CESifo WORKING PAPERS

10437
2023

May 2023

Underrepresentation, Quotas and Quality: A Dynamic Argument for Reform

Malin Arve, Justin Valasek

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10437*

Underrepresentation, Quotas and Quality: A Dynamic Argument for Reform

Abstract

The trade-off between increased representation and perceived quality is central to the debate on how to address underrepresentation in high-profile professions. We address this trade-off using a dynamic model of career selection where juniors value both the identity and perceived quality of their mentors (seniors). A preference for homophily results in the persistence of underrepresentation, suggesting intervention is needed. However, if an abrupt quota causes a large decrease in the perceived quality of underrepresented seniors, then underrepresented juniors of high talent will select out of the profession, causing a permanent (real) quality difference. Encouragingly, we show that gradual reform—while decreasing perceived quality in the short term—enables a transition to equal representation and equal quality in the long term. We discuss the implications of our analysis for commonly-used measures to increase representation.

JEL-Codes: D620, E240, I200, J150, J160, J240.

Keywords: affirmative action, quotas, mentorship, identity, gender, adverse selection.

Malin Arve
Norwegian School of Economics (NHH)
Bergen / Norway
malin.arve@nhh.no

Justin Valasek
Norwegian School of Economics (NHH)
Bergen / Norway
justin.valasek@nhh.no

May 15, 2023
We would like to thank Sarah Auster, Rahul Deb, Thomas Giebe, Hans Grönqvist, Bard Harstad, Steffen Huck, Botond Kőszegi, Rachel Kranton, Konuray Mutluer and seminar participants at BI: Bergen, and the 2022 EUI Alumni Conference for their comments and suggestions.

1 Introduction

The persistent underrepresentation of women and minority groups in many high-skilled professions is a well-documented phenomenon, and has led to an active policy debate over how to tackle underrepresentation.[1] While certain pundits and scholars have argued against the use of quotas and affirmative action to address underrepresentation on the basis that it introduces a stigma of lower quality among underrepresented professionals (see for example Robin, 2019 on Supreme Court Justice Clarence Thomas), it is unclear whether this argument outweighs the clear benefit of increasing representation and diversity in high-skilled professions. Importantly, the stigma argument only considers a static perspective and largely ignores the dynamic arguments for increased representation. A large empirical literature has documented the importance of role models in the decision of labor-profession entrants to pursue a specific career, suggesting that affirmative action today may result in a more representative workforce tomorrow (see Porter and Serra, 2020, and Riise et al., 2022 for an overview). Therefore, any potential stigma associated with quotas may be transitory as a profession moves towards equal representation.

To address the trade-off between increased representation and perceived quality, we analyze identity-based hiring quotas in a formal dynamic model and are thus able to provide additional structure to this debate. Specifically, we provide insight as to when quotas will result in a transition to a steady state with equal representation *and* equal quality in the long term. We find that stigma and lower perceived quality is not an argument against quotas per se, but that it can have important implications for how quotas are implemented. Specifically, if a quota is implemented in a given profession and causes a large shift in the relative perceived quality, then it can cause an entrenched (real) quality difference as high-talent juniors of the underrepresented type select out of the profession. However, as long

[1] See for example Auriol et al. (2022) for global evidence from the academic profession for economists and Wallon et al. (2015); European Commission (2019) for a overview of the policy debate.

as perceived quality is proportionate to the size of the implemented quota, then gradual reform results in a transition to equal representation and equal quality.

In our analysis, we follow the example of Athey et al. (2000) and Müller-Itten and Öry (2022) and consider a setting where mentorship plays a key role both in the development of quality, and where potential career entrants (juniors) value both the identity and quality of their mentor.[2] Specifically, we incorporate a preference for identity-homophily into an overlapping generations (OLG) model and show that, in this setting, underrepresentation is persistent. This is the result of a cycle where, due to the relative lack of seniors of the same identity type, juniors of the under-represented identity disproportionately select out of the profession. This cycle causes persistent underrepresentation at the senior level and illustrates the need for policy intervention to address underrepresentation.

We first consider a 1:1 quota on juniors as a way of correcting for underrepresentation at the senior level. However, we find that a quota on juniors does not solve the problem of underrepresentation due to an adverse selection problem: while a quota on juniors mechanically equalizes representation at the junior level, it does not increase the number of high-talent juniors of the underrepresented type, and therefore may not lead to an increase in the number of *seniors* of the under-represented type.

Next we consider quotas for hiring seniors of the under-represented type as a direct approach to addressing underrepresentation. Conceptually, we assume that quotas do not have a direct impact on the quality of seniors—indeed most empirical studies have shown that quotas do not lower quality.[3] Instead, studies have shown that affirmative action and quotas can negatively impact the *perceptions* of quality (e.g. Heilman et al., 1992; Coate and Loury, 1993; Fang and Moro, 2011; Leslie, 2014). Accordingly, we consider the case where quotas impact the

[2]Mentorship is an important factor in many different career areas such as politics, law and academia. Moreover, identity-homophily is a well-documented fact in some fields of academia (Hilmer and Hilmer, 2007; Gaule and Piacentini, 2018) and in the judiciary branch (Battaglini et al., 2022).

[3]For example Besley et al. (2017) show that a quota actually *increased* the average quality of politicians in Sweden, a country considered to be one of the most egalitarian in the world. This, suggests that hiring quotas may in practice serve to address bias.

perception of career entrants, who associate quotas with lower perceived quality of seniors of the underrepresented type.

We find that even if quotas do not have a direct impact on quality, perceived quality may impact real quality through entry decisions of career entrants. Specifically, while identity-based hiring quotas at the senior level mechanically address underrepresentation, the quality of the implemented steady state depends on the dynamic structure of the quota. That is, quotas can either result in a transition to a steady state with equal average quality of seniors of both types, or result in a transition to a steady-state where high-talent juniors of the underrepresented type select out of the profession, resulting in a lower average quality of mentors of the historically underrepresented type.

The transition to an unequal steady-state can occur since, with a preference for homophily, juniors disproportionately value the quality of seniors of their identity type. Therefore, if a quota causes a large enough decrease in perceived quality of seniors of the underrepresented type, then high-talent juniors of the underrepresented type will disproportionately select out of the profession, causing a transition to a steady state with unequal (real) quality. Our analysis therefore points to a gradually increasing quota on seniors of the underrepresented type as a way to ensure a transition to a steady state with equal representation and equal quality.

In addition to the literature on role-models discussed above, our research contributes to the theoretical literature on affirmative action and underrepresentation (see Fershtman and Pavan, 2021 for an overview). Our work is most closely related to Athey et al. (2000) and Müller-Itten and Öry (2022), who study quotas in the context of juniors who value the identity-composition of the mentor pool and find that quotas may be required to maintain equal or optimal representation. Arguments for quotas to address underrepresentation are also presented in Siniscalchi and Veronesi (2020) and Carvalho and Pradelski (2022) using alternative models of underrepresentation based on a mechanism of, respectively, self-image bias and in-group norms. We expand on this research by accounting for the fact that juniors' career decisions may also depend on the perceived quality-composition of the mentor pool, which allows us to address the important trade-off between representation and quality that is often central to the debate surrounding affirmative action and quotas. This innovation leads to our novel insight that the dynamics of quotas matter: in contrast to previous research, we highlight that the speed of reform is crucial because it can impact whether the profession converges to equal quality, or to a steady state where high-talent juniors of the underrepresented type select out of the profession.

Lastly, we discuss the implications of our analysis for measures that have been proposed or implemented for addressing underrepresentation. Our results suggest that the "cascade model"—a quota at each level of seniority that is equal to the level of representation at the level below—can be counterproductive since it may result in a lower perceived quality of underrepresented seniors relative to *both* a more gradual transition and to an immediate transition to equal representation. In contrast, a preference for underrepresented seniors in cases of equal quality avoids the problem of the cascade model, but can lead to a cycle where representation is increased in one period and reduced in the next. Therefore, this model may require an occasional "nudge" to keep the profession on the path towards equal representation.

2 Theoretical Framework

We consider an OLG setup where each agent lives for two periods. In the first period each agent is a career entrant and can apply for a junior position in a given profession which includes on-the-job mentoring by a senior colleague. For shortness of the exposition, we will refer to the two levels of positions as juniors and seniors. Conditional upon being hired after the junior period, he or she becomes a senior. In each period, the profession consists of a continuous population of seniors of mass 1. Each senior has the capacity to mentor $\lambda \in \mathcal{N}$ juniors. There is also a continuous population of potential juniors of each identity and talent type of size N, where N is arbitrarily large.[4] Identity and talent is described in more detail below. For simplicity, we introduce the notation without the time subscript, t.

Types: Over the life-cycle, each agent i is characterized by a four-dimensional type (I_i, q_i, Q_i, o_i) where we refer to I_i as the agent's identity-type, q_i, as the talent of the agent as a junior (i.e., in the first period), Q_i the quality of the agent as a senior (i.e., in the second period) and o_i the value of the agent's outside option.

The identity-type space is binary and each agent i has an identity $I_i \in \{A, B\}$ that is observable and constant over time. This identity-type can for example be the agent's gender or national identity. We denote by $\mathbf{M}^I$ (respectively $\mathbf{m}^I$) the set of senior (resp. juniors) of identity-type I.

Junior talent is binary, and we denote a junior's talent by $q_i \in \{h, l\}$. Beyond own talent being observed, juniors' talent is partially observed: we assume that only seniors with a quality greater than some threshold, Q_H, observe juniors' talent. This allows us to account for the fact that juniors with perceived high-talent may be more likely to match with high-quality mentors.[5] Additionally, in the first stage, juniors have an outside option which is valued at $o_i = o_{q_i}$ with $o_h > o_l$.

Seniors' quality is continuous, and we denote a seniors' quality by $Q_j \in [0, 1]$. Senior quality is public information but it is only realized in the agent's second period. We introduce perceived quality, and the impact of quotas on perceived quality, after our benchmark analysis.

The assumptions of of binary talent of juniors and continuous quality of seniors are made for the following reason: As fresh on the profession, the juniors have not had the time to develop their individual skills to the same extent as the seniors. The binary, more coarse, assumption for juniors and the continuous, more nuanced, assumption for seniors reflect this. Furthermore, the organization's abil-

[4] This assumption assures that our results are not driven by a limited supply of juniors of a given identity type.

[5] For simplicity, we assume that seniors with perceived high-quality observe talent after juniors have been admitted, rather than in the application stage. The results are qualitatively similar if quality is observed in the application stage, which implies that juniors with perceived high-talent have a higher probability of being admitted; however, this complicates the analysis substantially.

ity to learn more about the junior during job appraisals is also reflected in this assumption.

In our analysis, two important metrics are the size of the different sets of senior identity types, which we denote with $M^I = |\mathbf{M}^I|$ and average senior quality for identity type I, $\bar{Q}^I$. Furthermore, since our analysis will often involve the size of sets, we use non-bold, italic notation to refer to the size of a set: e.g. M^I_Q is the size of set of seniors of identity-type I and quality Q.

Timing, Choices and Technology: The model consists of three stages: agents first decide whether to enter this particular profession and apply for a job which includes on-the-job mentoring. Then, conditional upon entry, juniors are matched to a senior and realize a senior-quality Q_i. Lastly, a mass 1 of juniors are hired as seniors for $t+1$. Since Q_i is observable and the profession has a strict preference for quality, the hiring rule consists of an endogenous quality cutoff above which all seniors are hired (we formally introduce identity quotas in Section 4). While we explicitly model the decision to enter the industry and apply for a job, our analysis will focus on characterizing the size of the sets of juniors that enter in equilibrium: $\{m^A_h, m^B_h, m^A_l, m^B_l\}$. Therefore, much of the machinery and notation we introduce here will operate in the background.

In each period, all juniors choose to apply for the job (which includes mentoring) or not, and all juniors who apply have an equal probability of being hired. There is a fixed application cost c for applying. This may represent actual monetary costs or the cost of specialized training (e.g. GRE prep, etc.). We use the notation $\hat{a}_i = 1$ to denote that i applies for a job, and $\hat{a}_i = 0$ if i does not apply. Furthermore, the notation $a_i = 1$ denotes that i obtained a junior position, and $a_i = 0$ that i did not.

Additionally, each junior m who applies indicates his or her preference over their preferred senior identity-type, $\hat{I}^m$, and are assigned to the set of A seniors or to the set of B seniors. If an identity type is not over-demanded, then all juniors who prefer this type are matched to this set of seniors. If one identity type is over-demanded, then all juniors are randomly allocated over their preferred type and the remaining slots in the set of the under-demanded identity type (all juniors are matched to a senior). After being matched to a set of seniors, the highest quality seniors with $Q_j > Q_H$, who observe junior talent, match with high-talent juniors and the remaining set of juniors are randomly allocated to seniors in their identity set.

After completing the first-period job/training, all juniors realize their own senior quality, Q_i. For high-talent juniors, Q_i is a random variable whose distribution depends on the quality of the senior that mentored them in the first period. First, juniors who are matched with the highest quality seniors, $Q^m > Q_H$, Q_i is drawn from a uniform distribution over $[0,1]$ with probability one. For high-talent juniors who are not matched a senior with $Q^m > Q_H$, Q_i is drawn from a uniform distribution over $[0,1]$ with probability $f(\bar{Q}^m)$, where $\bar{Q}^m$ is the average quality of the set of seniors of identity-type m; that is, with probability $(1 - f(\bar{Q}^m))$, $Q_i = 0$.[6] The function $f(\bar{Q}^m)$ maps $[0,1] \to [0,1]$, with $f(0) = 0$, and is monotonically increasing and differentiable in $\bar{Q}^m$.[7]

All low-talent juniors realize $Q_i = 0$ with probability equal to one. However, we assume that low talent juniors do receive a benefit from training: low talent juniors increase their outside option, and hence their wage, by $\beta\bar{Q}^m + c$ with $\beta > 0$. This is, the benefit to training is a function of the senior quality for both high and low-talent juniors.[8]

Payoffs: Agents who do not enter training and low-talent juniors receive a wage equal to the value of their outside option. High-talent juniors receive a wage

[6]Formally, $\bar{Q}^I = \int_0^1 Q dG^I(Q)$, where $G^I(Q)$ is the distribution of senior quality by identity-type and we set $\bar{Q}^I$ equal to 0 if the set of seniors is empty.

[7]For simplicity we assume that the probability that juniors realize positive senior quality is a function of the average quality of the group of seniors to which i is matched—this assumption is not without complete loss of generality, but we note that because juniors value their expected senior quality, for each function $f(Q^m)$ mapping individual senior quality to $[0,1]$ that is monotonically increasing, there is a monotonically increasing function $f(\bar{Q}^m)$ with equivalent expected payoffs.

[8]Our main results also hold under the assumption that low-talent juniors realize positive senior quality according to some function $f^l(\bar{Q}^m)$, where $f^l(\bar{Q}^m) < f(\bar{Q}^m)$ for all $\bar{Q}^m$. However, the assumption that all low-talent juniors realize $Q_i = 0$ allows us to characterize the dynamics of $\bar{Q}$ in a relatively straightforward matter, and to more cleanly illustrate our main results.

equal to their senior quality. Juniors value homophily in their mentor-match as well as wages, and have the following intertemporal von Neumann Morgernstern utility function:

$$u_i(w_i, \hat{a}_i, I_i, I_i^m) = w_i - \eta \mathbb{1}(I_i \neq I_i^m) + \hat{a}_i c + (1 - \hat{a}_i) o_i, \quad (1)$$

where $\eta \geq 0$ measures the utility from homophily between senior that provides mentoring and the junior.

Equilibrium and steady state: We consider symmetric period equilibria, $\sigma_q^I = \Pr(\hat{a}_i = 1 | q, I)$ and $\{\hat{I}^m\}$ that maximizes the expected utility given $\{\mathbf{M}^I\}$. We define a steady state as $\{\mathbf{M}^{I*}\}$ such that a period equilibrium exists with $\mathbf{M}_t^{I*} = \mathbf{M}_{t-1}^{I*}$.

3 Benchmark Analysis

We begin by analyzing the model without homophily nor with quotas to establish a benchmark and to build intuition regarding our main results. Since identity is not directly payoff relevant if $\eta = 0$, we first focus on characterizing "identity neutral" symmetric equilibria where $\bar{Q}_t^A = \bar{Q}_t^B = \bar{Q}_t$ and $\sigma_{q,t}^A = \sigma_{q,t}^B = \sigma_{q,t}$ as the relevant benchmark.

To simplify the presentation of the analysis, we will first show that it is possible to define equilibrium entry choices indirectly by characterizing the set of juniors that enter the profession in equilibrium, $\{\mathbf{m}_{q,t}\}$. This allows us to focus on the object of interest—the number of juniors of the different types that enter—instead of referring directly to the probability of applying for entry ($\sigma_{q,t}$). To achieve this, we introduce some additional notation.

Note that conditional upon entry, the probability that a high-talent junior realizes a positive senior quality depends on the quality of the senior that junior is assigned, $\bar{Q}_t$, which again depends on the size of the set of high-quality seniors, $M_{H,t}$, and the size of the set of high-quality juniors, $m_{h,t}$. Therefore, we introduce $g(\bar{Q}_t, m_{h,t})$ to denote the probability of realizing a positive senior quality conditional on entry (we omit $M_{H,t}$ from the arguments of $g(\cdot)$ due to the straightforward relationship between $\bar{Q}_t$ and $M_{H,t}$):

$$g(\bar{Q}_t, m_{h,t}) = \frac{\lambda M_{H,t}}{m_{h,t}} + \left(1 - \frac{\lambda M_{H,t}}{m_{h,t}}\right) f(\bar{Q}_t)$$

$g(\bar{Q}_t, m_{h,t})$ is increasing in $\bar{Q}_t$, but decreasing in $m_{h,t}$ since an increase in $m_{h,t}$ implies that the "competition" for high-quality seniors increases.

Period Equilibria

Before characterizing the steady states, we detail period equilibria given average senior quality, and therefore drop the period notation, t, for the first part of the analysis. We begin by characterizing the best response functions:

$$\sigma_i = \begin{cases} 1 & \text{if } E[u_i|\hat{a}_i = 1, \bar{Q}, q_i, \sigma_{-i}] > o_{q_i}, \\ \sigma \in [0,1] & \text{if } E[u_i|\hat{a}_i = 1, \bar{Q}, q_i, \sigma_{-i}] = o_{q_i}, \\ 0 & \text{if } E[u_i|\hat{a}_i = 1, \bar{Q}, q_i, \sigma_{-i}] < o_{q_i}. \end{cases}$$

That is, in equilibrium juniors will apply for training to the point where the expected utility from applying is equal to or less than the outside option.

Since all juniors are indifferent between applying and the outside option in any interior equilibrium this implies that the expected wage conditional on entry relative to the outside option must be the same for high and low-type juniors. Formally:

Lemma 1. *In equilibrium, if $\sigma_h, \sigma_l > 0$, then the following condition must hold:*

$$\beta\bar{Q} - o_l = g(\bar{Q}, m_h)/2 - o_h. \tag{2}$$

Lemma 1 illustrates the basic structure of period equilibria: Given $\bar{Q}$, the expected wage of low-type juniors is fixed. However, the expected wage for high types, $g(\bar{Q}, m_h)/2$, is decreasing in m_h because high-talent juniors "compete" over the seniors with $Q_j > Q_H$. Therefore, given σ_l, high types will apply to training to the point where $g(\bar{Q}, m_h)$ is low enough for (2) to hold.[9]

Lemma 1 also shows that period equilibria can be characterized indirectly in terms of the equilibrium size of the sets of juniors, m_h^*, m_l^*. That is, an interior value of m_h^* is implicitly defined as a function of $\bar{Q}$ in (2). Rearranging (2) and using the expression for $g(\bar{Q}, m_h)$, we get a closed form solution for an interior value of m_h^*:

$$m_h' = \lambda M_H \left[\frac{1 - f(\bar{Q})}{(\beta\bar{Q} - o_l) - (f(\bar{Q}) - o_h)} \right], \qquad (3)$$

where $m_h^* = m_h'$ if the value of m_h' is in $(0, \lambda)$.

The following result also establishes that, naturally, a corner equilibrium exists when $m_h' \notin (0, \lambda)$, and that the period equilibrium is unique.

Proposition 1. *The period equilibrium, m_h^*, m_l^*, is unique and m_h^* is characterized by:*

$$m_h^* = \begin{cases} 0 & \text{if } m_h' \leq 0, \\ m_h' & \text{if } m_h' \in (0, \lambda), \\ \lambda & \text{if } m_h' \geq \lambda, \end{cases}$$

where m_h' is defined by (3).

Characterization of Steady States

Next, we characterize the steady states of the model and establish that despite unique period equilibria, multiple steady states may exist. First we consider interior steady states and characterize the dynamics of quality, $\bar{Q}_t$—since we are

[9] In the proof of Lemma 1 we show that $m_h \geq \lambda M_H$ in all period equilibria, implying that $m_h > 0$ if $\bar{Q} > 0$.

considering an identity-neutral model, we can characterize a steady state by the average quality, $\bar{Q}$, rather than referring to set notations.

Note that $\bar{Q}_{t+1}$ is determined by the size of the set of juniors who realize positive senior quality in period t, which is defined by the following equation:

$$|\{i : Q_i > 0\}| = m_{h,t} g(\bar{Q}_t, m_{h,t}). \tag{4}$$

Since the set of juniors that realize positive senior quality in time t have Q_i distributed uniformly over $[0,1]$, the profession will hire all seniors who realize a quality of $Q_{L,t}$ or greater where $Q_{L,t}$ satisfies the expression:

$$(1 - Q_{L,t}) m_{h,t} g(\bar{Q}_t, m_{h,t}) = 1. \tag{5}$$

Moreover, since $\bar{Q}_{t+1}$ is characterized by the following expression:

$$\bar{Q}_{t+1} = \frac{1 - Q_{L,t}}{2}, \tag{6}$$

we can substitute for $Q_{L,t}$ using Equation 5 to get:

$$\bar{Q}_{t+1} = \frac{1}{2 m_{h,t} g(\bar{Q}_t, m_{h,t})}. \tag{7}$$

Lastly, since (3) characterizes $m_{h,t}$ as a function of $\bar{Q}_t$, this implies that (7) details $\bar{Q}_{t+1}$ as a function of $\bar{Q}_t$, and $\bar{Q}_{t+1}(\bar{Q}_t)$ can be used to characterize the dynamics of quality as a function of o_l, o_h, $f(\cdot)$ and β.

Moreover, (3) also identifies the interior steady states of the model. However, if interior steady states exist, they are not unique. As shown in the following proposition, a steady state of the model also exists at $\bar{Q} = 0$.

Proposition 2. *$\bar{Q}^* = 0$ is a steady state of the model. Interior steady states, $\bar{Q}^* \in (0,1)$, exist if and only if $\bar{Q}_{t+1}(\bar{Q}^*) = \bar{Q}^*$.*[10]

[10] $\bar{Q}^* = 1$ cannot be a steady state since the set of juniors that realize $Q_i = 1$ has no mass.

A corner solution with $\bar{Q}^* = 0$ is a steady state since only low-talent juniors enter if $\bar{Q}_t = 0$, and low-talent juniors do not realize positive quality, which implies that $\bar{Q}_{t+1} = 0$. The fact that there may be an interior and a corner steady state will be important when considering policy interventions since, as we show in the following section, with quotas and $\eta > 0$ it may be possible for the different identity groups to converge to different average senior quality, and in particular to a steady state where high-talent juniors of one identity type exit the profession.

Lastly, note that the steady states characterized in the above are "identity-neutral" in the sense that any composition of $\mathbf{M}^A$ and $\mathbf{M}^B$ constitute a steady state of the model as long as the corresponding $\bar{Q}$ is a steady state. Moreover, there is no persistence of identity at the steady states—if $M_t^A < M_t^B$ in period t, there exist $\mathbf{M}_{t+1}^A$ and $\mathbf{M}_{t+1}^B$ with $M_{t+1}^A = M_{t+1}^B$ that correspond to a period equilibrium as long as $\bar{Q}_t = \bar{Q}_{t+1}$, implying that a transition to equal representation can be achieved in a single period. As we show next, this changes drastically when we introduce a preference for homophily to the model.

Persistence of Underrepresentation with Homophily Payoff

The analysis of the model with a preference for homophily is similar to the analysis above, with the exception that, depending on the size of the various identity groups, juniors may face a negative utility associated with the probability of being matched with a mentor of a different type. We denote this probability, which is a function of $\{M^I, m^I\}$, in short-hand as $\Pr(I_{\neg i}|I_i)$.

Otherwise, from a technical perspective the analysis is analogous to the model without homophily. In particular, as we show in the following lemma, the expected relative value of entry is identical in equilibrium for all types that apply with positive probability, where the relative expected utility of entry for low-talent and high-talent, respectively, is:

$$\Pr(I_{\neg i}|I_i)\left(\beta\bar{Q}^{I_{\neg i}} - o_l - \eta\right) + (1 - \Pr(I_{\neg i}|I_i))\left(\beta\bar{Q}^{I_i} - o_l\right),$$

$$\Pr(I_{\neg i}|I_i)\left(g(\bar{Q}^{I_{\neg i}}, m_h) - o_h - \eta\right) + (1 - \Pr(I_{\neg i}|I_i))\left(g(\bar{Q}^{I_i}, m_h) - o_h\right),$$

where $g(\bar{Q}^I, m_h)$ is the probability of matching with a high-quality senior *conditional* on being assigned to the group of seniors with identity I (note that m_h in $g(\bar{Q}^I, m_h)$ refers to the number of high-talent juniors that match with mentors of identity I).

Lemma 2. *All types that enters with positive probability in equilibrium ($\sigma_h^A, \sigma_l^A, \sigma_h^B$ or $\sigma_l^B > 0$) have an equal relative expected utility of entry.*

Next we link the analysis to the benchmark model without homophily by showing that for any steady state with $\eta = 0$, there exists a corresponding steady with equal representation and equal quality; i.e. $M^A = M^B$ and $\bar{Q}^A = \bar{Q}^B = \bar{Q}^*$. Note that with homophily, it is no longer sufficient to refer to average quality to identify a steady state, so we will revert to using the full set notation $\{\mathbf{M}^A, \mathbf{M}^B\}$.

Proposition 3. *Take $\bar{Q}^*$ to be a steady state given $\eta = 0$. For any $\eta > 0$, $\{\mathbf{M}^A, \mathbf{M}^B\}$ is a steady state if the corresponding $\bar{Q}^A = \bar{Q}^B = \bar{Q}^*$.*

Intuitively, we can think of the steady state with equal quality in both groups and $\eta > 0$ as two separate professions for $I = A, B$, where juniors of type A only match with mentors of type A. If both professions are at a steady state, i.e. if $\bar{Q}^A = \bar{Q}^B = \bar{Q}^*$, then the overall profession is at a steady state as well, and no juniors have an incentive to match with a mentor outside of their identity group.

Encouragingly, Proposition 3 shows that homophily does not need to have a distortionary impact on the profession: no matter how large η is, a steady state exists with equal representation and equal quality. While this may give the impression that homophily will not distort the profession, note that Proposition 3 shows that unequal representation is also a steady state of the model with homophily and, as we will show, transitioning to equal representation is no longer possible in equilibrium.

That is, as shown in the following corollary, for any $\eta > 0$, if the profession is at a steady state with $\bar{Q}^A = \bar{Q}^B$ and $M_t^A < M_t^B$, then in all period equilibria after time t, $M_{t+1}^A < M_{t+1}^B$.

Corollary 1 (Persistence of under-representation)**.** *If the profession is at a steady state with $\bar{Q}^A = \bar{Q}^B > 0$ and $\eta > 0$, then $M^I_t = M^I_{t+1}$ in all period equilibria.*[11]

Proposition 1 shows that for any $\eta > 0$, no matter how small, if the profession starts at a steady state with $M^A < M^B$, that imbalance will persist in perpetuity. The intuition for the persistence of under-representation is the fact that for representation of type A to increase in period $t+1$, it must be the case that a higher proportion of high-talent juniors of type A entered in period t. However, this cannot be a period equilibrium since it implies that (1) some high-talent juniors of type A match with mentors of type B, or (2) high-talent juniors of type A have a lower probability of matching with high quality mentors. If either (1) or (2) is true, then the expected value of entry is lower for high-talent juniors of type A than for high-talent juniors of type B, which violates the condition for a period equilibrium.

Importantly, Corollary 1 shows that if one of the two types is underrepresented due to, say, historical discrimination in the profession, then underrepresentation will persist even after discrimination is removed from the profession—i.e. to transition to equal representation, a policy intervention will be necessary. In the following section we consider using identity-based quotas as a policy to transition the profession to equal representation.

4 Using Quotas to Achieve Equal Representation

In this section we will analyze the effectiveness of quotas as a policy tool to transition to a steady state with equal representation and equal quality. Note that we do not consider an explicit welfare objective or explicitly model the benefits of equal representation. However, beyond fairness concerns, we emphasize that the literature has highlighted many potential benefits of equal representation in high-skill professions (see Auriol et al., 2022).

[11] This corollary follows directly from Lemma 4, which is presented in the proof of Proposition 3.

We focus on transitions from an initial point of equal quality ($\bar{Q}^A = \bar{Q}^B$) that corresponds to a steady state of the model with $\eta = 0$, $\bar{Q}^* > 0$, that is asymptotically stable. We discuss the stability of steady states in the appendix and show that, generically, at least one interior steady state is stable. Moreover, this assumption is not restrictive since always exists a transition path from an unstable steady state to a stable steady state.

A natural starting point is a quotas on underrepresented juniors, which we analyze in the following section. However, we find that a quota on juniors is ineffective is due the familiar problem of selection, and therefore subsequently analyze the effectiveness of a quota on underrepresented seniors. (We also introduce the impact of a quota on perceived quality in Section 4.2.) Lastly, we discuss the implications of our analysis for suggested policy measures aimed at increasing the representation of underrepresented types in academia.

4.1 Analysis of quotas on underrepresented juniors

Here we consider a quota on juniors. In particular, we consider the effectiveness of instituting a 1:1 quota on juniors as a policy for achieving a transition to equal representation at the senior level. We show that while such a policy will mechanically equalize the number of juniors of each identity type, it will be surprisingly ineffective when it comes to equalizing representation at the *senior level*. The reason a quota on juniors is ineffective is due the familiar problem of selection—simply put, a quota on juniors will not necessarily increase the proportion of high-talent juniors of the under-represented type.

Formally, we model a 1:1 quota as two separate entry lotteries for types A and B, where all applicants of type A are randomly selected to fill $\lambda/2$ junior slots, and all applicants of type B are randomly selected to fill the remaining $\lambda/2$ slots. Surprisingly, as illustrated in the following result, unless the quotas are very high, there is no impact of junior quotas on representation.

Proposition 4 (Ineffectiveness of junior quota)**.** *If the profession is at a steady* *state* $\{\mathbf{M}^{A*}, \mathbf{M}^{B*}\}$ *with* $\bar{Q}^{A*} = \bar{Q}^{B*} = \bar{Q}^*$, $Q_L^* > 0$ *and* $m_h^{B*} \leq \lambda/2$, *then* $\{\mathbf{M}^{A*}, \mathbf{M}^{B*}\}$ *remains a steady state under a 1:1 quota on juniors.*

The intuition for this result is as follows: A quota on juniors of type A does not increase the relative attractiveness of entry for the high-talent juniors of type A since the number of highest-quality mentors of type A (M_H^A) remains unchanged. Since the number of high-talent juniors that enter depends only on the relative utility of entry (rather than the probability of entry), a 1:1 quota on juniors will not result in more entry of *high-talent* juniors of the underrepresented type. That is, if type A is underrepresented ($M^A < M^B$) at the initial steady state, then the junior quota will be filled by low-talent juniors of type A, who are not hired by the profession as seniors and therefore do not impact representation at the senior level.[12]

4.2 Analysis of quotas on underrepresented seniors

As discussed in the introduction, the most relevant setting for our analysis is a situation where quotas impact the perceived quality of the underrepresented type, but do not have a direct impact on real quality. A conceptual motivation is a situation where quotas counteract underlying bias and hence do not impact quality, but where career entrants are unaware of the underlying bias and hence perceive quotas as reducing real quality. Accordingly, we model the impact of quotas *as if* they result in a lower quality threshold for the underrepresented type.[13] Therefore, in our analysis, quotas impact the decisions of career entrants through the perceived quality of the seniors of the underrepresented type.

[12] A quota on juniors can only have an impact on the profession if (1) it decreases the number of high-talent juniors of the over-represented type, which only happens if $m_h^B > \lambda/2$, or (2) if the profession hires seniors with $Q_i = 0$, which is only the case if $\bar{Q} < 1/2$.

[13] Modeling perceived quality this way also ensures that the impact of a quota on perceived quality is proportional to the size of the quota relative to the size of trained juniors of the underrepresented type—a non-binding quota has no impact of perceived quality, while a larger quota will have a larger impact on perceived quality.

For simplicity, we also model the production of quality, $f(\bar{Q}^m)$, as a function of perceived quality. This is a simplification that allows us to avoid separate notation and accounting for perceived and real quality. This assumption implies that the impact of quotas on quality is overestimated since it will impact quality through two channels: 1) the decisions of career entrants, and 2) the production of quality. However, from the perspective of our analysis this is a conservative assumption, and as we discuss below our results are robust to removing channel (2)—importantly, the dynamic path that we highlight below that transitions to equal quality (perceived and real) also results in a transition if channel (2) is removed.

Here we consider a minimum quota on seniors of the underrepresented type as a method for achieving equal representation. We first consider the implementation of a 1:1 quota in period t and higher; i.e. the profession is constrained to hire $M^A_{t+1} = M^B_{t+1} = 1/2$. Note that a quota on seniors does not directly impact the entry decisions of the juniors, since trained juniors receive a wage equal to their realized senior quality regardless of whether they are hired by the profession. However, as discussed above, a quota will affect entry decisions indirectly through its impact on the perceived average senior quality.

In particular, career entrants hold the belief that given $\bar{Q}^A_{t-1} = \bar{Q}^B_{t-1}$, the distribution of senior quality in period t is uniform as characterized in the previous section. Therefore, if the profession starts from a point of underrepresentation, $M^A_t < M^B_t$, career entrants in period $t+1$ will believe that the institution of a 1:1 quota in period t will result a different perceived quality cutoffs for seniors of the different identity types to fill the quota; i.e. that $Q^A_{L,t} < Q^B_{L,t}$. This implies that the perceived quality of the underrepresented type is lower in period $t+1$ ($\bar{Q}^A_{t+1} < \bar{Q}^B_{t+1}$), and this will have an impact of the entry decision of career entrants in period $t+1$. Accordingly, as shown in the following lemma, a 1:1 quota on seniors will not result in 1:1 entry of juniors due to the perceived difference in quality of seniors of different identity types.

Lemma 3 (Crowding Out)**.** *If $\bar{Q}^A_t < \bar{Q}^B_t$, then for a junior of the overrepresented* *type, the probability of being matched with an underrepresented senior is strictly positive,* $\Pr(A|B) > 0$.

That is, juniors of the identity-group with higher perceived average quality will enter in a higher proportion relative to the proportion of senors of that identity group—i.e. there is a "crowding out" effect of the perceived quality difference. In fact, if the impact of the quota on perceived quality is high enough, the the crowding out effect can cause the profession can converge to an asymmetric steady state where high-talent juniors of the underrepresented type select out of the profession.

Proposition 5 (Adverse selection)**.** *If the profession is at a steady state* $\{\mathbf{M}^{A*}, \mathbf{M}^{B*}\}$ *with* $\bar{Q}^{A*} = \bar{Q}^{B*} = \bar{Q}^{*} > 0$*, then for all* $M^{A*} < M^{A'}$ *for some* $M^{A'} > 0$*, a 1:1 quota on seniors will result in a transition to* $\bar{Q}^{A} = 0$.

Proposition 5 follows from the fact that with homophily, the equilibrium entry decisions of high-talent juniors are driven by the average perceived quality of mentors in their identity-group. In particular, given Lemma 3 we can show that if perceived quality, $\bar{Q}^{A}_{t}$, is low enough, then all high-talent juniors of the underrepresented type select out of the profession since $g(\bar{Q}^{A}_{t}, P(A|B)m^{B}_{h,t}) - o_h < \beta\bar{Q}^{A}_{t} - o_l$. This effect is then permanent since if $m^{A}_{h,t} = 0$, then $\bar{Q}^{A}_{t+1} = 0$ and high-talent juniors of the underrepresented type will select out of the profession in all future periods.[14]

Proposition 5 demonstrates that attempting an abrupt transition to equal representation can have the unintended consequence of causing high-talent juniors of the underrepresented type to select out of the profession. Note that this result shows that a decrease in perceived quality can impact real quality: since all high-talent juniors of the underrepresented type will select out of the profession, the real quality of the seniors of the underrepresented type will converge to zero.

[14]Note that proposition 5 is robust to the assumption that the production of quality is a function of perceived quality: since all juniors of the underrepresented type select out of the profession, both real and perceived quality of underrepresented seniors will be equal to zero in $t+1$.

Encouragingly, however, our next results shows that a transition to equal representation and equal quality (real and perceived) is always feasible as long as the transition is gradual enough. We first prove the result for the extreme case of $\eta = \infty$.

Proposition 6 (Gradual Transition)**.** *If $\eta = \infty$ and the profession is at a steady state $\{\mathbf{M}^{A*}, \mathbf{M}^{B*}\}$ with $\bar{Q}^{A*} = \bar{Q}^{B*} = \bar{Q}^{*} > 0$ and $M^{A*} < M^{B*}$, then there exists a monotonically increasing sequence of quotas on seniors of type A, $\{\bar{M}_t\}$, that results in a convergence to a steady state with $\bar{Q}^{A} = \bar{Q}^{B} = \bar{Q}^{*}$ and $M^{A} = M^{B}$.*

Proposition 6 shows that when $\eta = \infty$, which implies that the two identity groups effectively function as independent professions, a transition is possible as long it is gradual. Note that since the dynamics of quality are independent of the size of the profession, and a small increase in $\{\bar{M}_t\}$ in each period corresponds to a small decrease in perceived quality ($\bar{Q}^{A}_{t+1}$). Moreover, given that $\bar{Q}^{*}$ is a steady-state in the "A-profession," small deviations in $\bar{Q}^{A}$ imply that the dynamics of the model point back to $\bar{Q}^{*}$. Therefore, a gradual increase in the quota, and correspondingly small increases $\{\bar{M}_t\}$, ensures that $\bar{Q}^{A}_{t}$ stays within the "basin of attraction" of $\bar{Q}^{*}$ along the whole path of transition. This implies that if the increase in the quota is gradual enough, then the profession will transition to both equal representation and equal quality and avoid the pitfall of adverse selection highlighted in Proposition 5.[15] Lastly, note that when the quota is non-binding, then perceived and real quality are equal: therefore, Proposition 6 characterizes a transition to both real and perceived quality.

Given the result of Proposition 6, it might be natural to assume that the same dynamic path of quotas will also result in a transition to equal representation and equal quality for any η. This, however, turns out to not always be the case due to the crowding out effect highlighted in Lemma 3. That is, the crowding out effect

[15] Proposition 6 is also robust to the assumption that the production of quality is a function of perceived quality since this overestimates the impact of the quota on $\bar{Q}^{A}_{t+1}$. Therefore, even if we assume that the production of quality is a function of real quality, then real and perceived quality will be bounded below by the sequence of $\bar{Q}^{A}_{t}$ characterized in the proof of Proposition 6, and the sequence of quotas will result in a transition to equal perceived and real quality.

of a quota may lead to a dynamic where instead of equalizing, $\bar{Q}_t^A$ and $\bar{Q}_t^B$ grow farther apart. From a technical perspective, the dynamics depends on the relative sensitivity to quality of the high and low-talent types at the steady state. This can lead to a failure of the quota to transition to equal representation and equal quality. However, there is a simple remedy: by instituting a quota on seniors *and* a corresponding quota on juniors, the crowding out effect can be eliminated.

Corollary 2 (Coordinated Quotas)**.** *For any* $\eta > 0$*, if the profession is at a steady state* $\{\mathbf{M}^{A*}, \mathbf{M}^{B*}\}$ *with* $\bar{Q}^{A*} = \bar{Q}^{B*} = \bar{Q}^* > 0$ *and* $M^{A*} < M^{B*}$*, then there exists a monotonically increasing sequence of quotas,* $\{\bar{M}_t^A, \bar{m}_t^A\}$*, that results in a convergence to a steady state with* $\bar{Q}^A = \bar{Q}^B = \bar{Q}^*$ *and* $M^A = M^B$*.*

The intuition for Corollary 2 is straightforward. By constraining the set of juniors to be proportional to the set of seniors (in terms of identity-types), quotas on juniors eliminate the crowding out effect of perceived differences in senior quality—essentially, the quota establishes separate professions for the two different identity types, which means that the result of Proposition 6 applies and a transition to equal representation and equal (real and perceived) quality can be achieved.

4.3 Discussion of proposed quotas

Lastly, we discuss the implications of our analysis for measures that have been proposed or implemented for addressing underrepresentation.

The cascade model: Many professions that exhibit underrepresentation at the senior level suffer from the so-called "leaky pipeline," where the level of representation is high at junior levels, but decreases in seniority. One example of this is academia, where even in fields that are close to parity at the undergraduate level, women are increasingly underrepresented at the level of PhD, Assistant, Associate and Full professor. The cascade model, used in Sweden and Germany (Wallon et al., 2015), is meant to address the leaky pipeline by setting a soft quota at each level of seniority that is equal to the level of representation at the level below.

In the context of our model, it is natural to interpret bachelor's students as career entrants, PhD students as juniors, and professor positions as seniors. In accordance with the leaky pipeline, consider a field that features equal representation at the bachelor's level, some underrepresentation at the PhD level, and higher underrepresentation at the professor level. Applied strictly, the cascade model would transition to equal representation at the professor level in two periods. Somewhat surprisingly given the result of Proposition 6, we find that the cascade model can result in a *lower perceived quality relative to instituting a 1:1 quota in the first period*.

To explain the intuition behind this result, note that in the first period the quota at the professor level will be binding and the perceived quality of the seniors of the underrepresented type will lower than in the previous period. This lowers the expected value of entry for bachelor's students of the underrepresented type. This implies that, despite the binding 1:1 quota on underrepresented PhD students, the total number of underrepresented high-talent PhD students may be lower than in the previous period.

In the second period, the profession will move to a 1:1 quota at the professor level given equal representation at the PhD level in the previous period. However, if there is a decrease in the total number of high talent PhD students of the underrepresented types that entered in the previous period, then the total number of seniors of the underrepresented type that realize high quality will be lower in the first period of the quota. Therefore, relative to the case where a 1:1 quota was instituted in the first period, the cascade model results in a *more* binding quota in the second period. This shows that the cascade model can result in a lower perceived quality relative to jumping straight to a 1:1 quota.

We present this result formally in the following proposition (the result extends straightforwardly to cascade quotas that take more than two periods to reach a 1:1 quota). Take $\bar{M}$ equal to a 1:1 quota in period t, and $\bar{M}'$ equal to a quota where

$\bar{M}'_{t+1}$ is equal to a 1:1 quota and $\bar{M}'_t < \bar{M}'_{t+1}$ is binding.

Proposition 7. *Take $\eta = \infty$ and a profession at a steady state $\{\mathbf{M}^{A*}, \mathbf{M}^{B*}\}$ with $\bar{Q}^{A*} = \bar{Q}^{B*} = \bar{Q}^* > 0$ and $M^{A*} < M^{B*}$. If $\beta(\bar{Q}^* - \bar{Q}^A_{t+1}) < f(\bar{Q}^*) - f(\bar{Q}^A_{t+1})$ under $\bar{M}'$, then perceived quality $\bar{Q}^A_{t+2}$ under $\bar{M}'$ is* lower *than $\bar{Q}^A_{t+1}$ under $\bar{M}$.*

Proposition 7 provides more detail about the sequence of quotas that will transition to equal representation and quality (Proposition 6): instead of a continuous increase, it may be beneficial to hold the quota constant at an intermediate level for a number of periods to allow the dynamics to move the perceived quality of seniors of the underrepresented type closer to $\bar{Q}^*$.

Tie-breaker models: A common measure used to address underrepresentation is to favor underrepresented candidates in cases of "equal quality," which we refer to as a tie-breaker quota. In our model, we look at a continuous quality distribution, in which case this quota would be non-binding, but our analysis extends naturally to a case where the observed signal of senior quality is coarse. In this case, favoring underrepresented seniors of equal quality will result in a binding quota and an decrease in perceived quality of the underrepresented type since the quota implies that all seniors with the marginal perceived quality are underrepresented seniors.

On one hand, a tie-breaker quota avoids the problem illustrated in Proposition 7 by construction, since the quota endogenously limits the difference in perceived quality between seniors of the two identity categories. That is, as long as the signal of quality is not too coarse, then both $\bar{Q}^A_t$ and $\bar{Q}^B_t$ will stay within the catchment area of $\bar{Q}^*$. On the other hand, we cannot exclude the possibility that a tie-breaker quota will lead to cycling rather than a transition to equal representation.

That is, while a tie-breaker quota will lead to an increase in M^A_t if it is introduced in period $t-1$, the impact on M^A_{t+1}, relative to M^A_{t-1}, is unclear due to the problem of crowding out. Since the quota implies that $\bar{Q}^A_t < \bar{Q}^B_t$, we know that $M^A_{t+1} < M^A_t$ by Lemma 3. Therefore, we cannot exclude the possibility that $M^A_{t+1} \leq M^A_{t-1}$ which could lead to a cycle about the original levels M^{A*}, M^{B*}.

Therefore, it may be possible that a tie-breaker quota could require an occasional "nudge"—a discrete increase in the number of underrepresented seniors—to put it on the path to transition to equal representation.

Quotas on employment committees: Another common measure meant to address underrepresentation indirectly is to specify identity quotas on employment committees and company boards (Wallon et al., 2015). The idea behind this measure is that increasing representation on employment committees will reduce bias in the hiring process and therefore increase representation in the profession. Unfortunately, the empirical evidence suggests that these measures have not worked as intended (see for example Bagues et al., 2017; Bertrand et al., 2018). Our model provides some insight as to what would be necessary for such a measure to be successful.

In particular, our analysis illustrates that addressing discrimination in the hiring process may not be sufficient to transition to equal representation due to the friction introduced by a preference for homophily. If underrepresented seniors hold the same beliefs regarding perceived quality as career entrants of the underrepresented identity, then underrepresentation will be persistent as illustrated in Proposition 1. Instead, for a quota on hiring committees to have an impact, it must be the case that seniors of the underrepresented types must have more accurate beliefs than career entrants. Additionally, since increasing hiring of underrepresented seniors may decrease perceived quality in the short run, the career incentives of seniors of the underrepresented type may not be aligned with boosting hiring of underrepresented seniors in the short run to achieve the long-term goal of equal representation. Therefore, quotas on hiring committees may not translate into increased representation, and may need to be combined with an explicit quota, such as a tie-breaker quota, to be effective.

5 Conclusion

In this paper we explore the trade-off between representation and perceived quality, and establish a dynamic argument for quotas to correct for underrepresentation even if they introduce stigma in the short run. Our research provides also provides important insights into policy measures. First, it is not sufficient to institute a quota at the junior level due to an adverse selection problem. Instead, a quota at the senior level is also necessary. Second, also due to adverse selection, an abrupt transition to equal representation can cause a permanent real quality difference between seniors of the two identity types, while a gradual transition can result in a equal representation and quality. While this may result in a slight decrease in the average perceived quality in the profession in the short run, this decrease is temporary and transitions the profession to a point of equal representation and equal quality.

Moreover, we show that a "cascade model," where employment at the senior level is equalized to the identity proportions at the junior level, can be counterproductive relative to a quick transition to equal representation. Instead, a preference for underrepresented seniors in the case of equal quality seems preferable, although it may require additional nudges to transition all the way to equal representation.

Our research also suggests that an important avenue for future research on role models is to explore the interaction of identity and quality. That is, while the empirical literature has focused on characterizing the impact of the identity of role models on educational and career choices, the impact of the quality of role models on choices is largely unexplored. Our research suggests that while more role models of an underrepresented type may be an important factor in combating underrepresentation, such a strategy could backfire even if role models of the underrepresented type are *perceived* to be of lower quality by potential entrants. Therefore, our research highlights that empirical evidence on the interaction of identity-quotas, career choice and perceived quality is essential when it comes to addressing the effectiveness of policy to achieve equal representation.

References

Athey, S., Avery, C., and Zemsky, P. (2000). Mentoring and diversity. *American Economic Review*, 90(4):765–786.

Auriol, E., Friebel, G., Weinberger, A., and Wilhelm, S. (2022). Underrepresentation of women in the economics profession more pronounced in the united states compared to heterogeneous europe. *Proceedings of the National Academy of Sciences*, 119(16):e2118853119.

Bagues, M., Sylos-Labini, M., and Zinovyeva, N. (2017). Does the gender composition of scientific committees matter? *The American Economic Review*, 107(4):1207–1238.

Battaglini, M., Harris, J., and Patacchini, E. (2022). Professional interactions and hiring decisions: Evidence from the federal judiciary. *Journal of Labor Economics*, Forthcoming.

Bertrand, M., Black, S. E., Jensen, S., and Lleras-Muney, A. (2018). Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway. *The Review of Economic Studies*, 86(1):191–239.

Besley, T., Folke, O., Persson, T., and Rickne, J. (2017). Gender quotas and the crisis of the mediocre man: Theory and evidence from sweden. *American Economic Review*, 107(8):2204–42.

Carvalho, J.-P. and Pradelski, B. (2022). Identity and underrepresentation: Interactions between race and gender. *Journal of Public Economics*, 216:104764.

Coate, S. and Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review*, 83(5):1220–1240.

European Commission (2019). She figures 2018. Technical report, Directorate-General for Research and Innovation.

Fang, H. and Moro, A. (2011). Chapter 5 - theories of statistical discrimination and affirmative action: A survey. volume 1 of *Handbook of Social Economics*, pages 133–200. North-Holland.

Fershtman, D. and Pavan, A. (2021). “soft” affirmative action and minority recruitment. *American Economic Review: Insights*, 3(1):1–18.

Gaule, P. and Piacentini, M. (2018). An advisor like me? advisor gender and post-graduate careers in science. *Research Policy*, 47(4).

Heilman, M., Block, C., and Lucas, J. (1992). Presumed incompetent? stigmatization and affirmative action efforts. *Journal of Applied Psychology*, 77(4):536–544.

Hilmer, C. and Hilmer, M. (2007). Women helping women, men helping women? same-gender mentoring, initial job placements, and early career publishing success for economics phds. *American Economic Review*, 97(2):422–426.

Leslie, Lisa, M. D. K. D. (2014). The stigma of affirmative action: A stereotyping-based theory and meta-analytic test of the consequences for performance. *The Academy of Management Journal*, 57(4):964–989.

Müller-Itten, M. and Öry, A. (2022). Mentoring and the dynamics of affirmative action. *American Economic Journal: Economic Policy*, 14(2):402–44.

Porter, C. and Serra, D. (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, 12(3):226–54.

Riise, J., Willage, B., and Willén, A. (2022). Can Female Doctors Cure the Gender STEMM Gap? Evidence from Exogenously Assigned General Practitioners. *The Review of Economics and Statistics*, 104(4):621–635.

Robin, C. (2019). Clarence thomass radical vision of race. *The New Yorker Magazine*.

Siniscalchi, M. and Veronesi, P. (2020). Self-image bias and lost talent. Working Paper 28308, National Bureau of Economic Research.

Wallon, G., Bendiscioli, S., and Garfinkel, M. (2015). Exploring quotas in academia. Technical Report 12.5.8050.0239.0, Robert Bosch Stiftung.

6 Appendix: Proofs

Proof Lemma 1: The best response functions specify that for interior equilibria $(\sigma_l, \sigma_h > 0)$:

$$E[u_i|\hat{a}_i = 1, \bar{Q}, q_i, \sigma_{-i}] = o_{q_i}.$$

Let p be the the probability of entry given σ_l, σ_h. This allows us to rewrite the condition for an interior equilibrium as a function of the expected wage conditional on entry:

$$pE[w_i|a_i = 1, \bar{Q}, m_h, q_i] - (1-p)o_{q_i} - c = o_{q_i},$$

which simplifies to:

$$E[w_i|a_i = 1, \bar{Q}, m_h, q_i] - o_{q_i} = \frac{c}{p}.$$

Since the right-hand side of the expression above is the same for high and low-talent juniors, the left-hand side must be equal for $q_i = l, h$ in an interior equilibrium.

$$E[w_i|a_i = 1, \bar{Q}, m_h, l] - o_l = E[w_i|a_i = 1, \bar{Q}, m_h, h] - o_h. \tag{8}$$

Before proceeding with the proof, we show the following corollary:

Corollary 3. *$m_h \geq \lambda M_H$ in all equilibria.*

Corollary 3 follows from (3) and the assumption that $\beta - o_l < 1/2 - o_h$: if $m_h < \lambda M_H$, then all high type juniors who enter will be matched to the highest-quality group and realize an expected payoff of $1/2 - o_h$. That is, the expected utility conditional on entry for the high type is greater than for the low type ($\beta\bar{Q} - o_l \leq \beta - o_l$), which violates (3). Therefore, in equilibrium, m_h must be greater or equal to λM_H.

This implies that no low-talent juniors will match with the group of highest quality mentors and the expected wage for a low type is equal to:

$$E[w_i|\bar{Q},l] = \beta\bar{Q}, \tag{9}$$

which gives the equilibrium condition:

$$\beta\bar{Q} - o_l = g(\bar{Q}, m_h) - o_h.$$

□

Proof of Proposition 1: In the main text we establish that an interior period equilibrium exists if and only if $m'_h \in (0,\lambda)$ where m'_h is defined in (3). We complete the proof by showing that a corner equilibrium exists with $m^*_h = 0$ if and only if $m'_h \leq 0$, and with $m^*_h = \lambda$ iff $m'_h \geq 0$.

First, take $m'_h \leq 0$ and assume an equilibrium exists with $m^*_h > 0$. Note that $m'_h \leq 0$ implies that:

$$\beta\bar{Q} - o_l \geq g(\bar{Q}, 0) - o_h.$$

Since $g(\bar{Q},0)$ is strictly decreasing in m_h, this equation shows that at m^*_h, $\beta\bar{Q} - o_l < g(\bar{Q}, m^*_h) - o_h$. Therefore, m^*_h cannot be an equilibrium. However, since $\beta\bar{Q} - o_l > 0$ for all $\bar{Q}$, $m^*_h = 0$, $m^*_l = \lambda$ is an equilibrium.

The proof for $m'_h \geq 0$ is analogous. Since $\beta\bar{Q} - o_l \leq g(\bar{Q}, 1) - o_h$, this expression also holds for all $m_h < \lambda$, implying that $m^*_h = \lambda$, $m^*_l = 0$ is the unique equilibrium. □

Proof of Proposition 2: First, note that if $\bar{Q}_t = 0$, then it is a best response for all high-type juniors to set $\hat{a}_i = 0$ since the relative expected utility of applying is negative: i.e. $o_h > 1/2f(0) - c$. Since $o_l + c < 0$, however, the relative expected utility of applying is positive for the low type if the probability of entry is equal to one. Therefore, $m_{h,t} = 0$ and $m_{l,t} = \lambda$ in any period equilibrium, and $\bar{Q}_{t+1} = 0$.

Second, trivially, $\bar{Q}^*$ is not a steady state if $\bar{Q}_{t+1}(\bar{Q}^*) \neq \bar{Q}^*$. If $\bar{Q}_{t+1}(\bar{Q}^*) = \bar{Q}^*$, however, then $\bar{Q}^*$ is a steady state since each $\bar{Q}$ is associated with a unique distribution of Q_j (uniform between $[Q_L, 1]$). □

Proof of Lemma 2: The proof follows directly from the proof of Lemma 1.

Proof of Proposition 3: First, note that the result is trivial for $\bar{Q} = 0$ since given $\bar{Q}_t = 0$, only low-talent juniors will apply and in equilibrium they will apply in proportion to M_t^A and M_t^B. Therefore, $\bar{Q}_{t+1} = 0$ and $M_t^A = M_{t+1}^A$.

For $\bar{Q}^A = \bar{Q}^B > 0$, we first introduce the following result characterizing period equilibria:

Lemma 4. *If $\bar{Q}^A = \bar{Q}^B > 0$, then in equilibrium:*

$$\frac{m_h^A}{m_h^B} = \frac{m_l^A}{m_l^B} = \frac{M^A}{M^B}.$$

That is, juniors will enter in the same identity-proportion as the proportion of mentors.

This lemma follows from Lemma 2: given the equal quality in both identity groups, it cannot be the case that any juniors face a positive probability of matching with an out-group mentor, since that would imply a lower expected utility of entry for that group. Moreover, $\frac{m_h^A}{m_h^B} = \frac{m_l^A}{m_l^B}$, since otherwise the high-type of one identity group would have a higher probability of matching with a high-quality mentor. ■

Lemma 4 shows that given $\bar{Q}_t^A = \bar{Q}_t^B$, $\Pr(I_{\neg i}|I_i) = 0$ in equilibrium for both identity types. This implies that the following equation defines an interior solution for m_h^{I*}:

$$m_h^{I'} = \lambda M_H^I \left[\frac{1 - f(\bar{Q})}{(\beta\bar{Q} - o_l) - (f(\bar{Q}) - o_h)} \right], \tag{10}$$

That is, $\frac{m_{h,t}^A}{m_{h,t}^B} = \frac{M_{H,t}^A}{M_{H,t}^B}$. In turn, this shows that if $\bar{Q}_t^A = \bar{Q}_t^B = Q^*$, where Q^* corresponds to a steady state, then $\bar{Q}_{t+1}^A = \bar{Q}_{t+1}^B = Q^*$, since $m_{h,t}^A$ and $m_{h,t}^B$ are both proportional to the size of the sets of mentors, M_t^A and M_t^B. □

Next we briefly address the existence of a stable interior steady state. Visually, note that interior steady states exist at points where $\bar{Q}_{t+1}(\bar{Q})$ either cross or are tangent to the 45 degree line. Moreover, $\bar{Q}_{t+1}(\bar{Q})$ is below the 45 degree line at both endpoints, 0 and 1, by assumption. It follows by the continuity of $\bar{Q}_{t+1}(\bar{Q})$, that as long as the function crosses the 45 degree line at some interior point, a steady state exists that where $\bar{Q}_{t+1}(\bar{Q})$ crosses the 45 degree line from above.

This gives the following result which completes the proof:

Result 1. *If $\bar{Q}_{t+1}(\bar{Q}) > \bar{Q}$ for some $\bar{Q} \in (0,1)$, then a stable steady state exists.*

Proof: Note that by the argument above, $\bar{Q}_{t+1}(\bar{Q}) > \bar{Q}$ for some $\bar{Q} \in (0,1)$ implies a steady state exists that where $\bar{Q}_{t+1}(\bar{Q})$ crosses the 45 degree line from above at some $\bar{Q}'$. This implies that the linearization of $\bar{Q}_{t+1}(\bar{Q})$ at $\bar{Q}'$ has a strictly negative slope, which implies that the eigenvalue criterion for a stable steady state is satisfied. □

Proof of Proposition 4: First, note that if $M^{A*} = M^{B*}$ then the 1:1 quota is redundant. Therefore, assume without loss of generality that $M^{A*} < M^{B*}$ and that the quota is initiated in period t. Therefore, a quota implies that with some probability each junior of type A will match with a mentor of type B.

$$\Pr(A|B) = \frac{\lambda/2 - \lambda M^A}{\lambda/2}$$

The expected utility of type A juniors conditional on entry in period t is:

$$(1 - \Pr(A|B))E[w_i|\bar{Q}_t^A, q_i] + \Pr(A|B)(E[w_i|\bar{Q}_t^B, q_i] - \eta).$$

However, since $\bar{Q}_t^A = \bar{Q}_t^B = \bar{Q}^*$ in period t, this expression simplifies to:

$$E[w_i|\bar{Q}^*, q_i] - \Pr(A|B)\eta.$$

That is, relative to no quota, in equilibrium the expected utility conditional on entry is lower by $\Pr(A|B)\eta$ for *both* the high and low type.

Therefore, $\Pr(A|B)\eta$ cancels out of the equilibrium condition listed in Lemma 2, and $m^A_{h,t}$ is still characterized by Equation 10 (in the proof of Proposition 3 above), which shows that the equilibrium level of $m^A_{h,t}$ is unchanged by the quota.

Next, note that $m^B_{h,t}$ is also unchanged by the quota by the same argument, and by the fact that $m_h^{B*} \leq \lambda/2$ (i.e. the size of the set of high-talent juniors of type B is lower than the quota at the steady state).

Since $m^A_{h,t} = m_h^{A*}$ and $m^B_{h,t} = m_h^{B*}$ the quota does not impact the set of juniors that realize positive quality in period t. Lastly, since $\bar{Q}^* \geq 1/2$, only mentors with strictly positive quality are hired by the profession ($Q_L > 0$). And since the set of juniors that realize positive quality in period t (with the quota) is identical to $t-1$ (without the quota), $Q_{L,t}$ will also be unchanged, and $\mathbf{M}^A_t = \mathbf{M}^{A*}$ and $\mathbf{M}^B_t = \mathbf{M}^{B*}$.
□

Proof of Lemma 3: First take the case of $m_l = 0$ and assume $\Pr(A|B) = 0$. In this case, both high-talent types must enter, and the following equation holds by Lemma 2:

$$\Pr(B|A)\left(g(\bar{Q}^B, m_h) - \eta\right) + (1 - \Pr(B|A))\left(g(\bar{Q}^A, m_h)\right) = g(\bar{Q}^B, m_h).$$

Since $m_l = 0$ and only high-talent juniors enter, $g(\bar{Q}^B, m_h) > g(\bar{Q}^A, m_h)$ which implies that $\Pr(A|B)$ must be strictly greater than zero for the above equation to hold.

Since $\Pr(A|B) > 0$, we can use the equilibrium condition in Lemma 2 to get the following expression for $\Pr(A|B)$:

$$\Pr(A|B) = \frac{\beta(\bar{Q}^B - \bar{Q}^A)}{\beta(\bar{Q}^B - \bar{Q}^A) + \eta},$$

Next, take the case of $m_l > 0$ and $\Pr(A|B) = 0$. First, consider $\beta\bar{Q}^B - \eta < \beta\bar{Q}^A$. In this case, both low-talent types enter if $m_l > 0$, and each type prefers to be matched to an own-type mentor. Therefore, the following equilibrium condition must hold by Lemma 2:

$$\Pr(B|A)\left(\beta\bar{Q}^B-\eta\right)+(1-\Pr(B|A))\left(\beta\bar{Q}^A\right)=\beta\bar{Q}^B.$$

This is a contradiction since $\bar{Q}^A<\bar{Q}^B$, showing that $\Pr(A|B)$ must be strictly positive.

Solving for $\Pr(A|B)$ as above gives:

$$\Pr(A|B)=\frac{g(\bar{Q}^B,m_h)-g(\bar{Q}^A,m_h)}{g(\bar{Q}^B,m_h)-g(\bar{Q}^A,m_h)+\eta},$$

which is strictly greater than 0.

Lastly, assume that $\beta\bar{Q}^B-\eta\geq\beta\bar{Q}^A$, $m_l>0$ and $\Pr(A|B)=0$. In this case, low-types of both identities prefer to be matched with a B mentor, and therefore conditional on entry have the same probability of being matched with a B mentor. However, this implies that the above equilibrium condition cannot hold for low-talent junior, since the relative expected utility of entry is strictly higher for B-type juniors, which implies that $m_l^A=0$.

This, however, implies that the following equilibrium condition cannot hold:

$$\Pr(B|A)\left(g(\bar{Q}^B,m_h)-\eta\right)+(1-\Pr(B|A))\left(g(\bar{Q}^A,m_h)\right)\geq g(\bar{Q}^B,m_h).$$

Since $g(\bar{Q}^B,m_h)>g(\bar{Q}^A,m_h)$ given that $\bar{Q}^B>\bar{Q}^A$, and only high-talent juniors match with mentors of type A.

□

Proof of Proposition 5: By Lemma 3, if $\bar{Q}_t^A<\bar{Q}_t^B$, then $P(A|B)_t>0$ in equilibrium. This implies that a proportion of high-talent juniors of type B, $P(A|B)_t m_{h,t}^B$, will match with mentors of type A. Next, note that if $P(A|B)_t m_{h,t}^B$ is high enough relative to $M_{H,t}^A$, then the relative expected utility of entry for high-talent juniors of type A will be lower than $\beta\bar{Q}_t^A-o_l$ due to the crowding out effect. If this occurs, then $m_{h,t}^A=0$ in equilibrium—high-talent juniors of type A are crowded out—which implies that $\bar{Q}^A_{t+1} = 0$. The result then follows by induction since $m^A_{h,t+1} = 0$ if $\bar{Q}^A_{t+1} = 0$.

To complete the proof, we show that if $M^{A*} < M'$ for some M', then $g(\bar{Q}^A_t, P(A|B)_t m^B_{h,t}) - o_h < \beta\bar{Q}^A_t - o_l$. First, note that $m^B_{h,t}$ is bounded from zero given $\bar{Q}^B_t > 0$, which implies that $P(A|B)_t m^B_{h,t} > \delta$ for some $\delta > 0$ for all values of M^{A*}.

Next, $\bar{Q}^A_t \to 0$ as $M^{A*} \to 0$, which implies that:

$$\lim_{M^{A*}\to 0} P(A|B) \to \frac{\beta\bar{Q}^B_t}{\beta\bar{Q}^B_t + \eta},$$

which is strictly greater than zero for all η. Lastly, since $g(0,\delta) - o_h < -o_l$, it follows that $g(\bar{Q}^A_t, \delta) - o_h < \beta\bar{Q}^A_t - o_l$ for all M^{A*} that are small enough, which completes the proof. □

Proof of Proposition 6: We begin by showing that if $\bar{M}_t = \bar{M}_{t+1} = \bar{M}$ and $\eta = \infty$, then the dynamics of $\bar{Q}^A$ can be characterized by Equation 7. That is, $\bar{Q}^A_{t+1}(\bar{Q}^A_t) = \bar{Q}_{t+1}(\bar{Q}^A_t)$. Essentially, this is the same as proving that the dynamics of the profession are invariant to the size of the profession.

First, note that $P(A|B)$ is equal to zero in equilibrium if $\eta = \infty$, which means that the results of Lemma 1 apply and the following expression characterizes an interior value of m^{A*}_h:

$$m^{A'}_h = \lambda M^A_H \left[\frac{1 - f(\bar{Q}^A)}{(\beta\bar{Q}^A - o_l) - (f(\bar{Q}^A) - o_h)}\right], \tag{11}$$

which shows that Proposition 1 also applies.

Note that we wish to compare the dynamics of $\bar{Q}^A_t$ to the dynamics of $\bar{Q}_t$ in a profession without homophily at a point with $\bar{Q}_t = \bar{Q}^A_t$. At this point $M^A_{H,t} = \bar{M}M_{H,t}$, and Expression 11 gives us the following expression for $m^{A'}_{h,t}$:

$$m^{A'}_{h,t} = m'_{h,t}\bar{M},$$

where $m'_{h,t}$ is the interior value of m^*_h in the profession without homophily and

$\bar{Q} = \bar{Q}^A$.

Next, note that $g(\bar{Q}_t^A, m_{h,t}^A) = g(\bar{Q}_t^A, m_{h,t})$ since:

$$\begin{aligned} g(\bar{Q}_t^A, m_{h,t}^A) &= \frac{\lambda M_{H,t}^A}{m_{h,t}^A} + \left(1 - \frac{\lambda M_{H,t}^A}{m_{h,t}^A}\right) f(\bar{Q}_t^A) \\ &= \frac{\lambda M_{H,t}\bar{M}}{m_{h,t}\bar{M}} + \left(1 - \frac{\lambda M_{H,t}\bar{M}}{m_{h,t}\bar{M}}\right) f(\bar{Q}_t^A) \\ &= g(\bar{Q}_t^A, m_{h,t}). \end{aligned}$$

Lastly, using the same steps we used to derive $\bar{Q}_{t+1}(\bar{Q}_t)$, we get:

$$\bar{Q}_{t+1}^A(\bar{Q}_t^A) = \frac{\bar{M}}{2m_{h,t}^A g(\bar{Q}_t^A, m_{h,t}^A)} = \frac{1}{2m_{h,t} g(\bar{Q}_t^A, m_{h,t})} = \bar{Q}_{t+1}(\bar{Q}_t^A).$$

That is, given a constant quota and $\eta = \infty$, the dynamics of $\bar{Q}^A$ are equivalent to the dynamics of $\bar{Q}$ given $\eta = \infty$.

In turn, this shows that given a constant quota, $\bar{M}$, $\bar{Q}^A$ is asymptotically stable at $\bar{Q}^A = \bar{Q}^*$, which allows us to characterize the following dynamic path of quotas that transition to $M^A = M^B$ and $\bar{Q}^A = \bar{Q}^*$. First, take $\bar{Q}'$ such that $\bar{Q}' < \bar{Q}^*$ and $|\bar{Q}', \bar{Q}^*| < \delta$, where $\delta > 0$ is small enough so that $\lim_{t\to\infty} \bar{Q}_t = \bar{Q}^*$. Next, take $\bar{Q}'' \in (\bar{Q}', \bar{Q}^*)$, and take n to equal the number of periods it takes the profession to transition $\bar{Q}^A$ from $\bar{Q}'$ to a point greater or equal to $\bar{Q}''$ (n is finite since $\bar{Q}^*$ is asymptotically stable).

The following algorithm results in a transition:

1. At $t = 0$, set $\bar{M}_0$ so that $\bar{Q}_1^A = \bar{Q}'$ if this implies $\bar{M}_0 < 1/2$. Otherwise set $\bar{M}_t = 1/2$ for all t.

2. Set $\bar{M}_t = \bar{M}_0$ for n periods.

3. At $t = n+1$, repeat step 1-2 and continue until $\bar{M}_0 \geq 1/2$.

Note that this algorithm will result in a transition to $M^A = M^B$ and $\bar{Q}^A = \bar{Q}^*$, but could result in a transition to $\lim_{t\to\infty} \bar{Q}_t^B > \bar{Q}^*$. However, since $\bar{Q}^*$ is stable from above and below, a transition to $M^A = M^B$ and $\bar{Q}^A = \bar{Q}^A = \bar{Q}^*$ can be achieved for a low enough δ. □

Proof of Corollary 2: Note that if the quota on mentors is gradual enough so that all juniors prefer to match with mentors of the same identity (which is effectively a restriction on δ in the proof of Proposition 6), and a quota on juniors is set so that the following holds for all t:

$$\frac{m_t^A}{m_t^B} = \frac{M_t^A}{M_t^B}.$$

Then $P(A|B)_t = 0$ for all t, and the result of Proposition 6 applies straightforwardly.

Proof of Proposition 7: Note that under $\bar{\mathbf{M}}'$, $m_{h,t}^A < m_{h,t-1}^A$ by Equation 3 given the condition that $\beta(\bar{Q}^* - \bar{Q}_{t+1}^A) < f(\bar{Q}^*) - f(\bar{Q}_{t+1}^A)$. The result then follows since a 1:1 quota will have a larger impact on perceived quality in period $t+1$ under $\bar{\mathbf{M}}'$ since it is more binding than the 1:1 quota in period t under $\bar{\mathbf{M}}$. □