ensure positive selection into academia, even if workers care greatly both about local status and occupational prestige.

Proposition 3. Suppose that $s(0) = 0$ and hence $\int_0^1 s(r)\mathrm{d}r > 0$. Define the unique x that solves $\bar{x}/H_M(x) = x$ by $\tilde{x}$.

If $w_A \geq w_F(\tilde{x})$ and $l_A, l_F, k \geq 0$ then $\bar{x}_A^A(\sigma_e) - \bar{x}_F^F(\sigma_e) < 0$ in every equilibrium of the model.

Proof. Suppose, by the means of contradiction, that there exists an equilibrium σ_e where $\bar{x}_A^A(\sigma_e) - \bar{x}_F^F(\sigma_e) \geq 0$. If $s(0) = 0$ then $x_F^m(\sigma_e) = w_F^{-1}(w_A + k(o_A(\sigma_e) - o_F(\sigma_e)))$. With $w_A \geq w_F(\tilde{x})$, it must be the case that $\bar{x}_F^F(\sigma_e) > x_F^m(\sigma_e) > \tilde{x}$ and $M_A(\sigma_e) > H_M(\tilde{x})$. Finally, because $M_A(\sigma)\bar{x}_A^A(\sigma) + (1 - M_A(\sigma))\bar{x}_A^F(\sigma) = \bar{x}$ where $\bar{x}_A^j(\sigma)$ denotes the average academic skill among members of occupation j under sorting σ, we have that $\bar{x}_A^A(\sigma_e) < \bar{x}/H_M(\tilde{x}) = \tilde{x}$. It follows, therefore, that $\bar{x}_A^A(\sigma_e) - \bar{x}_F^F(\sigma_e) < 0$; contradiction! $\square$

This result demonstrates clearly the important role that the punitive aspect of local status plays in sustaining positive selection into academia. Let me stress, however, that I find the case of $s(0) = 0$ less plausible than $s(0) < 0$. The lack of a punishment implies that workers would receive positive utility from occupying any rank greater than zero within a profession. In other words, workers would be pleased as soon as they were not *the absolute worst* in their profession.

B.3 Alternative Specification of Occupational Prestige

In this section, I explore an alternative specification of occupational prestige, while retaining the assumption that the average prestige payoff in the population is equal to 0. Specifically, consider any random variable $X_R \in [x_l, x_h]$ that has a strictly increasing and continuously differentiable distribution $Z : [x_l, x_h] \to [0,1]$; 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 $J(x_A, x_F, x_h) = H(x_A, x_F) \geq H_M(x_F)H_M(x_A)$. I impose no restrictions on J other than those inherited from H and Z.

Define the occupational prestige in academia and finance as follows:

$$o_F(\sigma) \equiv \frac{\bar{x}_R^F(\sigma)}{E(x_R)} - 1 \tag{OA.1}$$

$$o_A(\sigma) \equiv \frac{\bar{x}_R^A(\sigma)}{E(x_R)} - 1, \tag{OA.2}$$

where $\bar{x}_R^i(\sigma)$ denotes the average x_R among workers who joined occupation i. This functional form ensures that, as in the baseline, the average occupational prestige payoff 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 $J(x_A, x_h, x_R) > H_M(x_A)Z(x_R)$ for all $(x_A, x_R) \in (x_l, x_h)^2$, 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 $\bar{x}_A^A(\sigma_e) < \bar{x}_F^F(\sigma_e)$. It thus follows that occupational prestige cannot, on its own, cause academia to attract workers of (on average) higher skill than finance does.[26] 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 $(\delta, k) \in \mathbf{R}^{\mathbf{2}}_{\geq \mathbf{0}}$ 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 ($\bar{x}_A^A(\sigma_e) > \bar{x}_F^F(\sigma_e)$).[27] This result is weaker than Theorem

[26]If we were to further assume that $J(x_h, x_F, x_R) > H_M(x_F)Z(x_R)$, then $k > 0$ would decrease the size of academia and Theorem 1 would carry over in its entirety.

[27]It should be clear from the proof of Lemma 4(ii), that an analogous result holds also

3, as it does not guarantee that *all* non-degenerate equilibria will have the property that $M_A(\sigma_e) \geq M'_A$, $o_A(\sigma_e) > o_F(\sigma_e)$, and $\bar{x}_A^A(\sigma_e) > \bar{x}_F^F(\sigma_e)$.

Finally, it is worth explaining why the assumption $J(x_A, x_h, x_R) > H_M(x_A)Z(x_R)$ 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.4 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 $g < g^*$. [28] Theorem 3 holds, of course, for any finite g.

B.5 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)

in this alternative specification. That is, for any fixed M, if we set δ high enough, then $o_A(\sigma_c^{M_A}) > o_F(\sigma_c^{M_A})$ and $\bar{x}_A^A(\sigma_c^{M_A}) > \bar{x}_F^F(\sigma_c^{M_A})$. 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 $o_A(\sigma_c^{M_A}) > o_F(\sigma_c^{M_A})$, and thus we can always find some $k > 0$ for which $c(M'_A) = 0$.

[28] In the case of Theorem 2, this follows essentially from (a) the fact that if $l_F \to \infty$ and $M_A \geq 0.5$ then $\bar{x}_A^A(\sigma_c^{M_A}) \leq \bar{x}_F^F(\sigma_c^{M_A})$ for any finite g (Lemma 3) and (b) the continuity of the compensated equilibrium with respect to g. In the case of Theorem 1, the proof is more involved and is available on request.

would leave Theorems 1, 2 and 3 unchanged.

To be specific, define the functions

$$T_i(\sigma) \equiv M_i(\sigma)\int_{x_l}^{x_h} m_i(x)\mathrm{d}G_i(x;\sigma),$$

where $i \in \{A, F\}$ and $m_i : [x_l, x_h] \to \mathbf{R}_{>0}$ is an increasing function. Suppose that the marginal product of worker (x_A, x_F) in occupation i under sorting σ is equal to $p_i(T_i(\sigma))m_i(x_i)$, where $p_i : [0, \bar{T}_i] \to \mathbf{R}_{>\mathbf{0}}$ is decreasing and continuous and $\bar{T}_i = \int_{x_l}^{x_h} m_i(x)\mathrm{d}H_M(x)$. Note that as p_i maps from the number of efficiency units of labor provided by occupation i into the price the market is willing to pay, it is effectively an inverse demand function. 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 $w_A \in (\frac{p_F(\bar{T}_F)}{p_A(0)}m_F(x_l), \frac{p_F(0)}{p_A(\bar{T}_A)}m_F(x_h))$.

Let us define a new concept, 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 3. A sorting σ_{cp} constitutes a *twice-compensated equilibrium* if and only if (a) σ_{cp} is non-degenerate and (b) there exist some *compensating differential* $c_d \in \mathbf{R}$ and a *compensating price* $c_p \in \mathbf{R}_{>\mathbf{0}}$ such that for all $(x_A, x_F) \in [x_l, x_h]^2$,

$$(x_A, x_F) \in \sigma_{cp}^{-1}(\{A\}) \Rightarrow l_A s(G_A(x_A; \sigma_{cp})) + c_d > c_p m_F(x_F) + l_F s(G_F(x_F; \sigma_{cp})),$$
$$(x_A, x_F) \in \sigma_{cp}^{-1}(\{F\}) \Rightarrow l_A s(G_A(x_A; \sigma_{cp})) + c_d < c_p m_F(x_F) + l_F s(G_F(x_A; \sigma_{cp})).$$

A sorting σ_c constitutes a compensated equilibrium if it constitutes a twice-compensated equilibrium with $c_p = p_F(T_F(\sigma_c))$. A sorting σ_e constitutes an equilibrium if it constitutes a compensated equilibrium with $c_d = w_A p_A(T_A(\sigma_e)) + k(o_A(\sigma_e) - o_F(\sigma_e))$.

A sorting σ_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.[29]

B.6 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 payoffs are. To see why, consider the no-status baseline and assume that $x_A = 1 - x_F$ and $h_M(x_F) = x_F$, 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 $\geq x_h - \psi^b$ join academia. This implies that $\bar{x}_F^F(\sigma_e) > \bar{x}_A^A(\sigma_e)$ if and only if $M_A(\sigma_e) > 0.5$. 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.[30] It follows, therefore, that, at the very best, we can have $\bar{x}_A^A(\sigma_e) > \bar{x}_F^F(\sigma_e)$ only if $M_A(\sigma_e) < 0.5$.

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

[30]If 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.

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

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

respectively, where $A(x, y) \leq B(x, y)$ for $x \in [x_l, x_h]$ 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 $x \in (x_l, x_h]$. If, further, $A(x, h(x)) < B(x, h(x))$ or $h(a) < k(a)$, then $h(x) < k(x)$ for all $x \in (x_l, x_h]$.

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