Minority Traps

Luís Cabral
New York University

July 2022

Abstract. I develop the idea that minority representation in certain fields (e.g., women in STEM) results from a sort of a chicken-and-egg "equilibrium" — a "minority trap." Formally, I model the evolution of a profession as a birth-and-death Markov process. I assume there is symmetry in abilities and preferences and that selection is purely meritocratic. However, minorities must pay a (material or psychological) cost $c(s)$, a positive and decreasing function of the minority share s. I calibrate the model with data from STEM doctoral programs and show that the stationary distribution is symmetric but multi-modal. This implies that, absent intervention, the system remains for a long time in a set of asymmetric states. The model provides a basis for a (temporary) affirmative action policy that "nudges" the system toward the central mode of the stationary distribution. This implies short term costs but longer term higher productivity levels.

Paganelli-Bull Professor of Economics, Stern School of Business, New York University; and Research Fellow, CEPR; luis.cabral@nyu.edu.

1. Introduction

Women are underrepresented in a variety of professions and academic fields. There is no shortage of explanations for such an asymmetric outcome, including explicit or implicit discrimination, differences in preferences, and differences in ability. In this paper, I focus on one specific possibility, namely the possibility that minorities are not attracted to certain fields because they are in a minority to begin with. In other words, the possibility that minority representation is a sort of a chicken-and-egg equilibrium — a minority trap.[1]

Formally, I model the evolution of a profession as a birth-and-death (BD) Markov process. In each period, one worker retires and two candidates, one of each type (x and y), apply for the position. A key assumption is that the net utility (and productivity) of a minority worker is lower than gross utility by a cost $c(s)$ which is a positive and decreasing function of the minority share s. Other than this loss function, I *assume* symmetry in abilities and preferences. In other words, I focus on one single source of asymmetry in outcomes. The cost function $c(s)$ may be interpreted as a psychological or a material cost that minority members must incur (e.g., lack of mentorship or lack role models), or as the reduced-form of a mechanism that keeps minorities away from the profession (e.g., differential treatment across genders).

I show that the Markov process governing the share of each type is ergodic, that is, it converges to a unique stationary distribution regardless of the starting state. Moreover, consistent with symmetry of the model's primitives, the stationary distribution is symmetric. However, I also show that, if the minority cost is sufficiently high, then the stationary distribution is multi-modal: in addition to the mode at 50%, there exist modes where one type of agent is in a minority. Ergodicity notwithstanding, a multi-modal stationary distribution implies that it makes a difference where the system starts. Specifically, it takes a "long" time before the current minority's share reaches 50%. A simple calibration based on estimates of peer effects in STEM doctoral programs suggests that this is more than a simple theoretical possibility.

My model of ergodic dynamics with "quasi path dependence" suggests a role for affirmative action as a (temporary) means to nudge the system towards the central mode of the stationary distribution. I consider two alternative policies. I show that the policy of biasing selection in favor of minority candidates may backfire: it leads to a cost in terms of productivity (selection of lower ability candidates) and has no significant effect on moving the system towards symmetry. By contrast, a policy of subsidizing minority candidates in a way that compensates for the minority cost $c(s)$, while maintaining a meritocratic selection mechanism, does move the system to symmetry at a faster rate. It implies a loss in the short term (the cost of subsidies) but leads to higher productivity in the long term, due to selection from a larger pool of able and willing job candidates (Hsieh et al., 2019).

■ **Related literature.** I am by no means the first to propose a theoretical approach to the issue of asymmetric representation in specific fields. In several ways, my paper is closest to Müller-Itten and Öry (2022). They consider the dynamics of workforce composition in an overlapping-generations model with mentoring effects. Under certain conditions, the stationary equilibrium is asymmetric. By contrast, the equilibrium of the system I consider is symmetric, though the *state* of the system may be (temporarily) asymmetric. Therefore, I am more interested in time paths than in the stationary equilibrium per se. As well, the type of affirmative action policy I consider is of a different nature (in particular, it is temporary rather than permanent).[2]

There are other lines of research that I will make reference to throughout the paper. In addition to other theory papers, there is a growing empirical literature measuring gender composition effects, most of which I will refer to in the next section. From a modeling point of view, the paper is related to a literature on Markov processes with inertia and hystereses (and multi-modal stationary distributions). I will refer to this literature in Section 4.

1. I borrow the term "minority trap" from Shan (2022).
2. Other, somewhat less related, theory papers include Athey, Avery, and Zemsky (2000), Chung (2000), Fershtman and Pavan (2021).

Table 1
State transition starting from state n

Retiree	Entrant	New State	Probability
y	x	$n+1$	$(1-s)\,\rho(s)$
x	y	$n-1$	$s\,(1-\rho(s))$
x	x	n	$s\,\rho(s)$
y	y	n	$(1-s)\,(1-\rho(s))$

■ **Roadmap.** The paper is organized as follows. In Section 2, I introduce the basic model, including the motivation for the cost function $c(s)$, a central feature in the paper. In Section 3, I calibrate the model based on specific function forms and on data from STEM doctoral programs composition. Section 4 introduces the formal concept of a minority trap: an asymmetric locally "stable" region of the state space. Section 5, I discuss two possible polices to avoid a minority trap. Section 6 concludes the paper.

2. Model

Consider a profession, that is, a set of employed workers with specific skills and job description, with ω jobs and an indefinite number of periods $t = 1, 2, \ldots$. Jobs are occupied by workers from two different populations, x and y. The profession evolves according to a birth and death (BD) Markov process:[3] Each period, one employed worker retires and is replaced by a new worker. We model workers as perpetual youth individuals, so that all workers are equally likely to retire. Whenever a position opens, Nature generates one worker of each type as candidates to fill the position. Worker ability a_i, $i \in \{x, y\}$, is independently drawn from the c.d.f. $F(a)$, with $F(0) = 0$ and $f(a) > 0$ for $a > 0$. Each candidate has an outside option worth zero.

We assume that the candidate with the highest ability is selected to occupy the open position. In other words, while discrimination in candidate selection may well be a motive for underrepresentation (especially in the past), in the present paper we focus on different factors, and thus assume (in the base model) that candidate selection is purely meritocratic.

The state of the system is given by the number n_i (or, equivalently, the share $s_i = n_i/\omega$) of each type of workers in the profession, where $i \in \{x, y\}$. With no loss of generality, I assume n measures the number of x types in the profession and s measures the share of x types, so $s = n/\omega$. For simplicity, I will use both s and n as the indicator of the state. Following Szpankowski (1989), I assume ω is sufficiently high that s is approximately continuous. With no loss of generality, I assume $s < \frac{1}{2}$ (i.e., type x is a minority).

I next describe the state transition probabilities. There are basically four possibilities, all listed in Table 1: the retiree can be an x or a y type and, similarly, the selected candidate can be of x or y type. If the retiree is y and the entrant is x, then the state moves from n to $n+1$. Conversely, if the retiree is x and the entrant is y, then the state moves from n to $n-1$. Otherwise, the state remains at n.

■ **Minority cost.** The crucial feature of the model is the assumption that workers who are in a minority in a field must pay a cost $c(s)$. Formally, with an appropriate use of units, I assume that a worker's utility is given by their ability, a_i, minus the cost due to being in a minority, $c(s_i)$, where s_i is the share of type i workers in the profession, $i \in \{x, y\}$. I assume that $c(s) \geq 0$ and that $c(s) = 0$ for $s \geq \frac{1}{2}$.

3. There is an extensive literature on birth-and-death processes, with many applications including in particular queuing theory. See, for example, Section 7.8 in Trivedi (2016). For a related economics application, see **?**.

The cost $c(s)$ can be material or psychological; and it can operate through a variety of mechanisms. One first mechanism is role modeling. For example, Kofoed and McGovney (2017) find that when a West Point female cadet is assigned a female mentor, the cadet is more likely to pick her officer's branch than if she would have interacted with a male mentor. In another example, by means of a longitudinal field experiment, Dennehy and Dasgupta (2017) show that peer mentoring increases female college students' retention rate in engineering fields.

A second channel is given by peer effects. Exploiting the exogenous assignment of cadets to companies at West Point, Huntington-Klein and Rose (2018) find that women cadets are more likely to advance when they have more women peers. In a field experiment, Shan (2022) shows that women assigned to female-minority groups are more likely to drop out of an economics course than women in other groups.

A third, related, channel, is given by (implicit or explicit) discrimination of female workers by male workers (or the reverse). Blair-Loy et al. (2017) examine over one hundred job talks across five engineering departments at two research universities. They find that, compared to men, a higher proportion of women's talk time is taken up by the audience asking questions. Dupas et al. (2021) also find that women presenters are asked more questions during a seminar. Moreover, they report that the questions asked of women presenters are more likely to be patronizing or hostile.

I do not have the ability, nor is that the goal of the paper, to understand which channel is more relevant. They all have one feature in common, namely that the job candidate's utility — and thus their propensity to accept a job offer — is lower when they are in a minority. And this I capture by assuming the cost function $c(s)$.

■ State transition probabilities. Suppose that $s < \frac{1}{2}$. Given our assumption that $F(0) = 0$, if $a_y > a_x$, then the open position is occupied by a type x worker. In fact, the worker has higher ability than the other candidate and, moreover, receives a utility a_y that is greater than the outside option. Suppose however that $a_x > a_y$. Then the x candidate is selected to occupy the open position. However, the utility received by this x candidate is given by $a_x - c(s)$. Since the outside option is equal to zero, the x candidate only accepts the offer if $a_x - c(s) > 0$. It follows that the probability $\rho(s)$ that the open position is occupied by a type x worker is given by

$$\rho(s) \equiv \mathbb{P}\big(a_x > a_y \ \wedge \ a_x > c(s)\big) \tag{1}$$

The first condition, $a_x > a_y$, states that x is better qualified than y, whereas the second condition, $a_x > c(s)$, states that x's net utility is greater than their outside option. If these conditions are not satisfied, then the position is occupied by a y candidate. In other words, candidate y occupies the open position with probability $1 - \rho(s)$. Note that the assumption that $c(s) = 0$ for $s > \frac{1}{2}$ implies that candidate y's outside option is dominated by the job offer.

Let $\boldsymbol{M}$ be the Markov transition matrix in the state space $\Omega = \{0, 1, ..., \omega\}$ denoting the number of x types in the workforce. We then have

$$\begin{aligned} M(n, n+1) &= (1-s)\,\rho(s) \\ M(n, n-1) &= s\,\big(1 - \rho(s)\big) \\ M(n, n) &= 1 - M(n, n-1) - M(n, n+1) \end{aligned}$$

where the first equation corresponds to the first row of Table 1, the second equation to the second row, and the third equation to the remaining two rows. Finally, for $\mid i - j \mid > 2$, $M(i, j) = 0$.

■ Stationary distribution. Birth-and-death (BD) Markov processes, defined by the property that $M(i, j) = 0$ if $\mid i - j \mid > 2$ (that is, $\boldsymbol{M}$ is a tri-diagonal matrix), have a series of regularities. Since $M(n-1, n), M(n, n)$, and $M(n, n+1)$ are strictly positive for all n, there exists a unique stationary distribution, given by the vector $\boldsymbol{\pi}$. The stationary distribution satisfies $\boldsymbol{M}\,\boldsymbol{\pi} = \boldsymbol{\pi}$ and is such that $\pi(n) > 0$ for all $n \in \Omega$. This implies that the present Markov process is ergodic: regardless of the initial state, we eventually converge to the same limiting distribution, $\boldsymbol{\pi}$. Since the state set Ω

is irreducible, we have a recursive process, which implies that detailed balance holds (Kelly, 1979, Theorem 1.3), namely

$$M(n-1,n)\,\pi(n-1) = M(n,n-1)\,\pi(n) \tag{2}$$

Equation (2) allows us to compute the stationary distribution recursively. Given $\pi(0)$, we have

$$\pi(n) = \pi(0)\,\prod_{i=1}^{n}\frac{M(i-1,i)}{M(i,i-1)}$$

Since $\sum_{i=0}^{\omega}\pi(i) = 1$, we finally have

$$\pi(0) = \left(1 + \sum_{n=1}^{\omega}\prod_{i=1}^{n}\frac{M(i-1,i)}{M(i,i-1)}\right)^{-1}$$

3. Calibration

I next illustrate and calibrate the model based on specific distributions, functional forms, and parameter values. Suppose that $F(a)$ is an exponential distribution, where I normalize the mean to be equal to 1. Suppose also that $c(s)$ is given by

$$c(s) = \frac{\gamma}{1+\exp\big(\beta\,(s-\alpha)\big)} \tag{3}$$

where α, β and γ are parameters; and, consistent with the idea that this is a minority cost, the domain is $s \in [0, .5)$, or, equivalently, for $s \geq .5$ we set $c(s) = 0$. The function $c(s)$ has three parameters: α, the location of the inverted S shape in terms of share of the profession; β, which measures the steepness of the decline in utility loss; and γ, which is approximately equal to the utility loss when worker type share is zero. I assume that $0 < \alpha < \frac{1}{2}$, $\beta > 0$, and $\gamma > 0$. Given the assumption that a_i is exponentially distributed, it can be shown (see Appendix) that

$$\rho(s) \equiv \mathbb{P}\big(a_x > a_y \;\wedge\; a_x > c(s)\big) = \exp\big(-c(s)\big) - \tfrac{1}{2}\,\exp\big(-2\,c(s)\big) \tag{4}$$

The Markov matrix thus becomes

$$\begin{aligned}
M(n,n+1) &= s\,\Big(\exp\big(-c(1-s)\big) - \tfrac{1}{2}\,\exp\big(-2\,c(1-s)\big)\Big)\\
M(n,n-1) &= (1-s)\,\Big(\exp\big(-c(s)\big) - \tfrac{1}{2}\,\exp\big(-2\,c(s)\big)\Big)\\
M(n,n) &= 1 - M(n,n-1) - M(n,n+1)
\end{aligned}$$

The above set of functional forms require that I calibrate three parameters: α, β, and γ. Bostwick and Weinberg (2022) use data of graduate student achievement in Ohio from June 2009–June 2015 to estimate peer effects of male and female doctoral students in various fields. I am particularly interested in the fields, listed in Table 2, where women are a small minority. The basic regression ran by Bostwick and Weinberg (2022) takes the form

$$\mathbb{P}\big(Y_{ipt}=1\big) = \beta_1 \times \text{Female}_i + \beta_2 \times \text{FemalePeers}_{ipt} + \beta_3 \times \text{Female}_i \times \text{FemalePeers}_{ipt} + Z_{ipt}$$

where $\mathbb{P}\big(Y_{ipt}=1\big)$ is the probability that student i, who first enrolled in program p in year t, completes a PhD within 6 years; Female_i is equal to 1 if and only if student i is female; FemalePeers_{ipt} is the fraction of female students in student i's program; and Z measures other controls as well as the regression error term. Suppose that the event "completing a PhD within 6 years" is a good indicator of entering the profession. Then the coefficient β_3 provides an estimate of the differential peer effect that the fraction of female students in the program has on female student (relative) success (and on their entry into the profession). Making use of the variation in the fraction of female students

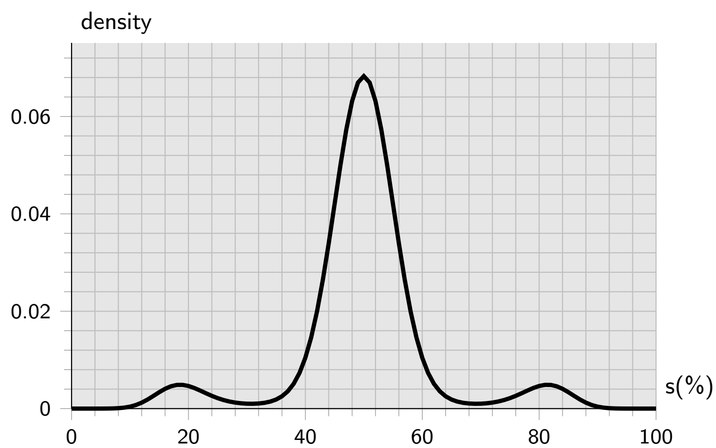
Table 2

Doctoral program cohort size and female share. Sources: cohort size and average % female from Table 1 in Bostwick and Weinberg (2022); standard deviation: author's calculations based on data provided by Bostwick and Weinberg.

Subject title	Average cohort size	Average % female	St. Dev. % female
Computer engineering	24.0	15	10.6
Electrical and electronics engineering	19.9	19	9.8
Mathematics	17.7	19	12.8
Physics	18.5	24	9.2

Figure 1

Stationary distribution of s in calibrated model



(even within a given discipline), Bostwick and Weinberg (2022) estimate $\widehat{\beta}_3 = .210$, with a standard deviation of .0941.

Based on the data in Bostwick and Weinberg (2022), I identify the value of γ by fitting the sample average of 19% female participation (the average value in the fields of EE and math, as per Table 2). I then identify the value of α by forcing

$$\left.\frac{dp(s)}{ds}\right|_{s\,=\,s'} = \widehat{\beta}_3 = .210 \tag{5}$$

Finally, I select the value of β that reproduces the second moment of s (that is, the standard deviation of the share of female), also as per Table 2. This yields the values $\alpha = .3212$, $\beta = 26.4$, and $\gamma = 1.5968$.

4. Minority traps

Figure 1 shows the stationary density of the calibrated model presented in the previous section. Although the density is strictly positive everywhere, in practice it takes very low values for a number of states. As a result, there are three "regions" of states, each reasonably separated from the other ones, corresponding to three modes of the stationary distribution. In other words, in the very long run (meaning, in the limit as $t \to \infty$) it does not matter where the process starts. However, for a

Figure 2
Probability distribution over states after τ periods starting from state 0 for three values of τ. Note that the third distribution is very close to the stationary distribution in Figure 1 (though the scale is different)

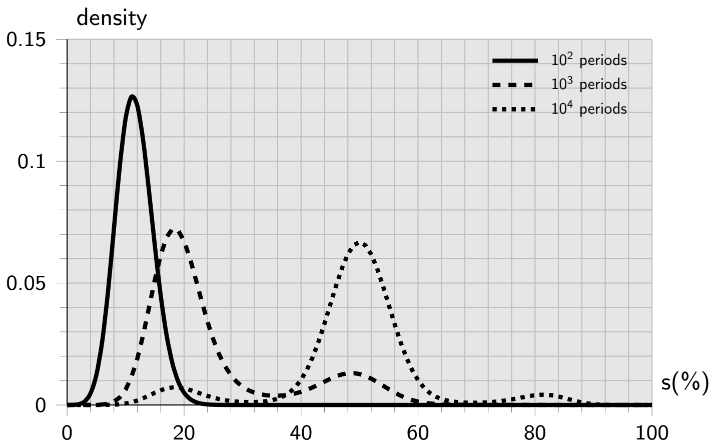
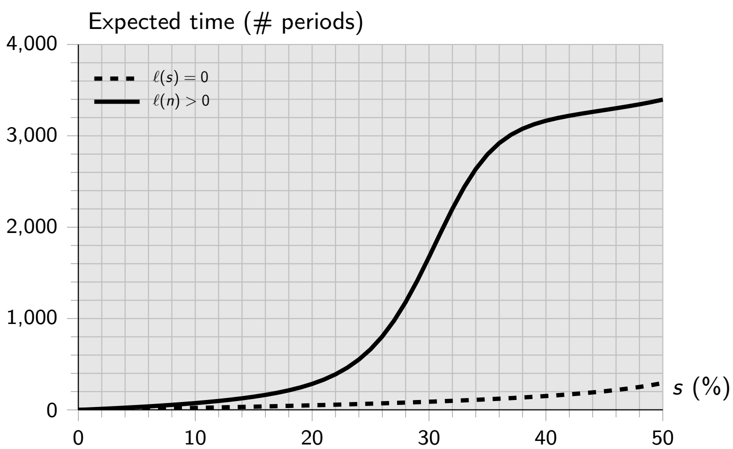


Figure 3
Expected time (number of periods) to reach market share s starting from $s = 0$.



finite t, it does make a difference what the initial state is. This is true for any ergodic system, but as I show next it is particularly true for multi-modal stationary distributions like to one in Figure 1

In order to understand how the dynamics of s unfold, and how they depend on initial conditions, we compute the distribution of s after τ periods. This is given by $\boldsymbol{\pi}_\tau \equiv \boldsymbol{M}^\tau\, \mathbf{1}_0$, where $\mathbf{1}_i$ is a vector with a 1 in the ith position and zeros elsewhere. Figure 2 plots the distribution for three values of τ, one hundred, one thousand, and ten thousand. After 10,000 periods, the distribution (in dotted line) is very close to the stationary distribution (as can be seen by comparing it to Figure 1). However, after one thousand periods, the distribution is still fairly concentrated around the left mode of the stationary distribution. Since I consider a profession of size 100, one thousand periods would allow for ten complete renewals of the composition of the workforce. In this sense, it is remarkable that the distribution is still so close to the left portion of the stationary distribution. It's as if there is a "minority trap" that keeps the system in an asymmetric state for a very long time.

Another way of understanding path dependance (or, to be more precise, quasi-path-dependance), is to derive the expected time to reach a certain state n' starting from a given initial state n. Let $T(n, n')$ be the estimated time (average number of periods) that it takes to go from state n to state

n' . For a BD process, the following relation applies:

$$T(n, n') = 1 + M(n, n-1)\, T(n-1, n') + M(n, n)\, T(n, n') + M(n, n+1)\, T(n+1, n')$$

to which we may add the terminal condition $T(n', n') = 0$. Let $\boldsymbol{M}_\tau$ be the transition matrix $\boldsymbol{M}$ restricted to to the first τ states (that is, from 0 to $\tau - 1$). Then we have

$$\boldsymbol{T}_\tau = \mathbf{1} + \boldsymbol{M}_\tau \boldsymbol{T}_\tau \tag{6}$$

where $\boldsymbol{T}_\tau$ is a τ-dimensional vector containing $T(0, \tau), ..., T(\tau-1, \tau)$ and $\mathbf{1}$ is a vector of 1's. Equation (6) implies

$$\boldsymbol{T}_\tau = (\boldsymbol{I} - \boldsymbol{M}_\tau)^{-1}\, \mathbf{1}$$

where $\boldsymbol{I}$ the identity matrix. Repeating this operation for all $\tau \in \Omega$ and collecting the first term of the resulting $\boldsymbol{T}_\tau$ vector (that is, $T(0, \tau)$), we construct the sequence $T(0, 1), T(0, 2), ..., T(0, \tau), ..., T(0, \omega)$ giving the expected time to reach state τ beginning from state zero. Figure 3 plots this sequence for $\tau = 0, ..., 50$ and for two different cases: the case when $c(s)$ is given by (3) and, for comparison purposes, the case when $c(s) = 0$, so that the resulting stationary distribution is unimodal. If there is no utility loss from being in a minority, then it takes on average 294 periods to reach state 50 starting from state 0. By contrast, given the calibrated loss function $c(s)$, it takes an average 3396 periods.

In sum, if the cost from being in a minority is sufficiently high — and the empirical evidence is consistent with that being the case — then, while the distribution of gender shares is symmetric, starting from an asymmetric state it may take a long time for parity to emerge (even when the overall mode of the distribution is 50%).

This phenomenon of inertia or hystereses has been documented in a variety of contexts, including electrical engineering (Carleial and Hellman, 1975) and biophysics (Fologea et al., 2011). Using a terminology closer to economics, we might refer to it as quasi path dependance, where the "quasi" stems from the fact that, in the limit as $t \to \infty$, the system's state is independent of initial conditions, whereas for finite t, even for a "large" finite t, the system is still dependent on initial conditions. This property is related to, but different from, quasi-stationarity (see van Doorn and Pollett, 2013, for a survey). The latter occurs when the Markov process has an almost surely absorbing state (e.g., population death) but nevertheless admits a "stationary steady state" that lasts for a long time. In some sense, the case we consider in this paper is the exact opposite: instead of a non-stationary system that appears to be stationary, we have a system that is stationary but appears not to be.

■ **Sufficient conditions for multi-modality.** Figures 1–3 illustrate the possibility of a minority trap: although parity between x and y takes place in the long run, the long run may be very long if we start from x close to zero. These minority traps result from the cost suffered by workers when in a minority and are associated with a multi-modal stationarity distribution. How general are these results? Under what conditions do we obtain a multi-modal stationary distribution as in Figure 1? The following result provides an answer to these questions:

Proposition 1. *There exists a $c' > 0$ such that, if $c(s) > c'$, then the stationary distribution of s is multi-modal.*

In other words, Proposition 1 establishes that, if the cost from being a minority is sufficiently high, then the stationary distribution of worker types is multi-modal. How high is high? The calibration with data from women in STEM fields seems consistent with the idea that the stationary distribution is muilti-modal. That said, there are a number of identification issues to which I will return in Section 6.

Figure 4
New hire productivity under different regimes

5. Affirmative action

A number of policies have been proposed to counteract the problem of under-representation in certain professions. Here I focus on two types of policy. First, what one might refer to as a quota policy: As long as the share of type x falls below a certain threshold, candidate x is selected even if their ability a_x is lower than the other candidate's, a_y. Note that this does not necessarily imply that the open position is filled by x, for it still has to be the case that the job candidate prefers the position to the outside option, that is, $a_x > c(s)$.

A second policy is to subsidize minority candidates. By this I mean that a minority candidate is given resources that exactly compensate for the utility loss $c(s)$ of being in a minority. And, other than that, candidates are selected based on ability a_i, $i \in \{x, y\}$.

Figure 4 shows the expected value of productivity of a new hire a_i ($i \in \{x, y\}$) over time according to each of the above regimes, to which I add a third one corresponding to no intervention.[4] As can be seen, no intervention implies that, on average, a new hire has a productivity of about 1, actually, slightly higher than 1. Recall that we assume average ability (and productivity) per worker is 1. The fact that the values in Figure 4 are slightly higher than 1 reflects the the feature that, with some (low) probability, a type x is a credible and willing contestant for the new job, thus increasing the expected productivity of the selected candidate.

Consider now the quota policy. Under this system, a minority candidate is selected regardless of their ability. This lowers expected productivity relative to the "meritocratic" system because, with some probability, type y workers are ignored, even if their ability is higher. However, it is still the case that, for a large fraction of x types, utility is lower that the outside option, which implies that the type y is selected. Not surprisingly, the quota policy leads to lower productivity. What is perhaps surprising is that it does *not* lead to an "exit" from the minority trap. Intuitively, what happens is that any increase in the share of x is compensated by an increases in their retirement rate and this effect dominates the decrease in cost $c(s)$ that minority candidates benefit from.

Finally, consider the subsidy case. Since x types are compensated exactly for their utility loss $c(s)$, it follows that x and y types are chosen with equal probability and the stationary distribution is effectively the uni-modal distribution corresponding to that case. In terms of net productivity (net of the subsidy given to the worker), there is a big drop during the early states of the process (when the subsidies are actually paid). However as the state of the system converges "rapidly" to the middle states (cf Figure 3), productivity jumps to a much higher value (approximately 1.5). This increase results from two factors. First, as the share of x in the profession increases, the utility loss

4. I compute expected values by averaging 10^5 300-period histories.

$c(s)$ vanishes and so do the subsidies. Second, the expected productivity of the new hire is now the higher of two independent exponentials with average 1. In fact, it can be shown that the highest of two independent exponential variables with mean 1 is equal to 1.5, which corresponds to the values plotted in Figure 4.[5]

The effect illustrated by Figure 4 is reminiscent of the effect described in Hsieh et al. (2019). They estimate that 20% to 40% of growth in aggregate market output per person in the US can be explained by the improved allocation of talent, specifically the increase in minority integration into professions such as doctors and lawyers since the 1960s.[6]

6. Conclusion

A report on the status of women in STEM fields states that

> Making full use of the nation's scientific and technical talent, regardless of the sex, social, and ethnic characteristics of the persons who possess it, will require both understanding of the causes of inequality and effective remedies (National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, 2007).

In this paper, I suggest that one possible explanation for high gender asymmetries is the "chicken-and-egg" problem created by minority costs. Formally, this amounts to a multi-modal stationary distribution of a birth-and-death Markov process. Even though all states are communicated, the process remains in a set of asymmetric states for very long. In this context, the right affirmative action policy may successfully nudge the system towards the central model of the stationary distribution.

To be clear, my analysis is based on the *assumption* of symmetric preferences and abilities. One alternative explanation for asymmetry is simply that women have different preferences (even when systemic barriers to equal access are removed). Moreover, there are numerous issues with the identification strategy I followed, both in terms of functional forms and and in terms of parameter values. That said, I have shown that a multi-modal stationary distribution is not a purely theoretical possibility.

Multi-modality implies a sort of path dependance. In the limite as $t \to \infty$, the process converges to a symmetric stationary distribution with most mass around parity. However, absent any intervention, it may take a long time for that to happen. Path dependence also begs the question of why we start from an asymmetric state to begin with. In recent empirical work, Andrews and Zhao (2022) note that women's minority stratus is not uniform across all STEM fields. Based on a difference-in-differences framework, they argue that exposure to home economics college studies, which included considerable chemistry and biology content, had a lasting effect on women's interest in these fields. This evidence of path dependance corroborates the theory of multi-modality, but obviously much work remains to be done to assess its validity.

5. See Athey, Avery, and Zemsky (2000) Johnson, Kotz, and Balakrishnan, 1994, Equation (19.11).
6. "In 1960, 94 percent of doctors and lawyers were white men. By 2010, the fraction was just 62 percent. Similar changes in other highly-skilled occupations have occurred throughout the U.S. economy during the last 50 years." See also Goldin (2006).

Appendix

Lemma 1. *Suppose that a_x and a_y are independently distributed as exponential with parameter 1. Then*

$$\mathbb{P}\big(a_x > a_y \;\wedge\; a_x > c(s)\big) = \exp\big(-c(s)\big) - \tfrac{1}{2}\,\exp\big(-2\,c(s)\big)$$

Proof of Lemma 1: We prove a slightly more general result. Suppose that x and y are exponentially distributed with parameters λ_x and λ_y; and that z is a scalar. Then

$$\mathbb{P}(x > y \wedge x > z) = \exp(-\lambda_x\, z) - \frac{\lambda_x}{\lambda_x + \lambda_y}\;\exp\big(-(\lambda_x + \lambda_y)\, z\big)$$

$$\begin{aligned}
\mathbb{P}(x > y \wedge x > z) &= \int_z^\infty \lambda_x\,\exp(-\lambda_x\, x) \int_0^x \lambda_y\,\exp(-\lambda_y\, y)\; dy\; dx \\
&= \int_z^\infty \lambda_x\,\exp(-\lambda_x\, x)\,\Big[-\,\exp(-\lambda_y\, y)\Big]_0^x\; dx \\
&= \int_z^\infty \lambda_x\,\exp(-\lambda_x\, x)\,\big(1 - \exp(-\lambda_y\, x)\big)\; dx \\
&= \int_z^\infty \lambda_x\,\Big(\exp(-\lambda_x\, x) - \exp\big(-(\lambda_x + \lambda_y)\, x\big)\Big)\; dx \\
&= \Big[-\exp(-\lambda_x\, x)\Big]_z^\infty + \frac{\lambda_x}{\lambda_x + \lambda_y}\;\Big[\exp\big(-(\lambda_x + \lambda_y)\, x\big)\Big]_z^\infty \\
&= \exp(-\lambda_x\, z) - \frac{\lambda_x}{(\lambda_x + \lambda_y)}\;\exp\big(-\lambda_x + \lambda_y\, z\big)
\end{aligned}$$

Setting $\lambda_x = \lambda_y = 1$ and substituting $c(s)$ for z the result follows. ■

Proof of Proposition 1: Szpankowski (1989) shows that the modes of the stationary distribution of s correspond to the zeros of the average drift function $d(s)$, defined as

$$d(s) = M(n, n+1) - M(n, n-1)$$

where $n = s\,\omega$. From Table 1, we see that

$$\begin{aligned}
d(s) &= (1-s)\,\rho(s) - s\,(1 - \rho(s)) \\
&= \rho(s) - s
\end{aligned} \tag{7}$$

where $\rho(s)$ is given by

$$\rho(s) \equiv \mathbb{P}\big(a_x > a_y \;\wedge\; a_x > c(s)\big)$$

From (7), we see that $d(0) = \rho(s) > 0$. Moreover, for a strictly positive s', $\rho(s) \to 0$ as $c(s') \to \infty$, which in turn implies that $d(s') \to -s < 0$. It follows that, if $c(s')$ is sufficiently high, then $d(s)$ goes from being positive at zero to being negative at s'. Therefore, there exists value of $s \in (0, s')$ such that $d(s) \approx 0$, which in turn implies that the stationary distribution has a mode at s. ■

References

Andrews, Mike and Yiling Zhao (2022), “Home Economics and Women’s Gateway to Science,” Working paper, University of Maryland and Peking University.

Athey, Susan, Christopher Avery, and Peter Zemsky (2000), “Mentoring and Diversity,” *American Economic Review*, 90, 765–786.

Blair-Loy, Mary, Laura E. Rogers, Daniela Glaser, Y. L. Anne Wong, Danielle Abraham, and Pamela C. Cosman (2017), “Gender in Engineering Departments: Are There Gender Differences in Interruptions of Academic Job Talks?” *Social Sciences*, 6, 1–19.

Bostwick, Valerie K. and Bruce A. Weinberg (2022), “Nevertheless She Persisted? Gender Peer Effects in Doctoral STEM Programs,” *Journal of Labor Economics*, 40, 397–436.

Cabral, Luís (2011), “Dynamic Price Competition with Network Effects,” *Review of Economic Studies*, 78, 83–111.

Carleial, Aydano and Martin Hellman (1975), “Bistable Behavior of ALOHA-type Systems,” *IEEE Transactions on Communications*, 23, 401–410.

Chung, Kim-Sau (2000), “Role Models and Arguments for Affirmative Action,” *American Economic Review*, 90, 640–648.

Dennehy, Tara C. and Nilanjana Dasgupta (2017), “Female Peer Mentors Early in College Increase Women’s Positive Academic Experiences and Retention in Engineering,” *Proceedings of the National Academy of Sciences*, 114, 5964–5969.

Dupas, Pascaline, Alicia Sasser Modestino, Muriel Niederle, Justin Wolfers, and The Seminar Dynamics Collective (2021), “Gender and the Dynamics of Economics Seminars,” Working Paper 28494, National Bureau of Economic Research.

Fershtman, Daniel and Alessandro Pavan (2021), “"Soft" Affirmative Action and Minority Recruitment,” *American Economic Review: Insights*, 3, 1–18.

Fologea, Daniel, Eric Krueger, Yuriy I. Mazur, Christine Stith, Yui Okuyama, Ralph Henry, and Greg J. Salamo (2011), “Bi-Stability, Hysteresis, and Memory of Voltage-Gated Lysenin Channels,” *Biochimica et Biophysica Acta (BBA) — Biomembranes*, 1808, 2933–2939.

Goldin, Claudia (2006), “The Quiet Revolution That Transformed Women’s Employment, Education, and Family,” *American Economic Review*, 96, 1–21.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow (2019), “The Allocation of Talent and U.S. Economic Growth,” *Econometrica : journal of the Econometric Society*, 87, 1439–1474.

Huntington-Klein, Nick and Elaina Rose (2018), “Gender Peer Effects in a Predominantly Male Environment: Evidence from West Point,” *AEA Papers and Proceedings*, 108, 392–395.

Johnson, Norman L., Samuel Kotz, and Narayanaswamy Balakrishnan (1994), *Continuous Univariate Distributions, Vol 1, 2nd Ed*, New York: John Wiley & Sons.

Kelly, Frank (1979), *Reversibility and Stochastic Networks*, Cambridge University Press.

Kofoed, Michael S. and Elizabeth McGovney (2017), “The Effect of Same-Gender and Same-Race Role Models on Occupation Choice: Evidence from Randomly Assigned Mentors at West Point,” *Journal of Human Resources*, 0416.

Müller-Itten, Michèle and Aniko Öry (2022), "Mentoring and the Dynamics of Affirmative Action," *American Economic Journal: Economic Policy*, 14, 402–444.

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine (2007), *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering*, The National Academies Press, Washington, DC.

Shan, Xiaoyue (2022), "The Minority Trap: Minority Status Drives Women Out of Male-Dominated Fields," Working Paper, University of Pennsylvania.

Szpankowski, Wojciech (1989), "Some Remarks on Uniformly Bounded Markov Chains: Multimodality Analysis," *Computers & Operations Research*, 16, 85–99.

Trivedi, Kishor Shridharbhai (2016), *Probability and Statistics with Reliability, Queueing, and Computer Science Applications, 2nd Edition*, Wiley.

van Doorn, Erik A. and Philip K. Pollett (2013), "Quasi-Stationary Distributions for Discrete-State Models," *European Journal of Operational Research*, 230, 1–14.