Self-Image Bias and Talent Loss
On-Line Appendix

Marciano Siniscalchi
Northwestern University
Pietro Veronesi
University of Chicago

A1. Extensions

A1.1. Endogenous Choice of Young Researchers

Consider a potential researcher choosing between an academic career and an outside option. The prospective researcher knows her type θ, and is aware of both the likelihood of producing quality research, and the evaluation criteria used by the referees. Attempting to pursue research entails a cost C, which is identical across agents. If the potential researcher is hired (accepted), he or she receives a payoff of P; finally, the outside option is normalized to 0. Thus, the total payoff is $P - C$ if the researcher is hired, and $-C$ otherwise. What types of agents decide to pay the cost C and thus take their chance with the academic career?

Assume that the entry decision, research activity, and hiring decision all occur at time t. Then, given the time-t distribution $\lambda_t = (\lambda_t^\theta)_{\theta \in \Theta}$ of referees' types, a prospective researcher of type θ pursues an academic career—"applies"—if and only if

$$\gamma^\theta \lambda_t^\theta (P - C) + (1 - \gamma^\theta \lambda_t^\theta)(-C) > 0. \tag{A.28}$$

Consequently, the accepted mass of researchers is as follows: for $g = f, m$,

$$a_t^{\theta,g} = \begin{cases} \gamma^\theta \cdot \lambda_{t-1}^\theta \cdot p^{\theta,g} & \text{if } \gamma^\theta \lambda_{t-1}^\theta \geq \frac{C}{P} \\ 0 & \text{otherwise} \end{cases} \tag{A.29}$$

$$\lambda_t^{\theta,g} = \lambda_{t-1}^{\theta,g} (1 - a_t) + a_t^{\theta,g} \tag{A.30}$$

Expression (A.29) shows that if the mass of type-θ reviewers drops below $\frac{C}{\gamma^\theta P}$ at time $t-1$, both M and F young type-θ researchers will not apply at date t. From Eq. (A.30), this implies that the total mass of such types will decrease, at least weakly, because some type-θ established researchers will have to retire in order to make room for researchers of other types who are accepted. In fact, the mass of such types will decrease strictly, except in case no young researcher wants to apply.

While the dynamics with endogenous entry is considerably more complicated than in the benchmark case, we prove the following Proposition:

Proposition A.1 *Let*

$$\Theta^{\max} = \left\{ \theta \in \Theta \text{ such that } \lambda_0^\theta \geq \frac{C}{\gamma^\theta P} \text{ and } \theta \in \arg\max_{\theta' \in \Theta} \gamma^{\theta'} (p^{\theta',m} + p^{\theta',f}) \right\} \tag{A.31}$$

then:

(i) only $\theta \in \Theta^{\max}$ survive in the limit as $t \to \infty$;

(ii) $\Theta^{\max}$ need not preserve symmetry across types: if $\theta \in \Theta^{\max}$, then $\sigma(\theta) \in \Theta^{\max}$ if and only if $\lambda_0^{\sigma(\theta)} \geq \frac{C}{\gamma^\theta P}$.

(iii) for every $\theta \in \Theta^{\max}$,

$$\bar{\lambda}^\theta = \frac{\lambda_0^\theta}{\sum_{\theta' \in \Theta^{\max}} \lambda_0^{\theta'}}. \tag{A.32}$$

(iv) for every $\theta \in \Theta^{\max}$,

$$\bar{\lambda}^{\theta,m} = \frac{\lambda_0^\theta \gamma^\theta p^{\theta,m}}{\sum_{\theta' \in \Theta^{\max}} \lambda_0^{\theta'} \gamma^\theta (p^{\theta,m} + p^{\theta,f})} \quad \text{and} \quad \bar{\lambda}^{\theta,f} = \frac{\lambda_0^\theta \gamma^\theta p^{\theta,f}}{\sum_{\theta' \in \Theta^{\max}} \lambda_0^{\theta'} \gamma^\theta (p^{\theta,m} + p^{\theta,f})}. \tag{A.33}$$

While the general structure of the limit distribution of types is similar to that in Propositions 2 and 3, the key difference is part (ii), which has implications for Eqs. (A.32) and (A.33). Two types θ and $\theta' = \sigma(\theta)$ may be symmetric, and yet differences in their initial frequencies $\lambda_0^\theta, \lambda_0^{\theta'}$ may imply that one survives and the other doesn't. This may exacerbate group imbalance.

Take for instance Corollary 1, a key result in our model. Consider $\theta, \theta' = \sigma(\theta)$ with $p^{\theta,m} > p^{\theta',m}$, and assume that both θ and θ' are in $\Theta^{\max}$. Then the conclusion of that Corollary holds verbatim, because its proof only relies on Eq. (7), which is structurally identical to Eq. (A.33). (The denominators will be different because of differences in the sets $\Theta^{\max}$, but this is immaterial to the argument.) However, if in addition type θ' does not survive because $\lambda_0^{\theta'} < \frac{C}{\gamma^\theta P}$, there is an additional force contributing to group imbalance, as shown next.

Corollary A.1 *Assume $\lambda_0 = p^m$. Let $\theta \in \Theta^{\max}$ be such that $p^{\theta,m} > p^{\theta,f}$. Then, there is a relative cost C/P such that $\theta' = \sigma(\theta) \notin \Theta^{\max}$ and $\bar{\lambda}^{\theta,m} > \bar{\lambda}^{\theta,f}$. Moreover, in the special case in which θ is the only surviving type, M-imbalance occurs:*

$$\bar{\lambda}^{\theta,m} = \frac{p^{\theta,m}}{p^{\theta,m} + p^{\theta,f}} > 0.5 > \frac{p^{\theta,f}}{p^{\theta,m} + p^{\theta,f}} = \bar{\lambda}^{\theta,f}$$

Finally, in this case, M-imbalance is larger than in the case with zero cost $C = 0$. That is:

$$\bar{\lambda}^{\theta,m} - \bar{\lambda}^{\theta,f} > \left(\bar{\lambda}^{\theta,m}_{C=0} + \bar{\lambda}^{\theta',m}_{C=0} \right) - \left(\bar{\lambda}^{\theta,f}_{C=0} + \bar{\lambda}^{\theta',f}_{C=0} \right)$$

where the subscript $C = 0$ denotes the case with zero cost, in which case $\theta' = \sigma(\theta) \in \Theta^{\max}$ [see Proposition 2 part (ii)].

Proposition A.1 and Corollary A.1 formalize statements (i)–(iii) in Section 6.. Statement (i) corresponds to part (ii) of Proposition A.1. Statement (ii) follows from Corollary A.1 (greater imbalance than for $C = 0$) and the definition of $\Theta^{\max}$ in Proposition A.1 (types more common in the F group may choose not to apply). Finally, Statement (iii) follows by noting that, in Corollary 3, if $\theta, \theta' = \sigma(\theta)$ are distinct elements of $\Theta^{\max}$ when $C = 0$, with $p^{\theta,m} > p^{\theta,f}$, then "field" θ is M-dominated and, symmetrically, "field" θ' is F-dominated; this remains true for $C > 0$ if $\theta, \theta' \in \Theta^{\max}$, but if $\theta' \notin \Theta^{\max}$, then "field" θ is M-dominated by Corollary A.1 and there is no corresponding F-dominated field.

For the special case of the N-characteristics model in Section 4., we have:

Corollary A.2 *Assume that at time 0, all referees are from the M-group with $\lambda_0 = p^m$.*

(a.1) If $\rho < \bar{\rho}(\phi, N)$ and $\frac{C}{P} \leq (1-\phi)^N \gamma_0 \sqrt{\rho}$, then the steady state is as in Corollary 6 (a).

(a.2) If $\rho < \bar{\rho}(\phi, N)$ and $(1-\phi)^N \gamma_0 \sqrt{\rho} < \frac{C}{P} \leq \phi^N \gamma_0 \sqrt{\rho}$, then only type θ^m survives in the limit, i.e. $\bar{\lambda}^{\theta^m} = 1$. The limiting mass of M researchers is strictly larger than in (a.1):

$$\bar{\Lambda}^m = \lim_{t \to \infty} \sum_{\theta} \lambda_t^{m,\theta} = \frac{\phi^N}{\phi^N + (1-\phi)^N} > \frac{1 + \left(\frac{\phi}{1-\phi}\right)^{2N}}{1 + \left(\frac{\phi}{1-\phi}\right)^{2N} + 2\left(\frac{\phi}{1-\phi}\right)^{N}}. \tag{A.34}$$

(b) If $\rho > \bar{\rho}(\phi, N)$ and $[\phi(1-\phi)]^{N/2} \geq \frac{C}{\gamma_0 \rho P}$, then the steady state is as in Corollary 6 (b).

In each of the above cases, if $\bar{\lambda}^\theta = 0$, then there is $t^\theta \geq 0$ such that $\lambda_t^\theta = 0$ for all $t \geq t^\theta$.

Part (a.1) and (b) of this proposition shows that if the cost C is low enough, then the the steady state is the same as in the basic model in Section 4. for the same two conditions about ρ, respectively. This is intuitive. The only difference is that all types other than surviving ones drop out in finite time, rather than only in the limit.

The interesting new part is (a.2). In this case, the only type that survives in the long-run is θ^m, the most prevalent type in the M-population. In particular, θ^f now disappears. Thus, the characteristics that are mildly more frequent in the F-population, but also common in the M-population, eventually disappear. In this case, endogenous entry greatly exacerbates the loss of talent compared to the base case. Indeed, the total mass of M researchers, $\bar{\Lambda}^m$, is now even larger than in its counterpart without endogenous entry, whose expression is in Eq. (16) in Corollary 6. Thus, if the conditions in part (a.2) are satisfied, the distribution of established researchers will be even more skewed towards the M group.

Parts (a.1)–(b) do not exhaust all possible cases; for instance, they do not analyze the possibility that the first condition in part (b) holds, but the second does not—that is, θ^* is not willing to apply. The following section illustrates an instance of one such possibility.

The proof of the above Proposition in the Appendix provides a general characterization that can be used to further explore different parametric choices.

A1.1.1. Example of Group Imbalance due to Endogenous Entry

We first illustrate how endogenous entry can exacerbate group imbalance, provided the cost of entry is not too small. Consider the parameterization in Section 4. In our basic model, M-researchers represent 91% of the overall population in the limit. If we add endogenous entry, Corollary A.2 shows that the steady state either remains the same, if the cost C is sufficiently low, as in case (a.1), or it becomes even more skewed towards the M group, as in case (a.2). In the latter case, the limiting fraction of M-researchers is $\bar{\Lambda}^m = \phi^N/(\phi^N + (1-\phi)^N) = 95\%$.

We now illustrate how endogenous choice may prevent convergence to group balance even when group balance would in fact attain in the basic model. We use the same parameterization as in Section 4., except that the number of characteristics is $N = 8$ instead of $N = 10$. With these parameter values, Corollary 6 part (b) implies that the system will converge to an equal mass of M and F researchers, because $\rho = 5 > 3.61 = \bar{\rho}(\phi, N)$. The solid and dashed lines in Figure A.1 confirm this.

However, assume now that entry is endogenous; the payoff if a researcher is hired is $P = 1,000$, and the cost of entry is $C = 4$ (i.e., 0.4% of the payoff of becoming a researcher over the outside option). Note that these parameters apply equally to M and F researchers. The key point is that now the efficient type θ^* (M or F) does not want to apply at date 0:

$$\lambda_0^{\theta^*} = p^{\theta^*,m} = \phi^{N/2}(1-\phi)^{N/2} = 0.3574\% < 0.4\% = \frac{C}{\gamma^{\theta^*} P}.$$

Moreover, type θ^f (M or F) does not want to apply either:

$$\lambda_0^{\theta^f} = p^{\theta^f,m} = (1-\phi)^N = 0.1081\% < 0.8944\% = \frac{C}{\gamma^{\theta^f} P}.$$

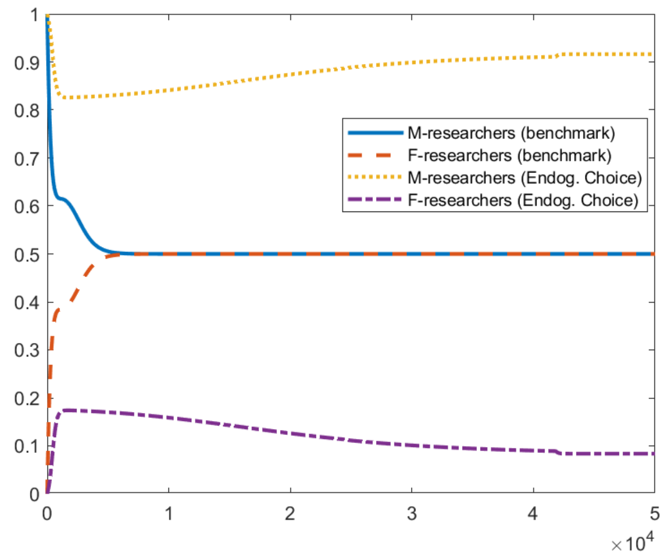
On the other hand, type θ^m (M or F) does:

$$\lambda_0^{\theta^m} = p^{\theta^m,m} = \phi^N = 1.18\% > 0.8944\% = \frac{C}{\gamma^{\theta^m} P}.$$

Therefore, while other types are also willing to apply, type θ^m will prevail, which will lead to a severe imbalance between M and F researchers in the limit, as shown in Figure A.1. Indeed, in this case the talent loss is rather severe, as the only surviving type $\theta^m = (1, ..., 1, 0, ...0)$ has none of the research characteristics that are (mildly) more common in the F-population. Figure A.2 shows that both F and M researchers are of type θ^m in the long run.

To sum up, even if the basic environment is meritocratic, in the sense that differences in talents γ^θ across types are sufficient to lead to group balance, endogenous entry introduces a bias in favor of M-researchers which leads to an imbalance steady state. In this case, policies aimed at lowering the cost C can lead to group balance in the long run.

Figure A.1: Fraction of M and F Researchers with Endogenous Entry

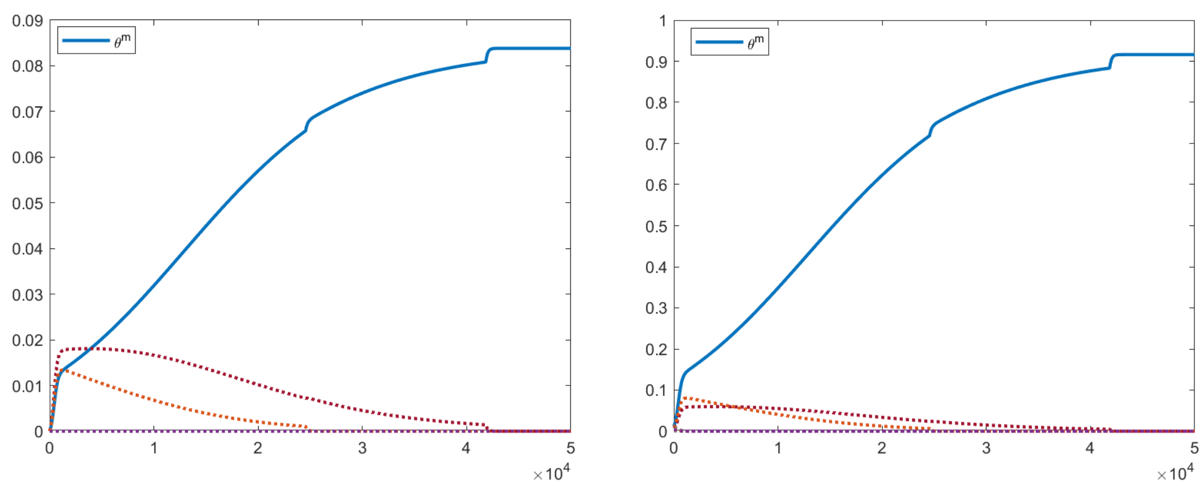


Fraction of M and F researchers when $\lambda_0 = p^m$. Parameters: $\phi = 0.5742$ ($d = 0.3$), $\gamma_0 = 0.2$, $\rho = 5$, $N = 8$, $P = 1000$, and $C = 4$.

Figure A.2: Types of Established F and M Researchers with Endogenous Entry

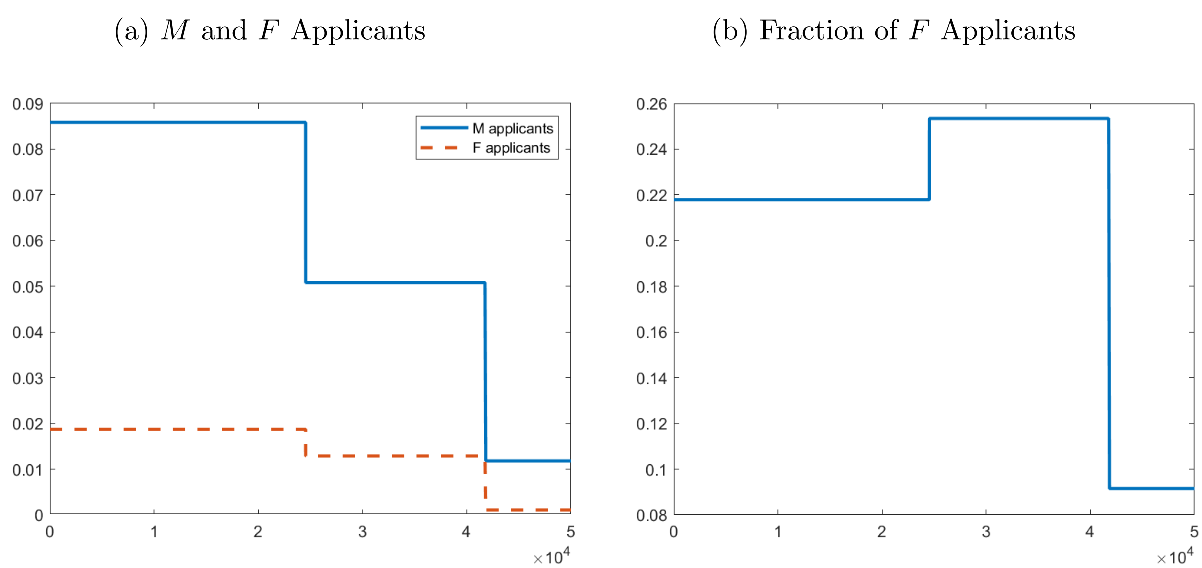
(a) F researchers

(b) M researchers



Types of established F (left) and M (right) researchers with endogenous entry. $\theta^m = (1, ..., 1, 0, ..., 0)$ dominates; all other types eventually vanish. Parameters: $\phi = 0.0.5742$ ($d = 0.3$), $\gamma_0 = 0.2$, $\rho = 5$, $N = 8$, $P = 1000$, and $C = 4$.

Figure A.3: Endogenous entry: applicants



Total mass of M and F applicants (left) and fraction of F applicants (right). Parameters: $\phi = 0.5742$ ($d = 0.3$), $\gamma_0 = 0.2$, $\rho = 5$, $N = 8$, $P = 1000$, and $C = 4$.

A1.1.2. Characterization of the Applicant Pool

Due to variation in the distribution of characteristics, Corollary A.2 also has implications for the mass of young M and F researchers who decide to apply for an academic job:

Proposition A.2 *For every t, let*

$$A_t^m = \sum_{\theta:\lambda_t^\theta \geq \frac{C}{\gamma^\theta P}} p^{\theta,m} \quad \text{and} \quad A_t^f = \sum_{\theta:\lambda_t^\theta \geq \frac{C}{\gamma^\theta P}} p^{\theta.,f}$$

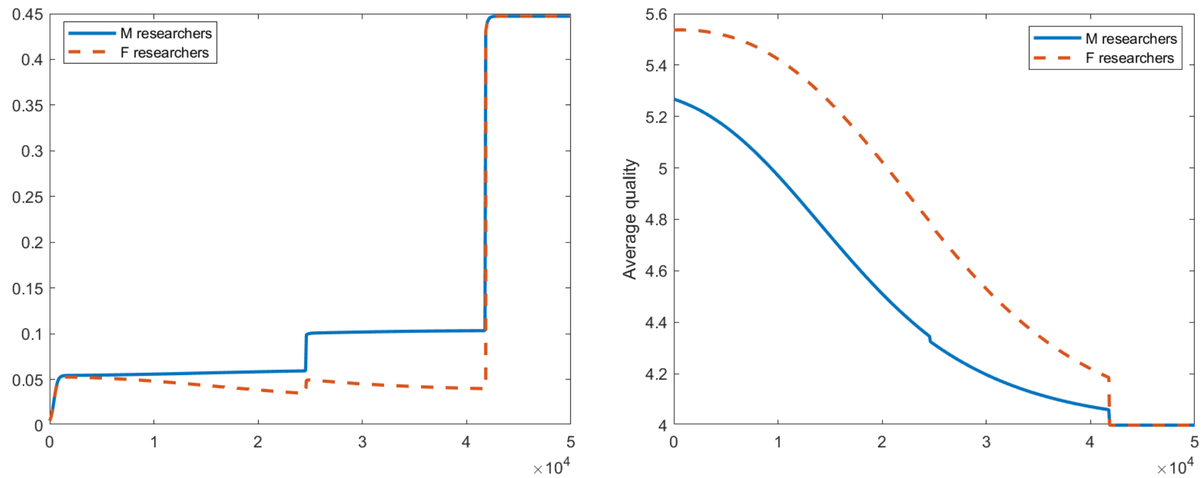
Then $A_t^m \geq A_t^f$. Moreover, if $\lambda_0^{\theta^m} > \frac{C}{\gamma_0\sqrt{\rho}P} > \lambda_0^{\theta^f}$, then $A_t^f \to 1 - \bar{\Lambda}^m$, where $\bar{\Lambda}^m$ is as in part (a.2) of Corollary A.2.

The intuition stems from the fact that when the majority of referees is from the M-group, it is more likely for an M-researchers to be accepted than for a F-researcher, on average. Thus, mass of applicants from the M-group is higher than from the F-group.

Figures A.3a and A.3b show the total masses of M and F applicants and, respectively, the percentage of F applicants over the total application pool. The parameter values are the same as for Figure A.1. Consistently with Corollary A.2, the mass of M applicants is always greater than that of F applicants; furthermore, the latter declines over time. The discrete jumps in these masses occur whenever, for some type θ, the population fraction λ_t^θ falls below the cutoff $C/(\gamma^\theta P)$. In the limit, the fraction of F applicants equals the fraction

Figure A.4: Endogenous entry: Acceptance Rates

(a) Acceptance rate of applicants (b) Average quality



Acceptance rate of M and F applicants (left) and average quality of accepted ones (right). Parameters: $\phi = 0.5742$ ($d = 0.3$), $\gamma_0 = 0.2$, $\rho = 5$, $N = 8$, $P = 1000$, and $C = 4$.

of F researchers of the only surviving type θ^m over the total:

$$\lim_{t\to\infty} \frac{A_t^f}{A_t^m + A_t^f} = \frac{p^{\theta^m,f}}{p^{\theta^m,f} + p^{\theta^m,m}} = \frac{(1-\phi)^N}{\phi^N + (1-\phi)^N} = \frac{0.4258^8}{.4258^8 + .5742^8} = 0.0838$$

Finally, the left panel of Figure A.4 shows the total acceptance rates of M and F applicants. In the initial period, the acceptance rates of M and F applicants are similar. They though diverge in the intermediate period, in which M applicants are accepted more often than the (fewer) F applicants, and then they finally converge, when only type θ^m survives. Interestingly, the right panel shows that the average quality of F researchers is uniformly higher until the time of convergence. This implies that in the initial period our model predicts similar acceptance rates of M and F researchers, even if the latter have higher objective quality. This result is reminiscent of Card et al. (2020), who show that unconditionally, acceptance rates of men- and women-authored papers are similar, but that the average quality of accepted women-authored papers, proxied by their future citations, is higher.

A1.2. Endogenous Selection by Hiring Institutions

The previous section demonstrates that endogenizing the choice of entry into academia may shrink the supply of talent. We now show that the a similar mechanism operates on the demand side: when hiring decisions are based on the expectation of academic success, the anticipation of self-image bias in the refereeing process (Section 3.2.) induces institutions to hire only those types θ that can produce research that is more likely to be "accepted" by the established refereeing population.

Consider the following alternative interpretation of our model. When a hiring institution evaluates a candidate, it takes into account whether or not the candidate will produce quality work that the profession recognizes, or—in the language of Section 3.2.—"accepts." A candidate who is accepted by the profession yields a payoff P to the institution; this reflects e.g. visibility, grant money, or increased ability to attract top students. Hiring a candidate involves a cost C, which may be monetary but may also reflect mentoring resources and/or opportunity cost. This cost is borne by the institution whether or not the candidate is eventually accepted, and it is the same for M and F researchers. If the candidate is eventually not accepted or if the institution does not hire any candidate, the institution's payoff is zero. As above, a candidate of type θ produces quality work with probability γ^θ. To analyze demand effects, we reinterpret the key assumption of Section 3.2. as follows: the hiring institution anticipates that referees are subject to self-image bias, so that a type-θ researcher will be accepted by the profession with probability $\gamma^\theta \lambda_t^\theta$ at the end of time t.

Under these conditions, the institution hires a young researcher of type θ if and only if

$$\gamma^\theta \lambda_t^\theta (P - C) + \left(1 - \lambda_t^\theta \gamma^\theta\right)(-C) > 0 \tag{A.35}$$

This is the same condition as in Equation (A.28) in the previous section. Thus, the mass of established researchers λ_t^θ follows the system dynamics described by Equations (A.29) - (A.30). Proposition A.1 then applies and group imbalance and loss of talent obtains.

Moreover, in the special case of the model of Section 4., under the conditions of case (a.2) Corollary A.2, the system converges, in finite time, to a steady state in which only type θ^m survives. That is, if institutions *only* take acceptance by the profession into account at the hiring stage, type θ^f eventually disappears, even when such type would survive without endogenous selection. Again, this implies talent loss: research characteristics that are (mildly) more common in the F-population disappear.

We can also re-interpret the example in subsection A1.1.1. as a consequence of the hiring practices of hiring institutions. In the absence of endogenous selection, the parametric choices in that example lead to group balance, with both types θ^m and θ^f being represented in the limit. However, if institutions wish to hire only young researchers who are sufficiently likely to be accepted by the *current* population of referees, then group imbalance emerges, as in Figure A.1. Again, in this example type θ^f then disappears completely, as in Figure A.2.

This mechanism with endogenous choice further explains the patterns documented in Section 2.

A1.3. Seniors and Juniors

We now extend the basic model (without endogenous entry) in a different direction, namely, to the case in which there are different levels of seniority in the population of established researchers, with the seniors judging the research of the juniors, before accepting them onto their group. For instance, junior assistant professors may judge candidates from the rookie market and senior professors judge both assistant professors and rookies.

To avoid introducing new symbols, we add a subscript "1" to denote the mass of junior established researchers, and a subscript "2" for the senior established researchers. The difference from the previous case is mainly the mass of candidates of each type θ at each time t. For simplicity, we assume that, at time 0 and thereafter, the mass of seniors is fixed at σ and the mass of juniors is $1-\sigma$, so that the overall population of established researchers has mass 1, as in previous sections. That is, for all t, we must have

$$\sum_\theta \lambda^\theta_{1,t} = 1-\sigma, \quad \sum_\theta \lambda^\theta_{2,t} = \sigma.$$

The flows are similar to before: young researchers are evaluated by all, and juniors are evaluated by seniors only. For each group $g \in \{f, m\}$ and type $\theta \in \Theta$, the flows of juniors $a^{\theta,g}_{1,t}$ and seniors $a^{\theta,g}_{2,t}$ evolve according to

$$a^{\theta,g}_{1,t} = \gamma^\theta \cdot p^{\theta,g} \cdot (\lambda^\theta_{1,t-1} + \lambda^\theta_{2,t-1}) \tag{A.36}$$

$$a^{\theta,g}_{2,t} = \gamma^\theta \cdot \lambda^{\theta,m}_{1,t-1} \cdot \lambda^\theta_{2,t-1}. \tag{A.37}$$

Again, we assume that current seniors are randomly replaced by newly promoted juniors, and current juniors are randomly replaced by newly accepted young researchers. However, we now must take into account the fact that juniors promoted to seniors leave the junior pool. We thus obtain the dynamics

$$\lambda^{\theta,g}_{1,t} = \lambda^{\theta,m}_{1,t-1}\left(1 - \frac{1}{1-\sigma}(a_{1,t} - a_{2,t})\right) + a^{\theta,g}_{1,t} - a^{\theta,g}_{2,t} \tag{A.38}$$

$$\lambda^{\theta,g}_{2,t} = \lambda^{\theta,g}_{2,t-1}\left(1 - \frac{1}{\sigma}a_{2,t}\right) + a^{\theta,g}_{2,t} \tag{A.39}$$

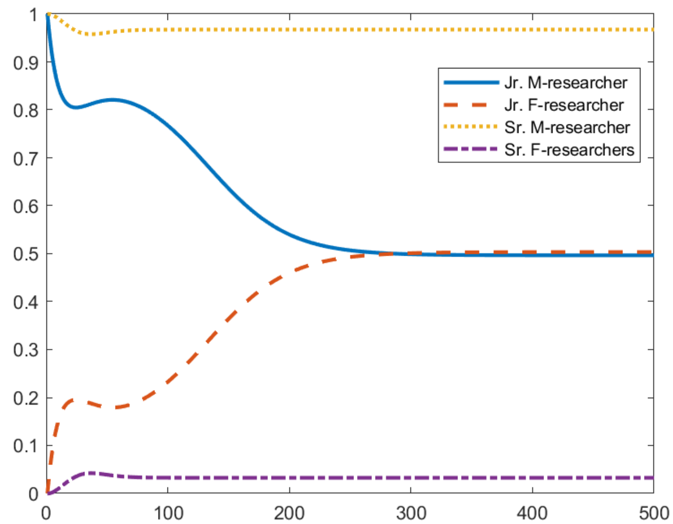
for $g \in \{f, m\}$, where $a_{j,t} = \sum_\theta (a^{\theta,f}_{j,t} + a^{\theta,m}_{j,t})$ for $j = 1, 2$.

The dynamics are far more complex than in the base case, and we rely on numerical simulations.

A1.3.1. Leaky Pipeline

Here we focus on the most interesting case, namely, the fact that this extension can also account for the "leaky pipeline" pattern highlighted in the CSWEP report (Chevalier, 2020).

Figure A.5: Leaky pipeline



Fraction of senior and junior M and F researchers, relative to σ (seniors) and $1-\sigma$ (juniors), when $\lambda_0 = p^m$. Parameters: $\phi = 0.7$, $\gamma_0 = 0.2$, $\rho = 4$, $N = 4$ and $\sigma = 0.5$.

Figure A.5 provides a stark illustration: under the given parametric assumptions, group balance attains among juniors, but not among seniors. A rough intuition is that the self-image bias may not be strong enough to result in a prevalence of θ^m types among juniors, given the constant influx of new researchers with a more balanced distribution of types. However, it may be strong enough if the candidates' types are themselves more biased towards the M researchers' distribution—as is the case for junior up for promotion to the senior rank.

A1.4. Co-authorship

This section briefly explores the implications of our model's dynamics for inferences about the relative (objective) quality of coauthors in a joint project. We show that, consistently with the findings in Sarsons et al. (2021), if research co-authored by a young M-researcher and a young F-researcher is accepted, then the expected quality of the M-researcher is higher. For simplicity, we consider an economy that has reached its steady state, and such that only types θ^m and θ^f are represented in the population of established scholars. Hence, a joint research project is accepted if and only if its vector of characteristics is θ^m or θ^f.

Proposition A.3 *Let the economy be at its steady state with only types θ^f and θ^m surviving. For each researcher of type θ, define $L(\theta) = \sum_{n=1}^{N} \theta_n$ its objective quality. Let a research that is coauthored by type θ^a and θ^b be of type $\theta = \theta^a \vee \theta^b$, where $\vee$ denotes the component-wise maximum. Let researcher $a \in M$ and $b \in F$. Then, conditional on* *acceptance of the joint work, i.e.* $\theta^a \vee \theta^b \in \{\theta^m, \theta^f\}$, *we have*

$$\mathrm{E}[L(\theta^a)|\theta^a \vee \theta^b \in \{\theta^m, \theta^f\}] > \mathrm{E}[L(\theta^b)|\theta^a \vee \theta^b \in \{\theta^m, \theta^f\}]$$

The intuition of the result is that referees are more frequently of type θ^m, and, in addition, θ^m is more frequent in the M population than in the F population. It follows that conditional on the joint work being accepted, it is then more likely it is due for the M characteristics than the F characteristics.

A1.5. Generalized self-image bias

In this section we discuss two extensions of the model to investigate the case in which referees do not only accept researchers who have characteristics identical to their own.

First, in the setting of Section 3., we assume that the set Θ of types is partitioned into subsets Θ_j; a referee of type θ^r drawn from subset Θ_j only accepts applicants of types $\theta^a \in \Theta_j$. That is, we replace Assumption 3 in the main text with

Assumption 3′: there exist disjoint sets $\Theta_1, \ldots, \Theta_J$ such that $\Theta = \Theta_1 \cup \ldots \Theta_J$ and, for every $j = 1, \ldots, J$, referees of type $\theta^r \in \Theta_j$ accept applicants of type θ^a if and only if $\theta^a \in \Theta_j$.

One interpretation is that types in each partition element Θ_j are in some sense "close" or "similar." Another is that each Θ_j represents a "field." In this interpretation, a field may involve different research attributes, but each attribute is only useful in one field.

The dynamics in Eqs. (3)–(4) must then be modified as follows: for every $j = 1, \ldots, J$ and $\theta \in \Theta_j$,

$$a_t^{\theta,g} = \gamma^\theta \lambda_{t-1}^j p^{\theta,g}, \qquad \lambda_t^{\theta,g} = (1 - a_t)\lambda_{t-1}^{\theta,g} + \gamma^\theta \lambda_{t-1}^j p^{\theta,g}, \tag{A.40}$$

where $\lambda_{t-1}^j = \sum_{\theta' \in \Theta_j} \lambda_{t-1}^{\theta'}$.

While this generalization of our model is not covered by our results, we can still invoke them indirectly to analyze it. First, summing over all $\theta \in \Theta_j$ yields

$$a_t^{j,g} = \gamma^j \lambda_{t-1}^j p_\gamma^{j,g}, \qquad \lambda_t^{j,g} = (1 - a_t)\lambda_{t-1}^{j,g} + \gamma^j \lambda_{t-1}^j p_\gamma^{j,g},$$

where $a_t^{j,g} = \sum_{\theta' \in \Theta_j} a_t^{\theta,g}$, $\lambda_\tau^{j,g} = \sum_{\theta' \in \Theta_j} \lambda_\tau^{\theta',g}$, $\gamma^j = \sum_{\theta' \in \Theta_j} \gamma^{\theta'}$, and $p_\gamma^{j,g} = \sum_{\theta' \in \Theta_j} \gamma^{\theta'} p^{\theta',g} / \gamma^j$. These equations are exactly like Eqs. (3)–(4), except that types $\theta \in \Theta$ are replaced with subsets or "fields" Θ_j, $j = 1 \ldots, J$.

Second, replace the symmetry assumption with the following

Assumption 1′: there is a function $\sigma : \Theta \to \Theta$ such that (i) for all $j, k \in \{1, \ldots, J\}$ and $\theta, \theta' \in \Theta_j$, $\sigma(\theta) \in \Theta_k$ implies $\sigma(\theta') \in \Theta_k$; (ii) for all $j = 1, \ldots, J$, if $\sigma(\theta) \in \Theta_k$ for all $\theta \in \Theta_j$, then $\gamma^j = \gamma^k$ and $p_\gamma^{j,m} = p_\gamma^{k,f}$; and (iii) $\sigma(\sigma(\theta)) = \theta$.

Finally, replace the heterogeneity and boundedness assumptions with

Assumption 2′: for some $j \in \{1, \dots, J\}$, if k is such that $\sigma(\theta) \in \Theta_k$ for all $\theta \in \Theta_j$, then $p^{j,m} > p^{k,m}$.

Assumption 4′: for all $j \in \{1, \dots, J\}$, $\gamma^j \cdot (p^{j,m}_\gamma + p^{j,f}_\gamma) = \sum_{\theta \in \Theta_j} \gamma^\theta (p^{m,\theta} + p^{f,\theta}) \leq 1$.

With these assumptions, our results go through unmodified, but apply to subsets, or fields $\Theta_1, \dots, \Theta_J$, rather than to individual types θ. The dynamics of individual types can still be retrieved from Eq. (A.40).

A1.5.1. Similarity in research characteristics

An alternative approach to modeling generalized self-image bias is to consider the parametric model of Section 4., introduce a notion of distance between types, and assume that referees accept young researchers whose characteristics are close enough to their own. Specifically, we assume that referee r of type θ^r accepts the research of young researcher θ if

$$D(\theta^r, \theta) = \sum_n (\theta^r_n - \theta_n)^2 \leq \eta \tag{A.41}$$

where η is a non-negative integer. That is, referee θ^r treats candidate θ as "close enough" if it differs from his or her own type in no more than η characteristics.

Our model calibrated in Section 4. corresponds to $\eta = 0$. If instead $\eta > 0$, the dynamics for λ^θ_t are still as in Eq. (4), but the mass $a^{\theta,g}_t$ of accepted researchers of type θ in group $g \in \{f, m\}$ is given by

$$a^{\theta,g}_t \;=\; \gamma^\theta \sum_{\theta^r : D(\theta^r, \theta) \leq \eta} \lambda^{\theta^r}_{t-1} p^{\theta,g} \tag{A.42}$$

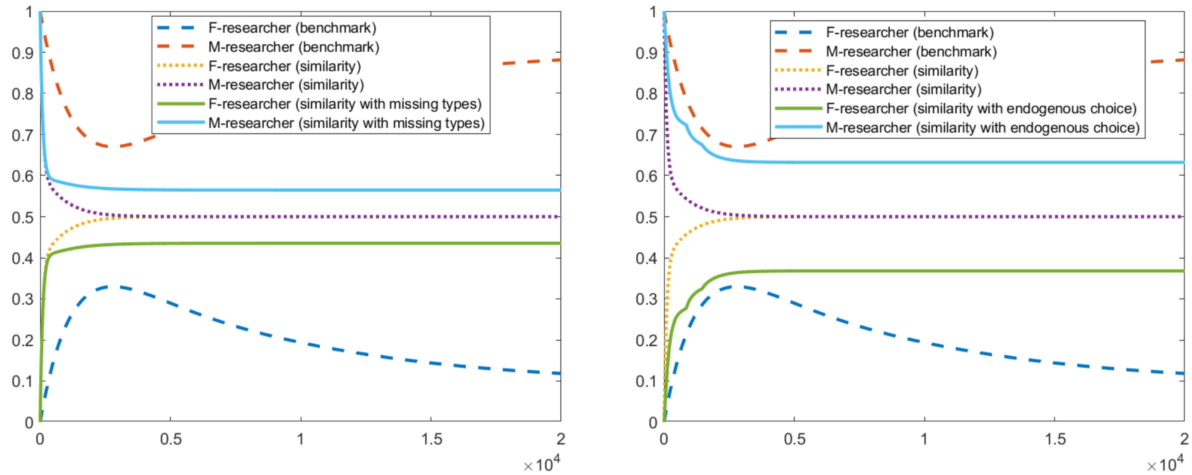
Unfortunately, obtaining general analytical results in this case seems difficult. Therefore, we consider illustrative special cases.

Connected Set of Types The set Θ of types we have considered in the calibration Section 4. enjoys a special structure that is relevant to the relaxed definition of "acceptance" in Eq. (A.41). For every $\eta \geq 1$, and every pair $\theta, \theta' \in \Theta$, there is a finite ordered list $\theta_1, \dots, \theta_K \in \Theta$ such that $\theta_1 = \theta$, $\theta_K = \theta'$, and $D(\theta_k, \theta_{k+1}) \leq \eta$ for all $k = 1, \dots, K-1$. In this sense, we say that $\Theta = \{0, 1\}^N$ is η*-connected* for every $\eta \geq 1$. Of course, being 1-connected implies being η-connected for $\eta > 1$; we shall see in the next subsection that a subset of $\{0, 1\}^N$ may be η-connected for some $\eta > 1$, but for any smaller integer η' (including $\eta' = 1$).

With $\Theta = \{0, 1\}^N$, and for the parameter values used in the examples of Section 4., the relaxed acceptance criterion in Eq. (A.41) leads to convergence. For instance, the left panel of Figure A.6 illustrates the parameterization used in Section 4.. The dashed lines represent the

Figure A.6: Fraction of M and F Researchers under the Research Similarity

(a) Missing Types (b) Endogenous Entry



Fraction of M and F researchers when $\lambda_0 = p^m$. Parameters: $\phi = 0.5742$, which implied $d = 0.3$, $\gamma_0 = 0.2$, $\rho = 5$, $N = 10$, and, under research similarity, $\eta = 1$. Endogenous choice assume $P = 1000$ and $C = 6$

benchmark case $\eta = 0$, where there is no convergence. The dotted lines reflect the assumption that referees accept young researchers that are closely similar to them: specifically, taking $\eta = 1$. Notably, group balance obtains. (The solid lines are discussed in the next section.) Moreover, we have not been able to find parameterizations for which convergence did *not* occur. We conjecture that this is a general property of the special structure of the type space $\Theta = \{0, 1\}^N$. Intuitively, a referee of type θ accepts a positive mass of young researchers of similar, but not identical type θ'; these become referees in the following period, and accept a positive mass of young researchers of type θ'' that type-θ referees would reject; and so on. A contagion argument suggests that, in the limit, the impact of self-image bias should vanish, so that group balance should emerge.

Disconnected Set of Types A subset of $\{0, 1\}^N$ may well be η-disconnected for some η. For a trivial example, $\{\theta^m, \theta^f\}$ is $(N-1)$-disconnected, because each of the N coordinates of θ^f is different from the corresponding coordinate of θ^f. A fortiori, it is η-disconnected for every $\eta \leq N - 1$.

Intuition suggests that the contagion argument given above breaks down with a disconnected set of types. We now verify this intuition. The solid lines in the left panel of Figure A.6 represent the same parameterization as in the previous subsection, with $\eta = 1$, but applied to a state space Θ obtained by randomly removing 20% of the elements of $\{0, 1\}^N$ and suitably renormalizing probabilities. As expected, the system does not attain group balance in the limit.

A model with a disconnected set of types is related to the partitional model considered at the beginning of this section—it is, in a sense, a generalization of that model, in the context of the specific setting of Section 4.. In the partitional model, a referee θ^r accepts type θ if and only if they belong to the same partition element Θ_j. In the similarity model, the distance between θ^r and θ may be greater than η, but there may be some intermediate type θ^ρ that θ^r accepts, and that in turn accepts θ. The similarity-based model essentially allows the sets Θ_j to overlap. The numerical result in this subsection indicates that even with these overlap, convergence to group balance may not be achieved.

A1.5.2. Endogenous Entry

Finally, return to the case in which $\Theta = \{0,1\}^N$ (a connected set of types) but consider endogenous entry, as in Section A1.1.. In this case, even if the connected set of types would lead to convergence (see subsection A1.5.1.), the endogenous entry prevents such convergence, as shown in Section A1.1.1.. This is shown in Figure A.6. Again, the dashed lines and the dotted lines show the total fraction of M- and F-researchers in the benchmark case ($\eta = 0$) and, respectively, the research similarity case ($\eta = 1$). The solid lines now show the the fraction of M- and F-researchers under research similarity ($\eta = 1$) but with endogenous entry. The intuition is the same as the one given in Section A1.1..

In sum, this section suggests that the main results of the paper are robust to a weaker assumption about the referees' selection mechanism.

A2. Omitted or sketched proofs

Proof of Corollary 1: Let $D = \sum_{\theta' \in \Theta^{\max}} \lambda_0^{\theta'} \gamma^{\theta'} (p^{\theta',m} + p^{\theta',f})$; then Eqs. (7) can be written as $\bar{\lambda}^{\tilde{\theta},g} = D^{-1} \lambda_0 \gamma^{\tilde{\theta}} p^{\tilde{\theta},g}$ for $\tilde{\theta} = \theta, \theta'$ and $g = f, m$. Since $\theta' = \sigma(\theta)$, $\gamma^{\theta'} = \gamma^{\theta}$. Letting $K = \gamma^\theta D^{-1}$, we have $\bar{\lambda}^{\tilde{\theta},g} = K p^{\tilde{\theta},g}$ for $\tilde{\theta} = \theta, \theta'$ and $g = f, m$. Then

$$\bar{\lambda}^{\theta,m} + \bar{\lambda}^{\theta',m} = K\left(\lambda_0^{\theta} p^{\theta,m} + \lambda_0^{\theta'} p^{\theta',m}\right) = K\left(\lambda_0^{\theta} p^{\theta',f} + \lambda_0^{\theta'} p^{\theta,f}\right) > K\left(\lambda_0^{\theta'} p^{\theta',f} + \lambda_0^{\theta} p^{\theta,f}\right) = \bar{\lambda}^{\theta,f} + \bar{\lambda}^{\theta',f},$$

where the first equality follows from $p^{\theta,m} = p^{\theta',f}$ and $p^{\theta',m} = p^{\theta,f}$ because $\theta' = \sigma(\theta)$, and the inequality follows because, by assumption, $p^{\theta',f} > p^{\theta,f}$ and $\lambda_0^{\theta} > \lambda_0^{\theta'}$. *Q.E.D.*

Proof of Proposition 4: Write

$$\bar{\Lambda}^g = \sum_{\theta \in \theta^{\max}:\sigma(\theta)=\theta} \bar{\lambda}^{\theta,g} + \sum_{\theta \in \Theta^{\max}:\sigma(\theta)\neq\theta} \bar{\lambda}^{\theta,g} = \sum_{\theta \in \theta^{\max}:\sigma(\theta)=\theta} \bar{\lambda}^{\theta,g} + \frac{1}{2} \sum_{\theta \in \Theta^{\max}:\sigma(\theta)\neq\theta} (\bar{\lambda}^{\theta,g} + \bar{\lambda}^{\sigma(\theta),g}).$$

Letting $D = \sum_{\theta' \in \Theta^{\max}} \lambda_0^{\theta'} \gamma^{\theta'} (p^{\theta',m} + p^{\theta',f})$, $\bar{\lambda}^{\theta,g} = D^{-1} \lambda_0^{\theta} \gamma^{\theta} p^{\theta,g}$.

By definition, $\gamma^{\theta} = \gamma^{\sigma(\theta)}$, $p^{\theta,m} = p^{\sigma(\theta),f}$; moreover, since $\sigma(\sigma(\theta)) = \theta$, also $p^{\sigma(\theta),m} = p^{\theta,f}$.

Thus, consider first $\theta \in \Theta^{\max}$ such that $\sigma(\theta) = \theta$: then $\bar{\lambda}^{\theta,m} = D^{-1}\lambda_0^{\theta}\gamma^{\theta}p^{\theta,m} = D^{-1}\lambda_0^{\theta}\gamma^{\theta}p^{\sigma(\theta),f} = D^{-1}\lambda_0^{\theta}\gamma^{\theta}p^{\theta,f} = \bar{\lambda}^{\theta,f}$. Thus, $\sum_{\theta \in \Theta^{\max}:\sigma(\theta)=\theta} \bar{\lambda}^{\theta,m} = \sum_{\theta \in \Theta^{\max}:\sigma(\theta)=\theta} \bar{\lambda}^{\theta,f}$.

Now consider $\theta \in \Theta^{\max}$ with $\sigma(\theta) \neq \theta$. By the definition of σ, $p^{\theta,m} = p^{\sigma(\theta),f}$. Furthermore, if $p^{\theta,m} = p^{\sigma(\theta),m}$, then also $p^{\theta,m} = p^{\sigma(\theta),m} = p^{\theta,f}$, so $p^{\theta,m} = p^{\sigma(\theta),m} = p^{\theta,f} = p^{\sigma(\theta),f}$ and so $D^{-1}\lambda_0^{\theta}\gamma^{\theta}p^{\theta,m} + D^{-1}\lambda_0^{\sigma(\theta)}\gamma^{\sigma(\theta)}p^{\sigma(\theta),m} = D^{-1}\lambda_0^{\theta}\gamma^{\theta}p^{\theta,f} + D^{-1}\lambda_0^{\sigma(\theta)}\gamma^{\sigma(\theta)}p^{\sigma(\theta),f}$. Therefore, for such θ, $\bar{\lambda}^{\theta,m} + \bar{\lambda}^{\sigma(\theta),m} = \bar{\lambda}^{\theta,f} + \bar{\lambda}^{\sigma(\theta),f}$.

If $\Theta^{\max}$ is homogeneous, then for all $\theta \in \Theta^{\max}$, either $\sigma(\theta) = \theta$, or $\sigma(\theta) \neq \theta$ but $p^{\theta,m} = p^{\sigma(\theta),m}$. In this case, the above decomposition of $\bar{\Lambda}^g$ implies that $\bar{\Lambda}^m = \bar{\Lambda}^f = \frac{1}{2}$ (because the total mass of established researchers is 1).

Otherwise, there is at least one $\theta \in \Theta^{\max}$ for which $\sigma(\theta) \in \Theta^{\max}$ and $p^{\theta,m} \neq p^{\sigma(\theta),m}$, and it is wlog to assume that $p^{\theta,m} > p^{\sigma(\theta),m}$. Corollary 1 shows that, for such types $\theta, \sigma(\theta)$, $\bar{\lambda}^{\theta,m} + \bar{\lambda}^{\sigma(\theta),m} > \bar{\lambda}^{\theta,f} + \bar{\lambda}^{\sigma(\theta),f}$. Therefore, $\bar{\Lambda}^m > \bar{\Lambda}^f$, which implies that $\bar{\Lambda}^m > \frac{1}{2}$. *Q.E.D.*

Proof of Proposition 5: We first claim that, for every $\theta \in \Theta$,

$$a_t^{\theta,m} + a_t^{\sigma(\theta),m} \geq a_t^{\theta,f} + a_t^{\sigma(\theta),f}. \tag{A.43}$$

Notice that, if $\sigma(\theta) = \theta$, the above inequality just says that $a_t^{\theta,m} \geq a_t^{\theta,f}$. Recall that, by the definition and properties of σ, $\gamma^{\theta} = \gamma^{\sigma(\theta)}$, $p^{\theta,m} = p^{\sigma(\theta),f}$, and $p^{\theta,f} = p^{\sigma(\theta),m}$, so also $p^{\theta,m} + p^{\theta,f} = p^{\sigma(\theta),m} + p^{\sigma(\theta),f}$.

Suppose that $p^{\theta,m} \geq p^{\sigma(\theta),m}$. Then, at time 0, $\lambda_0^{\theta} = p^{\theta,m} \geq p^{\sigma(\theta),m} = \lambda_0^{\sigma(\theta)} > 0$. Since, in the notation of Theorem 1, $q^{\theta} = \gamma^{\theta}(p^{\theta,m} + p^{\theta,f}) + \gamma^{\sigma(\theta)}(p^{\sigma(\theta),m} + p^{\sigma(\theta),f}) = q^{\sigma(\theta)}$, by part 3(a) of that Theorem, for every $t > 0$, $\frac{\lambda_t^{\theta}}{\lambda_{t-1}^{\theta}} = \frac{\lambda_t^{\sigma(\theta)}}{\lambda_{t-1}^{\sigma(\theta)}}$, and hence $\frac{\lambda_t^{\theta}}{\lambda_t^{\sigma(\theta)}} = \frac{\lambda_{t-1}^{\theta}}{\lambda_{t-1}^{\sigma(\theta)}} = \frac{\lambda_0^{\theta}}{\lambda_0^{\sigma(\theta)}} \geq 1$. Thus, $\lambda_t^{\theta} \geq \lambda_t^{\sigma(\theta)}$ for all $t > 0$ as well.

The argument just given applies verbatim if one replaces "$\geq$" in "$p^{\theta,m} \geq p^{\sigma(\theta,m)}$" and "$\lambda_t^{\theta} \geq \lambda_t^{\sigma(\theta)}$" with "$>$", "$<$" or "$=$." Thus, for all $t \geq 0$, $\lambda_t^{\theta} > \lambda_t^{\sigma(\theta)}$ if $p^{\theta,m} > p^{\sigma(\theta),m}$; $\lambda_t^{\theta} < \lambda_t^{\sigma(\theta)}$ if $p^{\theta,m} < p^{\sigma(\theta),m}$; and $\lambda_t^{\theta} = \lambda_t^{\sigma(\theta)}$ if $p^{\theta,m} = p^{\sigma(\theta),m}$. Therefore, letting $\bar{\gamma} = \gamma^{\theta} = \gamma^{\sigma(\theta)}$,

$$\begin{aligned} a_t^{\theta,m} + a_t^{\sigma(\theta),m} \geq a_t^{\theta,f} + a_t^{\sigma(\theta),f} \Leftrightarrow \quad & \bar{\gamma}(\lambda_{t-1}^{\theta}p^{\theta,m} + \lambda_{t-1}^{\sigma(\theta)}p^{\sigma(\theta),m}) \geq \bar{\gamma}(\lambda_{t-1}^{\theta}p^{\theta,f} + \lambda_{t-1}^{\sigma(\theta)}p^{\sigma(\theta),f}) \\ \Leftrightarrow \quad & \lambda_{t-1}^{\theta}[p^{\theta,m} - p^{\theta,f}] \geq \lambda_{t-1}^{\sigma(\theta)}[p^{\sigma(\theta),f} - p^{\sigma(\theta),m}] \\ \Leftrightarrow \quad & [\lambda_{t-1}^{\theta} - \lambda_{t-1}^{\sigma(\theta)}] \cdot [p^{\theta,m} - p^{\sigma(\theta),m}] \geq 0, \end{aligned}$$

where the last step follows from $p^{\theta,m} = p^{\sigma(\theta),f}$ and $p^{\theta,f} = p^{\sigma(\theta),m}$.

If $p^{\theta,m} = p^{\sigma(\theta),m}$, then both terms in square brackets equal zero, so equality obtains; in particular, this is true if $\theta = \sigma(\theta)$. If $p^{\theta,m} > p^{\sigma(\theta),m}$, then both terms are positive, and if $p^{\theta,m} < p^{\sigma(\theta),m}$, then both terms are negative. Thus, in all cases, the last inequality, and hence Eq. (A.43), holds, and is strict unless $p^{\theta,m} = p^{\sigma(\theta),m}$; furthermore, if $\theta = \sigma(\theta)$, then $a_t^{\theta,m} = a_t^{\theta,f}$.

Now consider an arbitrary $\bar{\gamma} \in \{\gamma^\theta : \theta \in \Theta\}$. Then

$$\begin{aligned}\sum_{\theta:\gamma^\theta=\bar{\gamma}} a_t^{\theta,m} &= \sum_{\theta:\gamma^\theta=\bar{\gamma},\theta=\sigma(\theta)} a_t^{\theta,m} + \sum_{\theta:\gamma^\theta=\bar{\gamma},\theta\neq\sigma(\theta)} a_t^{\theta,m} = \\ &= \sum_{\theta:\gamma^\theta=\bar{\gamma},\theta=\sigma(\theta)} a_t^{\theta,m} + \frac{1}{2}\sum_{\theta:\gamma^\theta=\bar{\gamma},\theta\neq\sigma(\theta)} [a_t^{\theta,m} + a_t^{\sigma(\theta),m}] \geq \\ &\geq \sum_{\theta:\gamma^\theta=\bar{\gamma},\theta=\sigma(\theta)} a_t^{\theta,f} + \frac{1}{2}\sum_{\theta:\gamma^\theta=\bar{\gamma},\theta\neq\sigma(\theta)} [a_t^{\theta,f} + a_t^{\sigma(\theta),f}] = \sum_{\theta:\gamma^\theta=\bar{\gamma}} a_t^{\theta,f}.\end{aligned}$$

The second equality follows from the fact that $\gamma^{\sigma(\theta)} = \gamma^\theta$, so that adding $a_t^{\theta,m} + a_t^{\sigma(\theta),m}$ over all θ with $\theta \neq \sigma(\theta)$ with productivity $\bar{\gamma}$ counts each such type twice. The inequality follows from Eq. (A.43) and the fact that $a_t^{\theta,m} = a_t^{\theta,f}$ if $\theta = \sigma(\theta)$. This inequality is strict if there is some type θ for which $p^{\theta,m} > p^{\sigma(\theta),m}$, which implies $\theta \neq \sigma(theta)$, because in that case we showed above that $a_t^{\theta,m} + a_t^{\sigma(\theta),m} > a_t^{\theta,f} + a_t^{\sigma(\theta),f}$. Finally, the last equality follows by repeating the first two steps backwards, for F-group researchers. *Q.E.D*

Lemma A.1 *For all parameter values and initial conditions, and for all $\theta \in \Theta$ and $t \geq 1$,*

$$\frac{\lambda_t^\theta}{\lambda_{t-1}^\theta} = (1-a_t) + \gamma^\theta(p^{\theta,m} + p^{\theta,f}) \quad \textit{and} \quad \frac{a_t^\theta}{a_{t-1}^\theta} = \frac{a_t^{\theta,m}}{a_{t-1}^{\theta,m}} = \frac{a_t^{\theta,f}}{a_{t-1}^{\theta,f}} = \frac{\lambda_{t-1}^\theta}{\lambda_{t-2}^\theta} \quad \textit{if } t \geq 2.$$

Proof: From Eq. (4), $\lambda_t^\theta = \lambda_t^{\theta,m} + \lambda_t^{\theta,f} = (\lambda_{t-1}^{\theta,m} + \lambda_{t-1}^{\theta,f})(1-a_t) + \gamma^\theta(p^{\theta,m} + p^{\theta,f})$, which yields the first equation because $\lambda_\tau^\theta > 0$ for all θ and τ. From Eq. (3), for $t \geq 2$,

$$\frac{a_t^{\theta,g}}{a_{t-1}^{\theta,g}} = \frac{\lambda_{t-1}^\theta \gamma^\theta p^{\theta,g}}{\lambda_{t-2}^\theta \gamma^\theta p^{\theta,g}} = \frac{\lambda_{t-1}^\theta}{\lambda_{t-2}^\theta} \quad \text{and} \quad \frac{a_t^\theta}{a_{t-1}^\theta} = \frac{\lambda_{t-1}^\theta \gamma^\theta(p^{\theta,m} + p^{\theta,f})}{\lambda_{t-2}^\theta \gamma^\theta(p^{\theta,m} + p^{\theta,f})} = \frac{\lambda_{t-1}^\theta}{\lambda_{t-2}^\theta}.$$

Q.E.D.

Proof of Proposition 6: Let $a_t^g = \sum_{\hat{\theta}} a_t^{\hat{\theta},g}$ and, to simplify the notation, $a^{s,g} = a^{\theta,g} + a^{\theta',g}$ ("s" stands for "symmetric types"). As in the proof of Proposition 5, since by assumption $p^{\theta,m} > p^{\theta',m}$, for all t, $a_t^{s,m} > a_t^{s,f}$. On the other hand, since $\sigma(\theta_1) = \theta_1$ and $\sigma(\theta_0) = \theta_0$, $a_t^{\theta_1,m} = a_t^{\theta_1,f}$, and $a_t^{\theta_0,m} = a_t^{\theta_0,f}$. Therefore, $a_t^m > a_t^f$, which implies that the weight on $\gamma^\theta = \gamma^{\theta'} \equiv \gamma^s$ for accepted M researchers is

$$\frac{a_t^{s,m}}{a_t^m} = 1 - \frac{a_t^{\theta_1,m} + a_t^{\theta_0,m}}{a_t^m} = 1 - \frac{a_t^{\theta_1,f} + a_t^{\theta_0,f}}{a_t^m} > 1 - \frac{a_t^{\theta_1,f} + a_t^{\theta_0,f}}{a_t^f} = \frac{a_t^{s,f}}{a_t^f}.$$

Similarly, $a_t^m > a_t^f$ and $a_t^{\theta_0,m} = a_t^{\theta_0,f}$, $a_t^{\theta_1,m} = a_t^{\theta_1,f}$ imply

$$\frac{a_t^{\theta_0,m}}{a_t^m} < \frac{a_t^{\theta_0,f}}{a_t^f}, \qquad \frac{a_t^{\theta_1,m}}{a_t^m} < \frac{a_t^{\theta_1,f}}{a_t^f}.$$

Moreover, we claim that, $a_t^{\theta_1,g} > a_t^{\theta_0,g}$. For $t=0$, $a_0^{\theta_1,g} = p^{\theta_1,m}\gamma^{\theta_1}p^{\theta_1,g} > p^{\theta_0,m}\gamma^{\theta_0}p^{\theta_0,g} = a_0^{\theta_0,g}$, because $p^{\theta_0,g} = p^{\theta_1,g}$ but $\gamma^{\theta_1} > \gamma^{\theta_0}$. Inductively, from Lemma A.1 (proof: Appendix A2.),

$$a_t^{\theta_1,g} = a_{t-1}^{\theta_1,g} \cdot \frac{a_t^{\theta_1,g}}{a_{t-1}^{\theta_1,g}} = a_{t-1}^{\theta_1,g}\left(1 - a_{t-1} + \gamma^{\theta_1}(p^{\theta_1,m} + p^{\theta_1,f}\right) > a_{t-1}^{\theta_1,g}\left(1 - a_{t-1} + \gamma^{\theta_0}(p^{\theta_0,m} + p^{\theta_0,f}\right) >$$

$$> a_{t-1}^{\theta_0,g}\left(1 - a_{t-1} + \gamma^{\theta_0}(p^{\theta_0,m} + p^{\theta_0,f}\right) = a_{t-1}^{\theta_0,g}\frac{a_t^{\theta_0,g}}{a_{t-1}^{\theta_0,g}} = a_t^{\theta_0,g}.$$

It then follows that

$$0 < \frac{a_t^{\theta_0,f}}{a_t^f} - \frac{a_t^{\theta_0,m}}{a_t^m} = \frac{a_t^{\theta_0,f}}{a_t^f} - \frac{a_t^{\theta_0,f}}{a_t^m} < \left(\frac{a_t^{\theta_1,f}}{a_t^{\theta_0,f}}\right)\cdot\left(\frac{a_t^{\theta_0,f}}{a_t^f} - \frac{a_t^{\theta_0,f}}{a_t^m}\right) =$$
$$= \frac{a_t^{\theta_1,f}}{a_t^f} - \frac{a_t^{\theta_1,f}}{a_t^m} = \frac{a_t^{\theta_1,f}}{a_t^f} - \frac{a_t^{\theta_1,m}}{a_t^m};$$

the first inequality follows from $a_t^{s,f} < a_t^{s,m}$ and $a_t^{\theta_0,f} = a_t^{\theta_0,m}$ and $a_t^{\theta_1,f} = a_t^{\theta_1,m}$, the next equality from $a_t^{\theta_0,m} = a_t^{\theta_0,f}$, the second inequality from $a_t^{\theta_1,f} > a_t^{\theta_0,f} > 0$ and the fact that the difference of fractions is positive, and the last equality from $a_t^{\theta_1,m} = a_t^{\theta_1,f}$. Now

$$E[\gamma|F] = \frac{\gamma^{\theta_0}a_t^{\theta_0,f} + \gamma^s a_t^{s,f} + \gamma^{\theta_1}a_t^{\theta_1,f}}{a_t^f}; \quad E[\gamma|M] = \frac{\gamma^{\theta_0}a_t^{\theta_0,m} + \gamma^s \times a_t^{s,m} + \gamma^{\theta_1}a_t^{\theta_1,m}}{a_t^m}$$

which, since $a_t^{s,g} = 1 - a_t^{\theta_0,g} - a_t^{\theta_1,g}$, implies

$$E[\gamma|F] = -(\gamma^s - \gamma^{\theta_0})\frac{a_t^{\theta_0,f}}{a_t^f} + \gamma^s + (\gamma^{\theta_1} - \gamma^s)\frac{a_t^{\theta_1,f}}{a_t^f}; \quad E[\gamma|M] = -(\gamma^s - \gamma^{\theta_0})\frac{a_t^{\theta_0,m}}{a_t^m} + \gamma^s + (\gamma^{\theta_1} - \gamma^s)\frac{a_t^{\theta_1,m}}{a_t^m}$$

and therefore, since $\frac{a_t^{\theta_1,f}}{a_t^f} - \frac{a_t^{\theta_1,m}}{a_t^m} > \frac{a_t^{\theta_0,f}}{a_t^f} - \frac{a_t^{\theta_0,m}}{a_t^m}$ and $\gamma^s - \gamma^{\theta_0} \leq \gamma^{\theta_1} - \gamma^s$,

$$E[\gamma|F] - E[\gamma|M] = -(\gamma^s - \gamma^{\theta_0})\left(\frac{a_t^{\theta_0,f}}{a_t^f} - \frac{a_t^{\theta_0,m}}{a_t^m}\right) + (\gamma^{\theta_1} - \gamma^s)\left(\frac{a_t^{\theta_1,f}}{a_t^f} - \frac{a_t^{\theta_1,m}}{a_t^m}\right) > 0.$$

This yields (i). For (ii), if $\Theta^{\max}$ is heterogeneous, then it must contain θ and θ' only; and if not, it can only contain θ_1 because $p^{\theta_0,m} + p^{\theta_0,f} = p^{\theta_1,m} + p^{\theta_1,f}$ but $\gamma^{\theta_1} > \gamma^{\theta_0}$. *Q.E.D.*

Proof of Corollary 2: From Eq. 7 in Proposition 3, since $\gamma^{\theta'} = \gamma^{\sigma(\theta)} = \gamma^{\theta}$,

$$\frac{\bar{\lambda}^{\theta,m}}{\bar{\lambda}^{\theta,m} + \bar{\lambda}^{\theta',m}} = \frac{\lambda_0^{\theta}p^{\theta,m}}{\lambda_0^{\theta}p^{\theta,m} + \lambda_0^{\theta'}p^{\theta',m}} = \frac{[\lambda_0^{\theta}]^2}{[\lambda_0^{\theta}]^2 + [\lambda_0^{\theta'}]^2} > 0.5 :$$

the second equality follows from $\lambda_0 = p^m$, and the inequality follows from $\lambda_0^{\theta} > \lambda_0^{\theta'}$. On the other hand,

$$\frac{\bar{\lambda}^{\theta,f}}{\bar{\lambda}^{\theta,f} + \bar{\lambda}^{\theta',f}} = \frac{\lambda_0^{\theta}p^{\theta,f}}{\lambda_0^{\theta}p^{\theta,f} + \lambda_0^{\theta'}p^{\theta',f}} = \frac{p^{\theta,m}p^{\theta,f}}{p^{\theta,m}p^{\theta,f} + p^{\theta',m}p^{\theta',f}} = \frac{p^{\theta,m}p^{\theta,f}}{p^{\theta,m}p^{\theta,f} + p^{\theta,f}p^{\theta,m}} = 0.5 :$$

the second equality follows from $\lambda_0 = p^m$, and the third from $p^{\theta',m} = p^{\sigma(\theta'),f} = p^{\theta,f}$ and similarly $p^{\theta',f} = p^{\theta,m}$. The remaining equality and inequality are immediate. *Q.E.D.*

Proof of Corollary 3: From Eq. (7) in Proposition 3,

$$\frac{\bar{\lambda}^{\theta,m}}{\bar{\lambda}^{\theta,m} + \bar{\lambda}^{\theta,f}} = \frac{\lambda_0^\theta p^{\theta,m}}{\lambda_0^\theta p^{\theta,m} + \lambda_0^\theta p^{\theta,f}} = \frac{p^{\theta,m}}{p^{\theta,m} + p^{\theta,f}} = \frac{p^{\theta,m}}{p^{\theta,m} + p^{\theta',m}} = \frac{\lambda_0^\theta}{\lambda_0^\theta + \lambda_0^{\theta'}} > 0.5:$$

the third equality follwos from $p^{\theta,f} = p^{\sigma(\theta),m} = p^{\theta',m}$ and the fourth from the assumption that $\lambda_0^\theta > \lambda_0^{\theta'}$.

The other equality in the Corollary follows because, if $\lambda_0 = p^m$, then Eq. (7) implies that $\bar{\lambda}^{\theta',f} = \bar{\lambda}^{\theta,m}$ and $\bar{\lambda}^{\theta',m} = \bar{\lambda}^{\theta,f}$. *Q.E.D.*

Proof of Corollary 4: rewrite $F(y)$ as

$$F(y) = \frac{(y\overline{\lambda}^{\theta,f} + (1-y)\overline{\lambda}^{\theta',f})}{(y\overline{\lambda}^{\theta} + (1-y)\overline{\lambda}^{\theta'})} = \frac{(y \cdot 0.5(\overline{\lambda}^{\theta,f} + \overline{\lambda}^{\theta',f}) + (1-y)0.5(\overline{\lambda}^{\theta,f} + \overline{\lambda}^{\theta',f}))}{(y\overline{\lambda}^{\theta} + (1-y)\overline{\lambda}^{\theta'})} =$$
$$= \frac{0.5(\overline{\lambda}^{\theta,f} + \overline{\lambda}^{\theta' f})}{(y\overline{\lambda}^{\theta} + (1-y)\overline{\lambda}^{\theta'})} = \frac{0.5\overline{\gamma}(\overline{\lambda}^{\theta,f} + \overline{\lambda}^{\theta' f})}{P(y)}$$

where the second equality follows from Corollary 2, and the fourth from the definition of $P(y)$. This shows that $F(y)P(y)$ is a constant, independent of y. Since $P(y)$ increases in y, $F(y)$ must decrease in y. *Q.E.D.*

Proof of Corollary 5: the argument was given in the main text.

Lemma A.2 *Assume that, for every $\theta \in \Theta$, γ^θ, $p^{\theta,m}$ and $p^{\theta,f}$ are as defined in Section 4.. Then, for every $\phi \in (\frac{1}{2}, 1)$, N even, $\gamma_0 \in (0,1)$, and $\rho \in (1, \frac{1}{\gamma_0})$:*

1. *the set of maximizers of $\gamma^\theta \cdot (p^{\theta,m} + p^{\theta,f})$ is $\{\theta^m, \theta^f\}$ if $\rho < \bar{\rho}(\phi, N)$ and $\{\theta^*\}$ if $\rho > \bar{\rho}(\phi, N)$. (Recall that $\bar{\rho}(\cdot)$ is defined in Eq. (14).)*
2. *$0 < \gamma^\theta \cdot [p^{\theta,m} + p^{\theta,f}] \leq 1$.*
3. *there is $\bar{N} > 0$ such that, for all even $N \geq \bar{N}$, the maximizers of $\gamma^\theta \cdot (p^{\theta,m} + p^{\theta,f})$ are θ^m and θ^f.*

Proof: Write

$$p^{\theta,m} = \phi^{\sum_{n=1}^{N/2}\theta_n}(1-\phi)^{N/2 - \sum_{n=1}^{N/2}\theta_n} \cdot (1-\phi)^{\sum_{n=N/2+1}^{N}\theta_n}\phi^{N/2 - \sum_{n=N/2+1}^{N}\theta_n} =$$
$$= \phi^{N/2 + \sum_{n=1}^{N/2}\theta_n - \sum_{n=N/2+1}^{N}\theta_n}(1-\phi)^{N/2 + \sum_{n=N/2+1}^{N}\theta_n - \sum_{n=1}^{N/2}\theta_n} =$$
$$= \phi^{N/2}(1-\phi)^{N/2}\left(\frac{\phi}{1-\phi}\right)^{\sum_{n=1}^{N/2}\theta_n - \sum_{n=N/2+1}^{N}\theta_n}.$$

Similarly

$$p^{\theta,f} = \phi^{N/2}(1-\phi)^{N/2}\left(\frac{\phi}{1-\phi}\right)^{\sum_{n=N/2+1}^{N}\theta_n - \sum_{n=1}^{N/2}\theta_n}.$$

Then $F(\theta) \equiv \gamma^\theta(p^{\theta,m} + p^{\theta,f})$ equals

$$\gamma_0\,\rho^{\sum_n \theta_n/N}\cdot\phi^{N/2}(1-\phi)^{N/2}\left[\left(\frac{\phi}{1-\phi}\right)^{\sum_{n=1}^{N/2}\theta_n - \sum_{n=N/2+1}^{N}\theta_n} + \left(\frac{\phi}{1-\phi}\right)^{-\sum_{n=1}^{N/2}\theta_n + \sum_{n=N/2+1}^{N}\theta_n}\right].$$

Since Θ is finite, there exists at least one maximizer θ of $F(\cdot)$. We claim that, if θ satisfies $\theta_n = \theta_m = 0$ for some $n \in \{1,\ldots,N/2\}$ and $m \in \{N/2+1,\ldots,N\}$, then it is not a maximizer. To see this, define θ' by $\theta'_\ell = \theta_\ell$ for $\ell \in \{1,\ldots,N\}\setminus\{n,m\}$ and $\theta'_n = \theta'_m = 1$. Then $\sum_n \theta'_n > \sum_n \theta_n$, so for $\rho > 1$, $\gamma^{\theta'} > \gamma^\theta$. On the other hand, the term in square brackets is the same for θ and θ' (and it is strictly positive). Hence, θ is not a maximizer of $F(\cdot)$. It follows that the only candidate maximizers of $F(\cdot)$ have either $\theta_n = 1$ for all $n = 1,\ldots,N/2$, or $\theta_n = 1$ for all $n = N/2,\ldots,N$, or both.

If $\theta_n = 1$ for $n = 1,\ldots,N/2$, then $F(\theta) = F(\theta')$, where $\theta'_n = 1$ for $n = N/2+1,\ldots,N$ and $\theta'_n = \theta_{n+N/2}$ for $n = 1,\ldots,N/2$. Hence, it is enough to consider θ such that $\theta_n = 1$ for $n = N/2+1,\ldots,N$. Let Θ^f be the collection of such types, and notice that it contains both θ^f (for which $\theta^f_n = 0$ for $n = 1,\ldots,N/2$) and $\theta^* = (1,\ldots,1)$. We show that the maximizer of $F(\cdot)$ on Θ^f is either θ^f or θ^*.

For each $\theta \in \Theta^f$, factoring out all terms not involving $\sum_{n=1}^{N/2}\theta_n$, $F(\theta)$ is proportional to

$$\rho^{\sum_{n=1}^{N/2}\theta_n/N}\cdot\left[\left(\frac{\phi}{1-\phi}\right)^{\sum_{n=1}^{N/2}\theta_n} + \left(\frac{1-\phi}{\phi}\right)^{\sum_{n=1}^{N/2}\theta_n}\right].$$

Hence, $F(\theta)$ is proportional to $\tilde{F}(\sum_{n=1}^{N/2}\theta_n)$, where $\tilde{F} : [0,\frac{1}{2}] \to \mathbb{R}_+$ is defined by

$$\tilde{F}(x) = \rho^x\left[\left(\frac{\phi}{1-\phi}\right)^x + \left(\frac{1-\phi}{\phi}\right)^x\right].$$

The functions $x \mapsto \rho^{\frac{x}{N}}\Phi^x = \left(\rho^{\frac{1}{N}}\right)^x\Phi^x = \left(\rho^{\frac{1}{N}}\cdot\Phi\right)^x$, for $\Phi = \frac{\phi}{1-\phi} \neq 1$ and $\Phi = \frac{1-\phi}{\phi} \neq 1$ respectively, are non-constant and exponential, hence strictly convex on $[0,\frac{1}{2}]$. Hence, $\tilde{F}(\cdot)$ is also strictly convex on $[0,\frac{1}{2}]$, so its maximum is either at 0 or at $\frac{1}{2}$. Correspondingly, $F(\cdot)$ attains a maximum either at θ^f or at θ^* on the set Θ^f.

To conclude the proof of Claim 1, we calculate the values attained by $F(\cdot)$ at θ^f, θ^*:

$$F(\theta^f) = \gamma_0\sqrt{\rho}\cdot[(1-\phi)^N + \phi^N]; \quad F(\theta^*) = \gamma_0\rho\cdot 2\phi^{N/2}(1-\phi)^{N/2}.$$

Dividing $F(\theta^*)$ and $F(\theta^f)$ by $\gamma_0\sqrt{\rho}\phi^{N/2}(1-\phi)^{N/2}$ and comparing the resulting quantities, we conclude that θ^* is (uniquely) optimal iff

$$2\sqrt{\rho} > \left[\left(\frac{\phi}{1-\phi}\right)^{-\frac{N}{2}} + \left(\frac{1-\phi}{\phi}\right)^{-\frac{N}{2}}\right]$$

which reduces to the condition in Claim 1

For Claim 2, we show that $(1-\phi)^N + \phi^N \leq 1$ and $\phi^{N/2}(1-\phi)^{N/2} \leq \frac{1}{2}$; this is sufficient, because $\gamma_0 \in (0,1)$ and $\rho \in (1, \frac{1}{\gamma_0})$ by assumption, so also $\gamma_0\sqrt{\rho} \leq \gamma_0\rho < 1$.

The function $N \mapsto (1-\phi)^N + \phi^N$ is strictly decreasing in N, so it is enough to prove the claim for $N = 2$. In this case, $(1-\phi)^2 + \phi^2 = 1 - 2\phi + \phi^2 + \phi^2 = 1 + 2\phi(\phi - 1) < 1$, because $\phi < 1$. Similarly, $N \mapsto [\phi(1-\phi)]^{N/2}$ is decreasing in N, and for $N = 2$ it reduces to $\phi(1-\phi) = \phi - \phi^2$; this is concave and maximized at $\phi = \frac{1}{2}$, where it takes the value $\frac{1}{4} < \frac{1}{2}$.

Finally, for Claim 3, as $N \to \infty$, the first term in the rhs of Eq. (14) converges to zero, but the second diverges to infinity. Thus, for N large, only θ^m and θ^f maximize $F(\cdot)$. *Q.E.D.*

Proof of Corollary 6: For part (a), Eq. (15) follows from part 1 of Lemma A.2 and part (i) of Proposition 2. Finally, for Eq. (16), part (ii) of Proposition 3 and the fact that $\lambda_0 = p^m$ and $\gamma^{\theta^m} = \gamma^{\theta^f}$ imply that

$$\begin{aligned}\bar{\lambda}^{\theta^m,m} &= \frac{\lambda_0^{\theta^m}\gamma^{\theta^m}p^{\theta^m,m}}{\lambda_0^{\theta^m}\gamma^{\theta^m}(p^{\theta^m,m} + p^{\theta^m,f}) + \lambda_0^{\theta^f}\gamma^{\theta^f}(p^{\theta^f,m} + p^{\theta^f,f})} = \\ &= \frac{\phi^{2N}}{\phi^N(\phi^N + (1-\phi)^N) + (1-\phi)^N((1-\phi)^N + \phi)^N} = \frac{\phi^{2N}}{(\phi^N + (1-\phi)^N)^2}\end{aligned}$$

and analogously

$$\bar{\lambda}^{\theta^m,f} = \bar{\lambda}^{\theta^f,f} = \frac{\phi^N(1-\phi)^N}{(\phi^N + (1-\phi)^N)^2}, \quad \bar{\lambda}^{\theta^f,m} = \frac{(1-\phi)^{2N}}{(\phi^N + (1-\phi)^N)^2}.$$

Therefore,

$$\bar{\Lambda}^m = \bar{\lambda}^{\theta^m,m} + \bar{\lambda}^{\theta^f,m} = \frac{\phi^{2N} + (1-\phi)^{2N}}{(\phi^N + (1-\phi)^N)^2} = \frac{1 + \left(\frac{\phi}{1-\phi}\right)^{2N}}{1 + 2\left(\frac{\phi}{1-\phi}\right)^N + \left(\frac{\phi}{1-\phi}\right)^{2N}}$$

and Proposition 4 ensures that $\bar{\Lambda}^m > 0.5$ regardless of the parameterization. This completes the proof of part (a).

For part (b), part 1 of Lemma A.2 and part (i) of Proposition 2 imply that $\Theta^{\max} = \{\theta^*\}$. Then, since $p^{\theta^*,m} = p^{\theta^*,f} = \phi^{N/2}(1-\phi)^{N/2}$, part (ii) in Proposition 3 implies that $\bar{\lambda}^{\theta^*,m} = \bar{\lambda}^{\theta^*,f} = \frac{1}{2}$. Consequently, $\bar{\Lambda}^m = \bar{\Lambda}^f = \frac{1}{2}$ as well, and these conclusions are independent of λ_0.

Part (c) now follows from part 3 of Lemma A.2 and part (a) of this Corollary. Finally, part (d) follows from Eq. (16) by taking limits as $N \to \infty$. *Q.E.D.*

A3. Proof of the results in Online Appendix A1.

Proof of Proposition A.1: let $\Theta_{-1} = \Theta$ and $t(-1) = 0$. Also let $\lambda_{0,0}^m = \lambda_{1,0}^m = \lambda_0^m$, $\lambda_{0,0}^f = \lambda_{1,0}^f = \lambda_0^f$, and $\lambda_{0,0} = \lambda_{1,0} = \lambda_{1,0}^m + \lambda_{1,0}^f$. Finally, let $\Theta_0 = \left\{\theta \in \Theta : \lambda_{1,0}^\theta \geq \frac{C}{\gamma^\theta P}\right\}$.

For $j \geq 0$, say that *Conditions $C(j)$ hold* if there is a set $\Theta_j \subseteq \Theta_{j-1}$, a period $t(j) > t(j-1)$, and for $\tau = 0, \ldots, t(j) - t(j-1)$, vectors $\lambda^m_{\tau,j}, \lambda^f_{\tau,j}, \lambda_{\tau,j} \in \mathbb{R}^\Theta_+$ such that

(i) for $0 \leq \tau \leq t(j) - t(j-1)$, $\lambda^m_{\tau,j} = \lambda^m_{t(j-1)+\tau}$, $\lambda^f_{\tau,j} = \lambda^f_{t(j-1)+\tau}$, and $\lambda_{\tau,j} = \lambda^m_{\tau,j} + \lambda^f_{\tau,j}$;

(ii) for $0 \leq \tau < t(j) - t(j-1)$, $\lambda^\theta_{\tau,j} \geq \frac{C}{\gamma^\theta P}$ for all $\theta \in \Theta_j$;

(iii) $\lambda^\theta_{\tau,j} < \frac{C}{\gamma^\theta P}$ for $0 \leq \tau \leq t(j) - t(j-1)$ and all $\theta \in \Theta \setminus \Theta_j$, and $\lambda^{\theta_0}_{t(j)-t(j-1),j} < \frac{C}{\gamma^{\theta_0} P}$ for some $\theta_0 \in \Theta_j$.

We claim that, for every $k \geq 0$, if either $k = 0$ or $k > 0$ and Conditions $C(k-1)$ hold, then either Conditions $C(k)$ hold as well, with $\Theta_k \subsetneq \Theta_{k-1}$ in case $k > 0$, or else there exist vectors $\lambda^m_{\tau,k}, \lambda^f_{\tau,k}, \lambda_{\tau,k} \in \mathbb{R}^\Theta_+$ for all $\tau \geq 1$ such that (i) holds for $j = k$, and $\lambda^\theta_{\tau,j} \geq \frac{C}{\gamma^\theta P}$ for all $\theta \in \Theta_k$. In the latter case, if the sequences of such vectors converge, then $\lim_{\tau \to \infty} \lambda^m_{\tau,k} = \lim_{t \to \infty} \lambda^m_t$ and similarly for $\lambda^f_{\tau,k}$ and $\lambda_{\tau,k}$.

Let $\lambda^{\theta,g}_{0,k} = \lambda^{\theta,g}_{t(k-1)}$ for $g = f, m$; also let $\lambda_{0,k} = \lambda^m_{0,k} + \lambda^f_{0,k}$. Let $\Theta_k = \left\{\theta \in \Theta : \lambda^\theta_{0,k} \geq \frac{C}{\gamma^\theta P}\right\}$. If $k = 0$, then $\Theta_0 \subseteq \Theta = \Theta_{-1}$. Otherwise, $C(k-1)$ must hold, so $\lambda_{0,k} = \lambda_{t(k-1)} = \lambda_{t(k-1)-t(k-2),k-1}$. By (iii), if $\theta \notin \Theta_{k-1}$ then $\lambda^\theta_{0,k} = \lambda^\theta_{t(k-1)-t(k-2),k-1} < \frac{C}{\gamma^\theta P}$, so $\theta \notin \Theta_k$ as well; firthermore, there exists $\theta_0 \in \Theta_{k-1}$ such that $\lambda^{\theta_0}_{0,k} = \lambda^{\theta_0}_{t(k-1)-t(k-2),k-1} < \frac{C}{\gamma^\theta P}$. Therefore, if $k > 0$, then $\Theta_k \subsetneq \Theta_{k-1}$.

Define $q^g_k \in \mathbb{R}^\Theta_+ \setminus \{0\}$ for $g = f, m$ by $q^{\theta,g}_k = \gamma^\theta p^{\theta,g}$ if $\theta \in \Theta_k$, and $q^{\theta,g}_k = 0$ otherwise. Then $q^{\theta,m}_k + q^{\theta,f}_k \leq 1$ for all θ. Consider the sequences $(\lambda^{\theta,g}_{\tau,k})_{\tau \geq 0}$ for $g = f, m$ and $(\lambda^\theta_{\tau,k})_{\tau \geq 0}$ defined by Eqs. (20)–(20) for the vectors q^f_k, q^m_k .

Suppose first that there are $\bar{\tau} > 0$ and $\theta_0 \in \Theta_k$ such that $\lambda^{\theta_0}_{\bar{\tau},k} < \frac{C}{\gamma^{\theta_0} P}$ and $\lambda^\theta_{\tau,k} \geq \frac{C}{\gamma^\theta P}$ for all $\theta \in \Theta_k$ and $0 \leq \tau < \bar{\tau}$. Let $t(k) = t(k-1) + \bar{\tau}$. Then, for each group $g = f, m$, the dynamics in Eqs. (20)–(20) induced by the vectors q^f_k, q^m_k for the subsequence $(\lambda^g_{\tau,k})_{\tau=0,\ldots,\bar{\tau}}$ coincide with those in Eq. (A.30) for the subsequences $(\lambda^g_t)_{t=t(k-1),\ldots,t(k)}$; thus, (i) holds for $j = k$. Furthermore, (ii) and the second part of (iii) hold for $j = k$ by the definition of $\bar{\tau}$. For the first part of (iii) with $j = k$, recall that by definition $q^{\theta,m}_k + q^{\theta,f}_k = 0$ for $\theta \in \Theta \setminus \Theta_k$; hence, for all $\theta' \in \Theta$ and all $\theta \in \Theta \setminus \Theta_k$, $q^{\theta,m}_k + q^{\theta,f}_k \leq q^{\theta'}_{m,k} + q^{\theta'}_{f,k}$, which by part 3(a) in Theorem 1 implies that $\lambda^\theta_{\tau+1,k}/\lambda^\theta_{\tau,k} \leq \lambda^{\theta'}_{\tau+1,k}/\lambda^{\theta'}_{\tau,k}$. Therefore, it must be the case that $\lambda^\theta_{\tau+1,k}/\lambda^\theta_{\tau,k} \leq 1$ for all $\theta \in \Theta \setminus \Theta_k$: otherwise, $\sum_{\theta' \in \Theta} \lambda^{\theta'}_{\tau+1,k} > \sum_{\theta' \in \Theta} \lambda^{\theta'}_{\tau,k} = 1$, which contradicts the fact that $\lambda_{\tau+1,k} \in \Delta(\Theta)$ per Theorem 1. Since by definition $\lambda^\theta_{0,k} < \frac{C}{\gamma^\theta P}$ for $\theta \notin \Theta_k$, it follows that also $\lambda^\theta_{\tau,k} < \frac{C}{\gamma^\theta P}$ for $\tau = 0, \ldots, \bar{\tau}$ and for any such θ. Thus, in this case Conditions $C(k)$ hold.

If instead $\lambda^\theta_{\tau,k} \geq \frac{C}{\gamma^\theta P}$ for all $\theta \in \Theta_k$ and $\tau \geq 0$, then for each group $g = f, m$, the dynamics in Eqs. (20)–(20) induced by the vectors q^m_k, q^f_k for the subsequence $(\lambda^g_{\tau,k})_{\tau \geq 0}$ coincide with those in Eq. (A.30) for the subsequence $(\lambda^g_t)_{t \geq t(k-1)}$. Again, in this case (i) holds for $j = k$. This completes the proof of the claim.

Since the set Θ is finite, there exists $K \geq 0$ such that the induction stops—that is, $\lambda^\theta_{\tau,K} \geq \frac{C}{\gamma^\theta P}$ for all $\theta \in \Theta_K$ and $\tau \geq 0$. For $k = 0, \ldots, K$, let $\Theta^{\max}_k = \arg\max\{q^{\theta,m}_k + q^{\theta,f}_k : \theta \in \Theta\}$. Since by definition $q^{\theta,g}_k = \gamma^\theta p^{\theta,g} > 0$ for $\theta \in \Theta_k$ and $q^{\theta,g}_k = 0$ for $\theta \in \Theta \setminus \Theta_k$, $\Theta^{\max}_k =$

$\arg\max\{\gamma^\theta(p^{\theta,m}+p^{\theta,f}):\theta\in\Theta_k\}$. In particular, by the definition of Θ_0, $\Theta_0^{\max}=\Theta^{\max}$, where $\Theta^{\max}$ is as defined in Eq. (A.31).

For every $k=0,\ldots,K-1$, and every $\theta\in\Theta_k^{\max}$, $\lambda_{\tau+1,k}^\theta/\lambda_{\tau,k}^\theta\geq 1$ for $0\leq\tau<t(k)-t(k)$; otherwise, by part 3(a) in Theorem 1 and the definition of $\Theta_k^{\max}$, $\lambda_{\tau+1,k}^\theta/\lambda_{\tau,k}^\theta<1$ for all such τ and all θ, so $\sum_{\theta\in\Theta}\lambda_{\tau+1,k}^\theta<\sum_{\theta\in\Theta}\lambda_{\tau,k}^\theta=1$, which contradicts the fact that $\lambda_{\tau+1}\in\Delta(\Theta)$ per Theorem 1.

It follows that, for every $\theta\in\Theta_0^{\max}$,

$$\frac{C}{\gamma^\theta P}\leq\lambda_{0,0}^\theta\leq\lambda_{t(1)-t(0),0}^\theta=\lambda_{0,1}^\theta\leq\lambda_{t(2)-t(1),1}^\theta\cdots\leq\lambda_{0,K}^\theta,$$

so $\theta\in\Theta_k$ for all $k=0,\ldots,K$. Therefore, since $\Theta_0\supsetneq\Theta_1\supsetneq\ldots\supsetneq\Theta_K$ and $\Theta_k^{\max}=\arg\max\{\gamma^\theta(p^{\theta,m}+p^{\theta,f}):\theta\in\Theta_k\}$, $\Theta_0^{\max}=\Theta_k^{\max}$ for all $0\leq k\leq K$.[1] Hence $\Theta^{\max}=\Theta_K^{\max}$.

In addition, again by part 3(a) of Theorem 1, if $\theta,\theta'\in\Theta^{\max}$, then $\frac{\lambda_{\tau+1,k}^\theta}{\lambda_{\tau,k}^\theta}=\frac{\lambda_{\tau+1,k}^{\theta'}}{\lambda_{\tau,k}^{\theta'}}$ for all $k=0,\ldots,K-1$ and $\tau=0,\ldots,t(k)-t(k-1)$, and for $k=K$ and all $\tau\geq 0$. Rearranging terms, $\frac{\lambda_{\tau+1,k}^\theta}{\lambda_{\tau+1,k}^{\theta'}}=\frac{\lambda_{\tau,k}^\theta}{\lambda_{\tau,k}^{\theta'}}$ for such k and τ. Therefore, (i) in Conditions $C(0)$...$C(K)$ imply that

$$\frac{\lambda_{0,K}^\theta}{\lambda_{0,K}^{\theta'}}=\frac{\lambda_{t(K-1)}^\theta}{\lambda_{t(K-1)}^{\theta'}}=\frac{\lambda_{t(K-1)-t(K-2),K-1}^\theta}{\lambda_{t(K-1)-t(K-2),K-1}^{\theta'}}=\frac{\lambda_{0,K-1}^\theta}{\lambda_{0,K-1}^{\theta'}}=\ldots=\frac{\lambda_{t(0)-t(-1),0}^\theta}{\lambda_{t(0)-t(-1),0}^{\theta'}}=\frac{\lambda_{0,0}^\theta}{\lambda_{0,0}^{\theta'}}=\frac{\lambda_0^\theta}{\lambda_0^{\theta'}}.$$

Therefore, for $\theta\in\Theta^{\max}=\Theta_K^{\max}$, from Theorem 1 part (4),

$$\bar\lambda^\theta=\bar\lambda_K^\theta=\frac{\lambda_{0,K}^\theta}{\sum_{\theta'\in\Theta^{\max}}\lambda_{0,K}^{\theta'}}=\frac{1}{\sum_{\theta'\in\Theta^{\max}}\frac{\lambda_{0,K}^{\theta'}}{\lambda_{0,K}^\theta}}=\frac{1}{\sum_{\theta'\in\Theta^{\max}}\frac{\lambda_0^{\theta'}}{\lambda_0^\theta}}=\frac{\lambda_0^\theta}{\sum_{\theta'\in\Theta^{\max}}\lambda_0^{\theta'}}$$

which is Eq. (A.32). Similarly, for $\theta\in\Theta^{\max}$, letting $q_K^\theta=q_K^{\theta,m}+q_K^{\theta,g}$, part (5) in the same Theorem implies that

$$\bar\lambda^{\theta,m}=\bar\lambda_K^{\theta,m}=\frac{\lambda_{0,K}^\theta q_K^{\theta,m}}{\sum_{\theta'\in\Theta^{\max}}\lambda_{0,K}^{\theta'}q_K^{\theta'}}=\frac{q_K^{\theta,m}}{\sum_{\theta'\in\Theta^{\max}}\frac{\lambda_{0,K}^{\theta'}}{\lambda_{0,K}^\theta}q_K^{\theta'}}=\frac{q_K^{\theta,m}}{\sum_{\theta'\in\Theta^{\max}}\frac{\lambda_0^{\theta'}}{\lambda_0^\theta}q_K^{\theta'}}=\frac{\lambda_0^\theta q_K^{\theta,m}}{\sum_{\theta'\in\Theta^{\max}}\lambda_0^{\theta'}q_K^{\theta'}},$$

and analogously for $\bar\lambda^{\theta,f}$. Since, for $\theta\in\Theta_k^{\max}$, $q_K^{\theta,g}=\gamma^\theta p^{\theta,g}$, this yields Eq. (A.33). *Q.E.D.*

Proof of Corollary A.1: since $\lambda_0=p^m$, $\lambda_0^\theta=p^{\theta,m}>p^{\theta,f}=p^{\sigma(\theta),m}=\lambda_0^{\sigma(\theta),m}$; thus, for C sufficiently small, $\sigma(\theta)\notin\Theta^{\max}$. In this case, from Eq. (A.33), $\bar\lambda^{\theta,m}/\bar\lambda^{\theta,f}=p^{\theta,m}/p^{\theta,f}>1$, so $\bar\lambda^{\theta,m}>\bar\lambda^{\theta,f}$. Furthermore, if $\Theta^{\max}=\{\theta\}$, then Eq. (A.33) reduces to $\bar\lambda^{\theta,g}=\frac{p^{\theta,g}}{p^{\theta,m}+p^{\theta,g}}$, which implies that the first displayed equation in the Corollary holds.

[1] Pick $\theta_k\in\Theta_k^{\max}$ arbitrarily. For every k, we just showed that $\theta_0\in\Theta_k$; furthermore, $\Theta_k\subsetneq\Theta_0$. Since $\theta_0,\theta_k\in\Theta_0$ and $\theta_0\in\Theta_0^{\max}$, $\gamma^{\theta_0}(p^{\theta_0,m}+p^{\theta_0,f})\geq\gamma^{\theta_k}(p^{\theta_k,m}+p^{\theta_k,f})$. Since $\theta_0,\theta_k\in\Theta_k$ and $\theta_k\in\Theta_k^{\max}$, $\gamma^{\theta_k}(p^{\theta_k,m}+p^{\theta_k,f})\geq\gamma^{\theta_0}(p^{\theta_0,m}+p^{\theta_0,f})$. Therefore $\theta_0\in\Theta_k$ and $\theta_k\in\Theta_0$, which implies the claim.

To prove the last inequality, let $D = \lambda_0^{\theta}\gamma p^{\theta,m} + \lambda_0^{\theta}\gamma p^{\theta,f} + \lambda_0^{\theta'}\gamma p^{\theta',m} + \lambda_0^{\theta'}\gamma p^{\theta',f}$ where $\gamma = \gamma^{\theta} = \gamma^{\theta'}$. Then, from Proposition 3 part (ii),

$$\begin{aligned}\left(\bar{\lambda}^{\theta,m}_{C=0} + \bar{\lambda}^{\theta',m}_{C=0}\right) - \left(\bar{\lambda}^{\theta,f}_{C=0} + \bar{\lambda}^{\theta',f}_{C=0}\right) &= \left(\frac{\lambda_0^{\theta}\gamma p^{\theta,m}}{D} + \frac{\lambda_0^{\theta'}\gamma p^{\theta',m}}{D}\right) - \left(\frac{\lambda_0^{\theta}\gamma p^{\theta,f}}{D} + \frac{\lambda_0^{\theta'}\gamma p^{\theta',f}}{D}\right) \\ &= \left(\frac{\lambda_0^{\theta}\gamma\left(p^{\theta,m} - p^{\theta,f}\right)}{D}\right) - \left(\frac{\lambda_0^{\theta'}\gamma\left(p^{\theta',f} - p^{\theta',m}\right)}{D}\right) \\ &= \left(p^{\theta,m} - p^{\theta,f}\right)\left[\frac{\left(\lambda_0^{\theta} - \lambda_0^{\theta'}\right)\gamma}{\lambda_0^{\theta}\gamma p^{\theta,m} + \lambda_0^{\theta}\gamma p^{\theta,f} + \lambda_0^{\theta'}\gamma p^{\theta',m} + \lambda_0^{\theta'}\gamma p^{\theta',f}}\right] \\ &= \frac{\left(p^{\theta,m} - p^{\theta,f}\right)}{\left(p^{\theta,m} + p^{\theta,f}\right)}\left[\frac{\lambda_0^{\theta} - \lambda_0^{\theta'}}{\lambda_0^{\theta} + \lambda_0^{\theta'}}\right]:\end{aligned}$$

the third equality follows by noting that, in the second term in parentheses, $p^{\theta',f} - p^{\theta',m} = p^{\sigma(\theta),f} - p^{\sigma(\theta),m} = p^{\theta,m} - p^{\theta,f}$; and the fourth equality follows similarly, by noting that the two terms of the denominator D multiplied by $\lambda_0^{\theta'}$ are $p^{\theta',m} = p^{\theta,f}$ and $p^{\theta',f} = p^{\theta,m}$. Finally, the first term in the last line equals $\bar{\lambda}^{\theta,m} - \bar{\lambda}^{\theta,f}$ when $\Theta^{\max} = \{\theta\}$, and the second term is between 0 and 1 because we assume that $\lambda_0 = p^m$ and $p^{\theta,m} > p^{\theta,f} = p^{\theta',m}$. This yields the required conclusion. *Q.E.D.*

Proof of Corollary A.2: Assume that $\lambda_0 = p^m$. In (a.1), the assumption implies that $\Theta^{\max} = \{\theta^m, \theta^f\} \subseteq \Theta_0$. Substituting $\lambda_0^{\theta^m} = \phi^N$ and $\lambda_0^{\theta^f} = (1-\phi)^N$ in Eq. (A.32) yields $\bar{\lambda}^{\theta^m} = \frac{\phi^N}{\phi^N + (1-\phi)^N}$. Similarly, substituting for γ^{θ}, p^m and p^f in Eq. (A.33) yields the same expression for $\bar{\lambda}^{\theta^m,g}$ and $\bar{\lambda}^{\theta^f,g}$, $g = m, f$, as in Corollary 6. This yields the required expression for $\bar{\Lambda}^m$.

For (a.2), the assumption implies that $\Theta^{\max} = \{\theta^m\}$. This immediately implies that $\bar{\lambda}^{\theta^m} = 1$. Furthermore, from Eq. (A.33), $\bar{\Lambda}^m = \bar{\lambda}^{\theta^m,m} = \frac{\gamma^{\theta^m} p^{\theta^m,m}}{\gamma^{\theta^m}(p^{\theta^m,m} + p^{\theta^m,f})} = \frac{p^{\theta^m,m}}{p^{\theta^m,m} + p^{\theta^m,f}} = \frac{\phi^N}{\phi^N + (1-\phi)^N}$, as asserted. Finally, we compare this quantity with its counterpart in Eq. (16):

$$\frac{1 + \left(\frac{\phi}{1-\phi}\right)^{2N}}{1 + \left(\frac{\phi}{1-\phi}\right)^{2N} + 2\left(\frac{\phi}{1-\phi}\right)^{N}} = \frac{(1-\phi)^{2N} + \phi^{2N}}{\left[(1-\phi)^N + \phi^N\right]^2} <$$
$$< \frac{(1-\phi)^N\phi^N + \phi^{2N}}{\left[(1-\phi)^N + \phi^N\right]^2} = \frac{(1-\phi)^N + \phi^N}{(1-\phi)^N + \phi^N} \cdot \frac{\phi^N}{(1-\phi)^N + \phi^N} = \frac{\phi^N}{(1-\phi)^N + \phi^N} = \bar{\Lambda}^m,$$

where the inequality follows from the assumption that $\phi > 0.5$.

The analysis of (b) is analogous to that of (a.2), with θ^* in lieu of θ^m; in this case, $p^{\theta^*,m} = p^{\theta^*,f} = \phi^{N/2}(1-\phi)^{N/2}$, so $\bar{\Lambda}^m = \bar{\lambda}^{\theta^*,m} = \frac{1}{2}$.

The statements about t^{θ} for $\theta \notin \Theta^{\max}$ follow from the construction of $t(0), \ldots, t(K)$ in the proof of Proposition A.1. *Q.E.D.*

Proof of Proposition A.2. For part 1, the key step is analogous to the proof of Proposition 5, modified to allow for endogenous entry. Let $m_0 = \sum_{n=1}^{N/2} \theta$ and $m_1 = \sum_{n=N/2+1}^{N} \theta_n$. By assumption, $m_0 > m_1$. By definition, $p^{\theta,m} = \phi^{m_0}(1-\phi)^{N/2-m_0}\phi^{N/2-m_1}(1-\phi)^{m_1} = \phi^{(m_0-m_1)+N/2}(1-\phi)^{N/2-(m_0-m_1)} = [\phi(1-\phi)]^{N/2}\left(\frac{\phi}{1-\phi}\right)^{m_0-m_1}$, and similarly $p^{\theta^{sym},m} = [\phi(1-\phi)]^{N/2}\left(\frac{1-\phi}{\phi}\right)^{m_0-m_1}$; since $\phi > \frac{1}{2}$, $p^{\theta,m} > p^{\theta^{\text{sym}},m}$. At time 0 we thus have $\lambda_0^{\theta} = p^{\theta,m} > p^{\theta^{\text{sym}},m} = \lambda_0^{\theta^{\text{sym}}}$. Moreover, since p_f is defined with the roles of ϕ and $1-\phi$ reversed, $p^{\theta,f} = p^{\theta^{\text{sym}},m} < p^{\theta,m} = p^{\theta^{\text{sym}},f}$.

Since $\gamma^{\theta^{\text{sym}}} = \gamma^{\theta}$, it follows that at time 0, if $\lambda_0^{\theta^{\text{sym}}} > \frac{C}{\gamma^{\theta^{\text{sym}}}P}$, then also $\lambda_0^{\theta} > \frac{C}{\gamma^{\theta}P}$. In addition, $p_m^{\theta}+p_f^{\theta} = p_m^{\theta^{\text{sym}}}+p_f^{\theta^{\text{sym}}}$. Thus, in the notation of Corollary A.2, for $t < \min(t^{\theta}, t^{\theta^{\text{sym}}})$, both θ and θ^{sym} apply, and applying part 3(a) of Theorem 1 to the relevant subsequence of $(\lambda_t)_{t\geq 0}$ as in the proof of Proposition A.1, $\frac{\lambda_t^{\theta}}{\lambda_{t-1}^{\theta}} = \frac{\lambda_t^{\theta^{\text{sym}}}}{\lambda_{t-1}^{\theta^{\text{sym}}}}$, and hence $\frac{\lambda_t^{\theta}}{\lambda_t^{\theta^{\text{sym}}}} = \frac{\lambda_{t-1}^{\theta}}{\lambda_{t-1}^{\theta^{\text{sym}}}} = \frac{\lambda_0^{\theta}}{\lambda_0^{\theta^{\text{sym}}}} > 1$. Thus, $\lambda_t^{\theta} > \lambda_t^{\theta^{\text{sym}}}$, so again, if $\lambda_t^{\theta^{\text{sym}}} > \frac{C}{\gamma^{\theta^{\text{sym}}}P}$, then also $\lambda_t^{\theta} > \frac{C}{\gamma^{\theta}P}$, i.e., $t^{\theta} \geq t^{\theta^{\text{sym}}}$. In particular, if the inequality is strict and $t^{\theta^{\text{sym}}} < t < t^{\theta}$, then researchers of type θ will apply at time t, but those of type θ^{sym} will not.

For part 2, we have

$$
\begin{aligned}
A_t^m - A_t^f &= \sum_{\theta:\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P} p^{\theta,m} - \sum_{\theta:\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P} p^{\theta,f} = \sum_{\theta} p^{\theta,m}1_{\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P} - \sum_{\theta} p^{\theta,f}1_{\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P} = \\
&= \sum_{\theta} p^{\theta,m}1_{\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P} - \sum_{\theta} p^{\theta^{\text{sym}},f}1_{\lambda_t^{\theta^{\text{sym}}}\geq\frac{C}{\gamma^{\theta^{\text{sym}}}}P} = \sum_{\theta} p^{\theta,m}\left(1_{\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P} - 1_{\lambda_t^{\theta^{\text{sym}}}\geq\frac{C}{\gamma^{\theta^{\text{sym}}}}P}\right) = \\
&= \sum_{\theta:\sum_{n=1}^{N/2}\theta_n>\sum_{n=N/2+1}^{N}\theta_n} p^{\theta,m}\left(1_{\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P} - 1_{\lambda_t^{\theta^{\text{sym}}}\geq\frac{C}{\gamma^{\theta^{\text{sym}}}}P}\right) + \\
&+ \sum_{\theta:\sum_{n=1}^{N/2}\theta_n=\sum_{n=N/2+1}^{N}\theta_n} p^{\theta,m}\left(1_{\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P} - 1_{\lambda_t^{\theta^{\text{sym}}}\geq\frac{C}{\gamma^{\theta^{\text{sym}}}}P}\right) + \\
&+ \sum_{\theta:\sum_{n=1}^{N/2}\theta_n<\sum_{n=N/2+1}^{N}\theta_n} p^{\theta,m}\left(1_{\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P} - 1_{\lambda_t^{\theta^{\text{sym}}}\geq\frac{C}{\gamma^{\theta^{\text{sym}}}}P}\right) = \\
&= \sum_{\theta:\sum_{n=1}^{N/2}\theta_n>\sum_{n=N/2+1}^{N}\theta_n} p^{\theta,m}\left(1_{\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P} - 1_{\lambda_t^{\theta^{\text{sym}}}\geq\frac{C}{\gamma^{\theta^{\text{sym}}}}P}\right) + \\
&+ \sum_{\theta:\sum_{n=1}^{N/2}\theta_n>\sum_{n=N/2+1}^{N}\theta_n} p^{\theta^{\text{sym}},m}\left(1_{\lambda_t^{\theta^{\text{sym}}}\geq\frac{C}{\gamma^{\theta^{\text{sym}}}}P} - 1_{\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P}\right) = \\
&= \sum_{\theta:\sum_{n=1}^{N/2}\theta_n>\sum_{n=N/2+1}^{N}\theta_n} (p^{\theta-p_m^{\theta^{\text{sym}}},m})\left(1_{\lambda_t^{\theta}\geq\frac{C}{\gamma^{\theta}}P} - 1_{\lambda_t^{\theta^{\text{sym}}}\geq\frac{C}{\gamma^{\theta^{\text{sym}}}}P}\right) \geq 0.
\end{aligned}
$$

The third equality follows from the fact that $\theta \mapsto (1-\theta_n)_{n=1}^{N}$ is a bijection. The fourth follows from the fact that $p^{\theta^{\text{sym}},f} = p^{\theta,f}$. To obtain the fifth, we break up the sum into types θ with more (resp. as many, resp. fewer) characteristics between 1 and $N/2$ than between $N/2+1$ and N. For the sixth, observe that if a type θ has the same number of features between 1 and $N/2$ and between $N/2+1$ and N, then $p^{\theta,m} = p^{\theta^{\text{sym}},m}$ and so $\lambda_0^\theta = \lambda_0^{\theta^{\text{sym}}}$; arguing as in Proposition A.2, $\lambda_t^\theta = \lambda_t^{\theta^{\text{sym}}}$ for all $t \geq 0$ (note that as soon as one type stops applying, so does the other); but then, since also $\gamma^\theta = \gamma^{\theta^{\text{sym}}}$, the term in parentheses for such types is identially zero. In addition, we express the sum over θ's for which $\sum_{n=1}^{N/2} \theta_n < \sum_{n=N/2+1}^{N} \theta_n$ iterating over types θ for which $\sum_{n=1}^{N/2} \theta_n > \sum_{n=N/2+1}^{N} \theta_n$, but adding up terms corresponding to the associated symmetric types θ^{sym}. The seventh equality is immediate. Finally, the inequality follows because, for θ such that $\sum_{n=1}^{N/2} \theta_n > \sum_{n=N/2+1}^{N} \theta_n$, the term in parentheses is non-negative by Proposition A.2, and in addition $p^{\theta > p_m^{\theta^{\text{sym}}},m}$. *Q.E.D.*

Proof of Proposition A.3 Let θ^a and θ^b be the types of the two young researchers. We assume that the type of the joint project is the elementwise maximum of θ^a and θ^b: that is, the project displays characteristics i if and only if at least one of the researchers displays it.

For $g = m, f$, let $\Theta^g = \{(\theta, \theta') : \theta \vee \theta' = \theta^g\}$, where $\vee$ denotes the component-wise maximum. Note that, if $(\theta, \theta') \in \Theta^m$, then $\theta_i = \theta'_i = 0$ for $i = N/2+1, \ldots, N$; similarly, if $(\theta, \theta') \in \Theta^f$, then $\theta_i = \theta'_i = 0$ for $i = 1, \ldots, N/2$. Moreover, $(\theta, \theta') \in \Theta^g$ iff $(\theta', \theta) \in \Theta^g$ for $g = m, f$. Finally, $(\theta, \theta') \in \Theta^m$ if and only if $(\bar\theta, \bar\theta') \in \Theta^f$, where $\bar\theta, \bar\theta'$ are defined by $\bar\theta_{i+N/2} = \theta_i$, $\bar\theta'_{i+N/2} = \theta'_i$ and $\bar\theta_i = \bar\theta'_i = 0$ for $i = 1, \ldots, N/2$; furthermore, these types satisfy

$$p^{\theta,m} = p^{\bar\theta,f} \qquad \text{and} \quad p^{\theta',f} = p^{\bar\theta',m}. \tag{A.44}$$

Then, invoking the above properties, the probability that the joint project is accepted—that is, the probability that $\theta^a \vee \theta^b \in \{\theta^m, \theta^f\}$—is

$$\begin{aligned}
&\gamma^{\theta^m}\bar\lambda^{\theta^m} \sum_{(\theta,\theta')\in\Theta^m} p^{\theta,m}\cdot p^{\theta',f} + \gamma^{\theta^f}\bar\lambda^{\theta^f} \sum_{(\theta,\theta')\in\Theta^f} p^{\theta,m}\cdot p^{\theta',f} \\
=&\gamma^{\theta^m}\bar\lambda^{\theta^m} \sum_{(\theta,\theta')\in\Theta^m} p^{\theta,m}\cdot p^{\theta',f} + \gamma^{\theta^f}\bar\lambda^{\theta^f} \sum_{(\theta,\theta')\in\Theta^m} p^{\bar\theta,m}\cdot p^{\bar\theta',f} \\
=&\gamma^{\theta^m}\bar\lambda^{\theta^m} \sum_{(\theta,\theta')\in\Theta^m} p^{\theta,m}\cdot p^{\theta',f} + \gamma^{\theta^f}\bar\lambda^{\theta^f} \sum_{(\theta',\theta)\in\Theta^m} p^{\bar\theta',m}\cdot p^{\bar\theta,f} \\
=&\gamma^{\theta^m}\bar\lambda^{\theta^m} \sum_{(\theta,\theta')\in\Theta^m} p^{\theta,m}\cdot p^{\theta',f} + \gamma^{\theta^f}\bar\lambda^{\theta^f} \sum_{(\theta',\theta)\in\Theta^m} p^{\theta',f}\cdot p^{\theta,m} \\
=&\gamma^{\theta^m}\bar\lambda^{\theta^m} \sum_{(\theta,\theta')\in\Theta^m} p^{\theta,m}\cdot p^{\theta',f} + \gamma^{\theta^f}\bar\lambda^{\theta^f} \sum_{(\theta,\theta')\in\Theta^m} p^{\theta,f}\cdot p^{\theta',m} \\
=&(\gamma^{\theta^m}\bar\lambda^{\theta^m} + \gamma^{\theta^f}\bar\lambda^{\theta^f}) \sum_{(\theta,\theta')\in\Theta^m} p^{\theta,m}\cdot p^{\theta',f} \\
=&\gamma_0\rho^{N/2} \sum_{(\theta,\theta')\in\Theta^m} p^{\theta,m} p^{\theta',f} \equiv \gamma_0\rho^{N/2}\Pi,
\end{aligned}$$

where the last equality follows from the definition of γ^θ and the fact that θ^m, θ^f are the only surviving types.

Now let $L(\theta) = \sum_i \theta_i$. We claim that the expectation of $L(\theta^a) - L(\theta^b)$ conditional on $\theta^a \vee \theta^b \in \{\theta^m, \theta^f\}$ is strictly positive—that is, the expected quality of a, the young M coauthor, is strictly higher than the expected quality of that of the young F coauthor b.

First,

$$
\begin{aligned}
\Delta \equiv & \sum_{(\theta,\theta')\in\Theta^m} p^{\theta,m}\cdot p^{\theta',f}[L(\theta)-L(\theta')] \\
= & \sum_{(\theta,\theta')\in\Theta^m:L(\theta)>L(\theta')} p^{\theta,m}\cdot p^{\theta',f}[L(\theta)-L(\theta')] + \sum_{(\theta,\theta')\in\Theta^m:L(\theta)<L(\theta')} p^{\theta,m}\cdot p^{\theta',f}[L(\theta)-L(\theta')] \\
= & \sum_{(\theta,\theta')\in\Theta^m:L(\theta)>L(\theta')} [p^{\theta,m}\cdot p^{\theta',f} - p^{\theta',m}\cdot p^{\theta,f}][L(\theta)-L(\theta')] > 0.
\end{aligned}
$$

The last equality follows because $(\theta, \theta') \in \Theta^m$ if and only if $(\theta', \theta) \in \Theta^m$, and of course $L(\theta) > L(\theta')$ iff $L(\theta') < L(\theta)$. The inequality follows because, if $L(\theta) > L(\theta')$, then by assumption $p^{\theta,m} > p^{\theta',m}$ and $p^{\theta',f} > p^{\theta,f}$.

Repeating the calculations for Θ^f and again appealing to the properties of pairs $(\theta, \theta') \in \Theta^m$ and the corresponding types $(\bar\theta, \bar\theta') \in \Theta^f$,

$$
\begin{aligned}
& \sum_{(\theta,\theta')\in\Theta^f} p^{\theta,m}\cdot p^{\theta',f}[L(\theta)-L(\theta')] = \sum_{(\theta,\theta')\in\Theta^f:L(\theta)>L(\theta')} [p^{\theta,m}\cdot p^{\theta',f} - p^{\theta',m}\cdot p^{\theta,f}][L(\theta)-L(\theta')] \\
= & \sum_{(\theta,\theta')\in\Theta^m:L(\theta)>L(\theta')} [p^{\bar\theta,m}\cdot p^{\bar\theta',f} - p^{\bar\theta',m}\cdot p^{\bar\theta,f}][L(\bar\theta)-L(\bar\theta')] \\
= & \sum_{(\theta,\theta')\in\Theta^m:L(\theta)>L(\theta')} [p^{\theta,f}\cdot p^{\theta',m} - p^{\theta',f}\cdot p^{\theta,m}][L(\theta)-L(\theta')] = \\
= & -\sum_{(\theta,\theta')\in\Theta^m} p^{\theta,m}\cdot p^{\theta',f}[L(\theta)-L(\theta')] = -\Delta.
\end{aligned}
$$

Finally, the expected difference in the number of characteristics of θ^a and θ^b is

$$
\mathrm{E}[L(\theta^a) - L(\theta^b) | \theta^a \vee \theta^b \in \{\theta^m, \theta^f\}] = \frac{\gamma^{\theta^m}\bar\lambda^{\theta^m}\Delta - \gamma^{\theta^f}\bar\lambda^{\theta^f}\Delta}{\gamma_0\rho^{N/2}\Pi} = \frac{\rho^{N/2}\Delta}{\Pi}(\bar\lambda^{\theta^m} - \bar\lambda^{\theta^f}) > 0,
$$

as asserted.

Q.E.D