Self-Image Bias and Talent Loss

On-Line Appendix

Marciano Siniscalchi

Northwestern University

Pietro Veronesi

University of Chicago

This on-line appendix contains additional analysis and the proofs of our propositions.

A1. Additional Analysis and Results

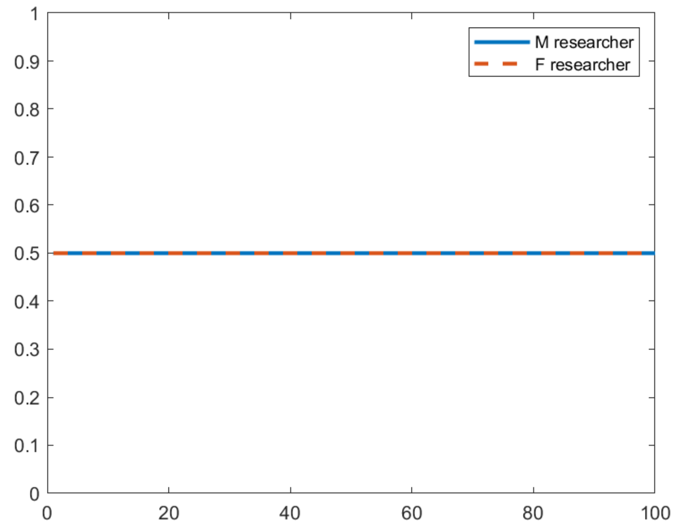
A1.1. Balanced Steady State

In Section 3. we considered a simple numerical example with only two characteristics ($N = 2$), which led to types $\Theta = \{(0,0),(0,1),(1,0),(1,1)\}$. In that section, we showed that when $\rho < \overline{\rho}(\phi, N)$ and the initial population of referees is only from the M-group, $\lambda_0^{\theta,m} = p^{\theta,m}$, then the dynamics never converges. Here we now consider a different initial condition.

Indeed, the dynamics of the mass of each type depends upon their frequencies in the population of young researchers, p_m and p_f, as well as the initial conditions λ_0. In particular, suppose that the initial mass of referees is composed of M- and F-researchers in equal proportions: $\lambda_0 = \frac{1}{2}p_m + \frac{1}{2}p_f$. One implication is that then the two M-prevalent and F-prevalent types $\theta^m = (1,0)$ and $\theta^f = (0,1)$ both represent 34% of the initial mass of referees, whereas the other two types $(0,0)$ and $(1,1)$ each represent 16% of the initial population. While we can no longer invoke the results in Sections 2.2.-2.5., we can plot the dynamics of the fractions of established M- and F-researchers, as well as those of established M-and F-researcher types. (Theorem A.1 in the Appendix characterizes the limiting behavior of the system for arbitrary initial conditions and type distributions.)

Figures A.1 and A.2 display the results. The figures are self explanatory: an equal proportion of M- and F-researchers is maintained throughout. However, importantly, type θ^f (resp. θ^m) will eventually become prevalent among F-researchers (resp. M-researchers), which means that established F- (resp. M-) economists are oversampled from those who

Figure A.1: Fraction of M and F researchers with Start from Equal Proportions

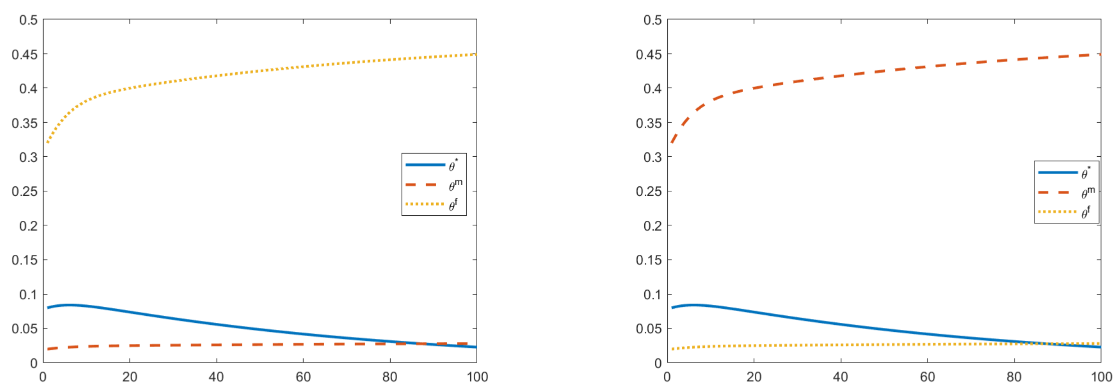


Fraction of M and F researchers when $\lambda_0 = \frac{1}{2}p_m + \frac{1}{2}p_f$. Parameters: $\phi = 0.8$, $\gamma_0 = 0.2$, $\rho = 4$, $N = 2$.

Figure A.2: Types of Established F and M Researchers with Start from Equal Proportions

(a) F researchers

(b) M researchers



Types of established F (left) and M (right) researchers. We show types $\theta^* = (1, 1, ..., 1)$, $\theta^m = (1, ..., 1, 0, ..., 0)$, and $\theta^f = (0, ..., 0, 1, ..., 1)$. Initially $\lambda_0 = \frac{1}{2}p_m + \frac{1}{2}p_f$. Parameters: $\phi = 0.8$, $\gamma_0 = 0.2$, $\rho = 4$, $N = 2$.

possess characteristic 2 (resp. 1). Furthermore, the efficient type θ^* will disappear in the limit.

A1.2. Seniors and Juniors

In Section 6. we extended the basic model to include different levels of seniorities in the established set of researchers, with seniors evaluating juniors before accepting them into their group, and both seniors and juniors evaluating the young researchers. The analysis is substantially more complex in this case, and we only rely on numerical simulations. The following cases add up to the one discussed in the body of the paper.All the simulations in this section assume equal fractions of juniors and seniors ($\sigma = 0.5$).

First, the presence of a second screening—and hence a second opportunity for self-image bias to exert its influence—can exacerbate group imbalance in the senior rank, at least in the short run. Figure A.3 demonstrates this. Model parameters are as in Figure 2, so in a single-cohort environment significant group imbalance emerges. The same is true with two ranks; however, in the short run, the imbalance is more pronounced in the senior rank. The reason is that, in order to be promoted to the senior rank, a researcher must match with a referee of the same type *twice*. Initially, both junior and senior referees have the same type distribution, which by assumption coincides with that of M researchers. Hence, whatever effect is present at the junior rank is compounded at the senior rank.[1] The difference between the two ranks vanishes in the long run because, as type θ^m becomes prevalent among established juniors and seniors, promotion eventually is driven solely by objective research quality—matching with a senior reviewer of the junior candidate's own type is virtually guaranteed.

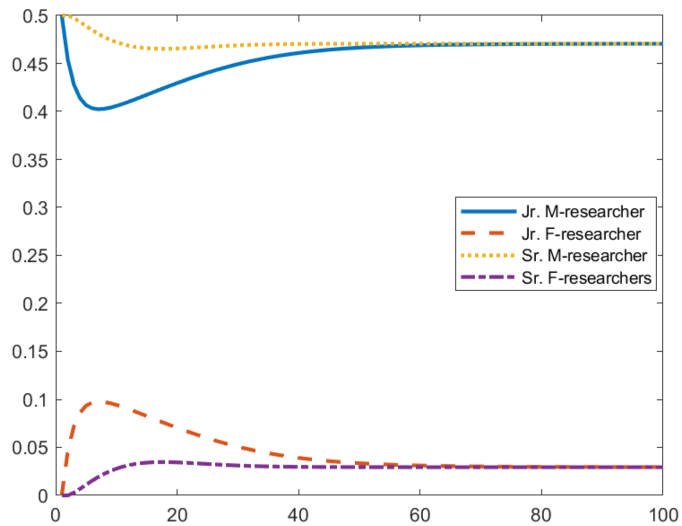
A more pronounced group imbalance can also arise, in the short / medium run, for parameter values for which convergence is eventually attained. This is demonstrated in Figure A.4, where we take $\phi = 0.6$ rather than $\phi = 0.8$. Again, the need to match with a like type twice, coupled with the assumption that the initial population consists entirely of M-researchers, leads to a lower representation of F researchers at the senior rank. However, over time, type θ^* prevails among juniors and seniors, so matching with like types is virtually guaranteed; and since convergence is attained amongst juniors, it must obtain among seniors as well.

A1.3. Similarity in Research Characteristics

In this section we extend the model to investigate the case in which referees accept researchers who have characteristics close but not necessarily identical to their own. In particular, we

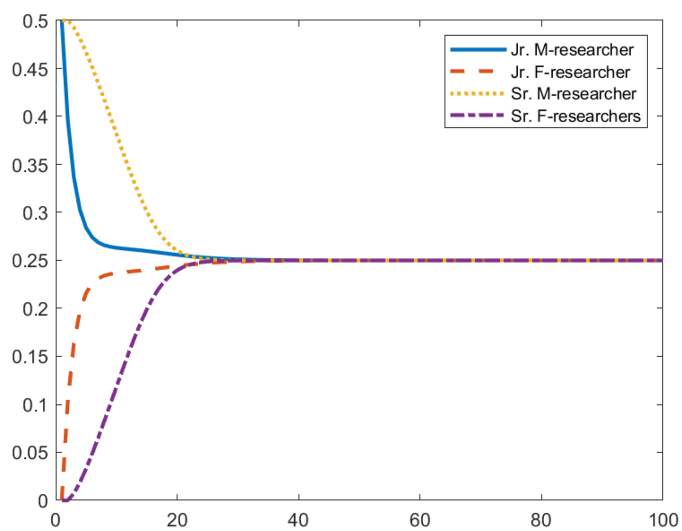
[1] In fact, the bias becomes stronger over time at the senior rank. The reason is that the initial population of junior candidates up for promotion is characterized by types distributed as among male researchers, whereas the initial population of young researchers applying for a junior position is balanced.

Figure A.3: More extreme imbalance for senior rank



Fraction of senior and junior M and F researchers when $\lambda_0 = p_m$. Parameters: $\phi = 0.8$, $\gamma_0 = 0.2$, $\rho = 4$, $N = 2$.

Figure A.4: Convergence, but greater short-run imbalance among seniors



Fraction of senior and junior M and F researchers when $\lambda_0 = p^m$. Parameters: $\phi = 0.6$, $\gamma_0 = 0.2$, $\rho = 4$, $N = 2$.

assume that referee r of type θ^r accepts the research of young researcher θ if

$$D(\theta^r, \theta) = \sum_n (\theta_n^r - \theta_n)^2 \leq \eta \tag{A.28}$$

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 models so far correspond to $\eta = 0$. If instead $\eta > 0$, the dynamics for λ_t^θ are still as in Eq. (7), but the mass $a_t^{\theta,g}$ of accepted researchers of type θ in group $g \in \{f, m\}$ is given by

$$a_t^{\theta,g} \;=\; \gamma^\theta \sum_{\theta^r : D(\theta^r, \theta) \leq \eta} \lambda_{t-1}^{\theta^r} p^{\theta,g} \tag{A.29}$$

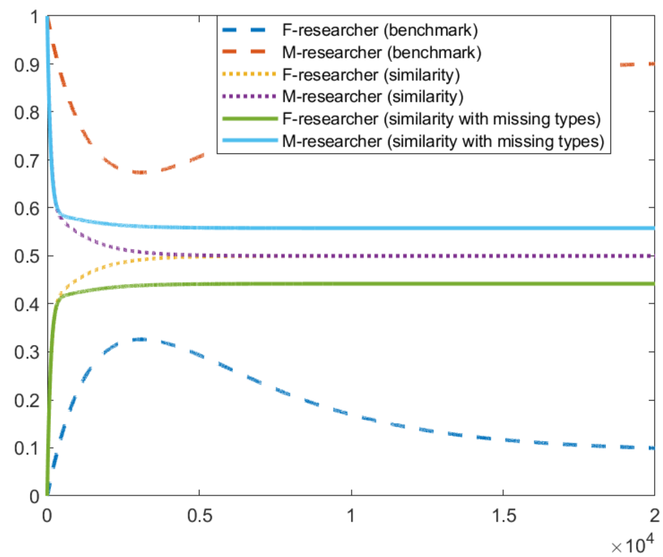
Unfortunately, obtaining general analytical results in this case seems difficult. Therefore, we consider illustrative special cases.

A1.3.1. Connected Set of Types

The set Θ of types we have considered so far enjoys a special structure that is relevant to the relaxed definition of "acceptance" in Eq. (A.28). For every $\eta \geq 1$, and every pair $\theta, \theta' \in \Theta$, there is a finite ordered list $\theta_1, \ldots, \theta_K \in \Theta$ such that $\theta_1 = \theta$, $\theta_K = \theta'$, and $D(\theta_k, \theta_{k+1}) \leq \eta$ for all $k = 1, \ldots, K-1$. 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 Sections 3. and 4., the relaxed acceptance criterion in Eq. (A.28) leads to convergence. For instance, Figure A.5 illustrates the parameterization used in Section 4.. The dashed lines represent the 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.

Figure A.5: Fraction of M and F Researchers under the Research Similarity Assumption



Fraction of M and F researchers when $\lambda_0 = p^m$. Parameters: $\phi = 0.5742$, which implied $d = 0.3$, $\gamma_0 = 0.2$, $\rho = 4$, $N = 10$, and, under research similarity, $\eta = 1$.

A1.3.2. Disconnected Set of Types

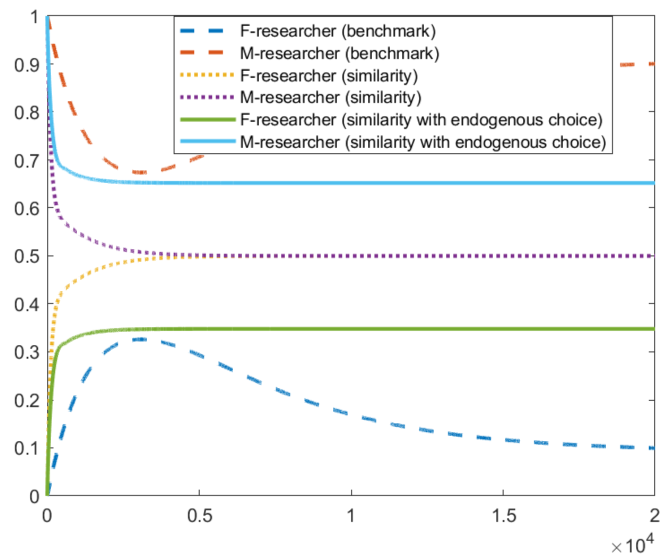
A subset of $\{0,1\}^N$ may well be η-disconnected for some η. For a trivial example, $\{\theta^m, \theta^f\}$ is $(N-1)$-disconnected, because each of the N coordinates of θ^f is different from the corresponding coordinate of θ^f. A fortiori, it is η-disconnected for every $\eta \leq N-1$.

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 Figure A.5 represent the same parameterization as in the previous subsection, with $\eta = 1$, but applied to a state space Θ obtained by randomly removing 15% of the elements of $\{0,1\}^N$ and suitably renormalizing probabilities. As expected, the system does not attain group balance in the limit.

A1.3.3. 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 5.. In this case, even if the connected set of types would lead to convergence (see subsection A1.3.1.), the endogenous entry prevents such convergence, as shown in Section 5.1.1.. 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,

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



Fraction of M and F researchers when $\lambda_0 = p^m$. Parameters: $\phi = 0.5742$, which implied $d = 0.3$, $\gamma_0 = 0.2$, $\rho = 4$, $N = 10$, and, under research similarity, $\eta = 1$.

respectively, the research similarity case ($\eta = 1$). The solid lines now show the the fraction of M- and F-researchers under research simularity ($\eta = 1$) but with endogenous entry. The intuition is the same as the one given in Section 5..

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

A2. 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.1 *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.

Proof of Proposition A.1 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

$$p^{\theta,m} = p^{\bar{\theta},f} \qquad \text{and} \quad p^{\theta',f} = p^{\bar{\theta}',m}. \tag{A.30}$$

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 $(\theta, \theta') \in$

Θ^m and the corresponding types $(\bar{\theta}, \bar{\theta}') \in \Theta^f$,

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

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

A3. Proofs

We first characterize key features of the population dynamics for an arbitrary, finite set Θ of types, with initial distribution $\lambda_0 \in \Delta(\Theta)$, such that $\lambda_0 = \lambda_0^m + \lambda_0^f$ for $\lambda_0^m, \lambda_0^f \in \mathbb{R}_+^\Theta$, and per-period inflows $q^g = (q^{\theta,g})_{\theta\in\Theta} \in \mathbb{R}_+^\Theta \setminus \{0\}$, for $g \in \{f, m\}$. It is also convenient to define $q = q^m + q^f$. Then, for $g \in \{f, m\}$, the dynamics are given by

$$\lambda_t^{\theta,g} = \lambda_{t-1}^{\theta,g} \left(1 - \sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'}\right) + \lambda_{t-1}^{\theta} q^{\theta,g} \tag{A.31}$$
$$\lambda_t^\theta = \lambda_t^{\theta,m} + \lambda_t^{\theta,f}. \tag{A.32}$$

The body of the paper focuses on the special case $q^{\theta,m} = \gamma^\theta p^{\theta,m}$, $q^{\theta,f} = \gamma^\theta p^{\theta,f}$.

Theorem A.1 *Assume that $q^\theta \leq 1$ for all $\theta \in \Theta$. Then, for all $t \geq 0$, $\lambda_t \in \Delta(\Theta)$, and $\lambda_t^m, \lambda_t^f \in \mathbb{R}_+^\Theta$. Moreover:*

1. *if $\lambda_0^\theta = 0$, then $\lambda_t^\theta = 0$ for all $t \geq 0$;*

2. *if $\lambda_0^\theta > 0$, then $\lambda_t^\theta > 0$ for all $t \geq 0$;*

3. *for $\theta, \tilde{\theta} \in \Theta$ with $\lambda_0^\theta \cdot \lambda_0^{\tilde{\theta}} > 0$:*

   (a) $\frac{\lambda_t^\theta}{\lambda_{t-1}^\theta} - \frac{\lambda_t^{\tilde{\theta}}}{\lambda_{t-1}^{\tilde{\theta}}} = q^\theta - q^{\tilde{\theta}}$ *for all $t \geq 1$, and*

   (b) $q^\theta > q^{\tilde{\theta}}$ *implies* $\frac{\lambda_t^\theta}{\lambda_t^{\tilde{\theta}}} \to \infty$, *and* $q^\theta = q^{\tilde{\theta}}$ *implies* $\frac{\lambda_t^\theta}{\lambda_t^{\tilde{\theta}}} = \frac{\bar{\lambda}_o^\theta}{\bar{\lambda}_o^{\tilde{\theta}}}$ *for all $t \geq 0$;*

4. *define the set*

$$\Theta^{\max} = \{\theta \in \Theta \ : \ \lambda_0^\theta > 0, \ \theta \in \arg\max_{\theta' \in \Theta} q^{\theta'}\} \tag{A.33}$$

   and let $\bar{\lambda} \in \Delta(\Theta)$ be such that

$$\bar{\lambda}^{\tilde{\theta}} = \begin{cases} \frac{\lambda_0^{\tilde{\theta}}}{\sum_{\theta \in \Theta^{\max}} \lambda_0^\theta} & \tilde{\theta} \in \Theta^{\max} \\ 0 & \tilde{\theta} \notin \Theta^{\max} \end{cases} : \tag{A.34}$$

   then $\lim_{t\to\infty} \lambda_t = \bar{\lambda}$;

5. *define*

$$\bar{\lambda}^{\tilde{\theta},f} = \begin{cases} \frac{\lambda_0^{\tilde{\theta}} q^{\tilde{\theta},f}}{\sum_{\theta \in \Theta^{\max}} \lambda_0^\theta q^\theta} & \tilde{\theta} \in \Theta^{\max} \\ 0 & \tilde{\theta} \notin \Theta^{\max} \end{cases} \quad \text{and} \quad \bar{\lambda}^{\tilde{\theta},m} = \begin{cases} \frac{\lambda_0^{\tilde{\theta}} q^{\tilde{\theta},m}}{\sum_{\theta \in \Theta^{\max}} \lambda_0^\theta q^\theta} & \tilde{\theta} \in \Theta^{\max} \\ 0 & \tilde{\theta} \notin \Theta^{\max} \end{cases} : \tag{A.35}$$

   then $\lim_{t\to\infty} \lambda_t^f = \bar{\lambda}^f$ and $\lim_{t\to\infty} \lambda_t^m = \bar{\lambda}^m$.

Proof: Eqs. (A.31) and (A.32) imply that

$$\lambda_t^\theta = \left(1 - \sum_{\theta' \in \Theta} \lambda_{t-1}^{\theta'} q^{\theta'}\right) \lambda_{t-1}^\theta + \lambda_{t-1}^\theta q^\theta. \tag{A.36}$$

By assumption $\lambda_0 \in \Delta(\Theta)$. Inductively, suppose $\lambda_{t-1} \in \Delta(\Theta)$ and $\lambda_{t-1}^m, \lambda_{t-1}^f \in \mathbb{R}_+^\Theta$. Summing over Θ on both sides of Eq. (A.36) yields $\sum_\theta \lambda_t^\theta = (1 - \sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'})(\sum_\theta \lambda_{t-1}^\theta) + \sum_\theta \lambda_{t-1}^\theta q^\theta = (1 - \sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'}) + \sum_\theta \lambda_{t-1}^\theta q^\theta = 1$. Furthermore, since $\lambda_{t-1} \in \Delta(\Theta)$, $\sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'} \in [\min_{\theta'} q^{\theta'}, \max_{\theta'} q^{\theta'}] \subseteq [0, 1]$; moreover, $q^\theta \geq 0$ and $\lambda_{t-1}^\theta \geq 0$, so Eq. (A.36) implies that $\lambda_t^\theta \geq 0$ as well. By the same argument, $q^\theta \geq 0$ and $\lambda_{t-1}^{\theta,g} \geq 0$ for $g \in \{f, m\}$ imply $\lambda_t^{\theta,g} \geq 0$ for $g \in \{f, m\}$ as well by Eq. (A.31). Thus, $\lambda_t \in \Delta(\Theta)$, and $\lambda_t^g \in \mathbb{R}_+^\Theta$ for each g.

Claim 1 is immediate. For Claim 2, again we argue by induction. For $t = 0$, the claim is trivially true. Inductively, assume $\lambda_{t-1}^\theta > 0$. By Eq. (A.36), since as was just shown $1 - \sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'} \geq 0$, and the inductive hypothesis implies that $\lambda_{t-1}^\theta > 0$, if $q^\theta > 0$ then $\lambda_t^\theta \geq \lambda_{t-1}^\theta q^\theta > 0$. Suppose instead $q^\theta = 0$. If $\sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'} = 1$, then, since $q^{\theta'} \leq 1$ for all θ' by assumption, and $\lambda_{t-1} \in \Delta(\Theta)$, it must be that $\lambda_{t-1}^{\theta'} > 0$ implies $q^{\theta'} = 1$: but then $\lambda_{t-1}^\theta = 0$, which contradicts the inductive hypothesis. Thus, $0 \leq \sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'} < 1$, so Eq. (A.36) implies that $\lambda_t^\theta = \left(1 - \sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'}\right) \lambda_{t-1}^\theta > 0$.

For Claim 3, divide both sides of Eq. (A.36) for type θ by λ_{t-1}^{θ}, which is assumed to be positive; this yields

$$\frac{\lambda_t^{\theta}}{\lambda_{t-1}^{\theta}} = 1 + q^{\theta} - \sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'}. \tag{A.37}$$

A similar equation holds for $\tilde{\theta}$. This immediately yields 3(a). To derive 3(b), since $\lambda_t^{\theta'} = \lambda_0^{\theta'} \cdot \prod_{s=1}^{t} \frac{\lambda_s^{\theta'}}{\lambda_{s-1}^{\theta'}}$ for $\theta' = \theta, \tilde{\theta}$,

$$\frac{\lambda_t^{\theta}}{\lambda_t^{\tilde{\theta}'}} = \frac{\lambda_0^{\theta}}{\lambda_0^{\tilde{\theta}}} \cdot \frac{\prod_{s=1}^{t} \frac{\lambda_s^{\theta}}{\lambda_{s-1}^{\theta}}}{\prod_{s=1}^{t} \frac{\lambda_s^{\tilde{\theta}}}{\lambda_{s-1}^{\tilde{\theta}}}} = \frac{\lambda_0^{\theta}}{\lambda_0^{\tilde{\theta}}} \cdot \prod_{s=1}^{t} \frac{\frac{\lambda_s^{\theta}}{\lambda_{s-1}^{\theta}}}{\frac{\lambda_s^{\tilde{\theta}}}{\lambda_{s-1}^{\tilde{\theta}}}} = \frac{\lambda_0^{\theta}}{\lambda_0^{\tilde{\theta}}} \cdot \prod_{s=1}^{t} \frac{\frac{\lambda_s^{\tilde{\theta}}}{\lambda_{s-1}^{\tilde{\theta}}} + q^{\theta} - q^{\tilde{\theta}}}{\frac{\lambda_s^{\tilde{\theta}}}{\lambda_{s-1}^{\tilde{\theta}}}} = \frac{\lambda_0^{\theta}}{\lambda_0^{\tilde{\theta}}} \cdot \prod_{s=1}^{t} \left(1 + \frac{q^{\theta} - q^{\tilde{\theta}}}{\frac{\lambda_s^{\tilde{\theta}}}{\lambda_{s-1}^{\tilde{\theta}}}} \right).$$

If $q^{\theta} = q^{\tilde{\theta}}$, then every term in parentheses equals 1, and the claim follows. If instead $q^{\theta} > q^{\tilde{\theta}}$, recall that, by Eq. (A.37), for all $s \geq 1$, since $\lambda_{s-1} \in \Delta(\Theta)$ and $q \in [0,1]^{|\Theta|}$, $\frac{\lambda_s^{\tilde{\theta}}}{\lambda_{s-1}^{\tilde{\theta}}} \leq 1 + q^{\tilde{\theta}}$. Therefore, each term in parentheses is not smaller than $1 + \frac{q^{\theta} - q^{\tilde{\theta}}}{1 + q^{\tilde{\theta}}} > 1$. It follows that

$$\frac{\lambda_t^{\theta}}{\lambda_t^{\tilde{\theta}'}} = \frac{\lambda_0^{\theta}}{\lambda_0^{\tilde{\theta}}} \cdot \prod_{s=1}^{t} \left(1 + \frac{q^{\theta} - q^{\tilde{\theta}}}{\frac{\lambda_s^{\tilde{\theta}}}{\lambda_{s-1}^{\tilde{\theta}}}} \right) \geq \frac{\lambda_0^{\theta}}{\lambda_0^{\tilde{\theta}}} \cdot \left(1 + \frac{q^{\theta} - q^{\tilde{\theta}}}{1 + q^{\tilde{\theta}}} \right)^{t} \to \infty.$$

For Claim 4, consider first $\tilde{\theta} \notin \Theta^{\max}$, and fix $\theta \in \Theta^{\max}$ arbitrarily. Then $\frac{\lambda_t^{\theta}}{\lambda_t^{\tilde{\theta}}} \to \infty$ by Claim 3(b). Suppose that there is a subsequence $(\lambda_{t(\ell)})_{\ell \geq 0}$ such that $\lambda_{t(\ell)}^{\tilde{\theta}} \geq \epsilon$ for some $\epsilon > 0$ and all $\ell \geq 0$. Since $\frac{\lambda_{t(\ell)}^{\theta}}{\lambda_{t(\ell)}^{\tilde{\theta}}} \to \infty$ as well, there is ℓ large enough such that $\frac{\lambda_{t(\ell)}^{\theta}}{\lambda_{t(\ell)}^{\tilde{\theta}}} > \frac{1}{\epsilon}$: but then $\Lambda_{t(\ell)}^{\theta} > 1$ for such ℓ: contradiction. Thus, for every $\epsilon > 0$, eventually $\lambda_t^{\tilde{\theta}} < \epsilon$: that is, $\lambda_t^{\tilde{\theta}} \to 0$.

Next, consider $\tilde{\theta} \in \Theta^{\max}$. By Claim 2, $\lambda_t^{\tilde{\theta}} > 0$ and $\sum_{\theta \in \Theta^{\max}} \lambda_t^{\theta} > 0$, and

$$\frac{\lambda_t^{\tilde{\theta}}}{\sum_{\theta \in \Theta^{\max}} \lambda_t^{\theta}} = \frac{1}{\sum_{\theta \in \Theta^{\max}} \frac{\lambda_t^{\theta}}{\lambda_t^{\tilde{\theta}}}} = \frac{1}{\sum_{\theta \in \Theta^{\max}} \frac{\lambda_0^{\theta}}{\lambda_0^{\tilde{\theta}}}} = \frac{\lambda_0^{\tilde{\theta}}}{\sum_{\theta \in \Theta^{\max}} \lambda_0^{\theta}} = \bar{\lambda}^{\tilde{\theta}},$$

where the third inequality follows from Claim 3(b). Therefore,

$$\lambda_t^{\tilde{\theta}} = \frac{\lambda_t^{\tilde{\theta}}}{\sum_{\theta \in \Theta^{\max}} \lambda_t^{\theta}} \cdot \left(\sum_{\theta \in \Theta^{\max}} \lambda_t^{\theta} \right) = \bar{\lambda}^{\tilde{\theta}} \cdot \left(1 - \sum_{\theta \notin \Theta^{\max}} \lambda_t^{\theta} \right) \to \bar{\lambda}^{\tilde{\theta}},$$

because, as was just shown above, $\lambda_t^{\theta} \to 0$ for $\theta \notin \Theta^{\max}$.

Finally, consider Claim 5. Fix $g \in \{f, m\}$. First, since $0 \leq \lambda_t^{\theta,g} \leq \lambda_t^{\theta}$ for all $t \geq 0$, if $\theta \notin \Theta^{\max}$ then by Claim 4 $\lambda_t^{\theta} \to \bar{\lambda}^{\theta} = 0$, and so $\lambda_t^{\theta,g} \to 0 = \bar{\lambda}^{\theta,g}$ as well. Thus, focus on the case $\theta \in \Theta^{\max}$, so that by Claim 4 $\bar{\lambda}^{\theta} > 0$.

If $\sum_{\theta'} \bar{\lambda}^{\theta'} q^{\theta'} = 1$, then Eq. (A.31) and the fact that $\sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'} \in [0,1]$ and $0 \leq \lambda_{t-1}^{\theta,g} \leq \lambda_{t-1}^{\theta} \leq 1$ for all θ imply that

$$\lambda_t^{\theta,g} = \left(1 - \sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'}\right) \lambda_{t-1}^{\theta,g} + \lambda_{t-1}^{\theta} q^{\theta,g} \in \left[\lambda_{t-1}^{\theta} q^{\theta,g}, 1 - \sum_{\theta'} \lambda_{t-1}^{\theta'} q^{\theta'} + \lambda_{t-1}^{\theta} q^{\theta,g}\right]$$

and both endpoints of the interval in the r.h.s. converge to $\bar{\lambda}^{\theta} q^{\theta,g}$ by Claim 4 if $\sum_{\theta'} \bar{\lambda}^{\theta'} q^{\theta'} = 1$. Furthermore, the same assumption implies that $\bar{\lambda}^{\theta} q^{\theta,g} = \bar{\lambda}^{\theta,g}$, so $\lambda_t^{\theta,g} \to \bar{\lambda}^{\theta,g}$.

Now consider the case $0 < \sum_{\theta'} \bar{\lambda}^{\theta'} q^{\theta'} < 1$. (The set $\Theta^{\max}$ is non-empty, and since $q \in \mathbb{R}_+^{\Theta} \setminus \{0\}$, there is $\theta^+ \in \Theta^{\max}$ with $q^{\theta^+} > 0$; by Claim 4, $\bar{\lambda}^{\theta'} > 0$ for $\theta' \in \Theta^{\max}$, so in particular $\bar{\lambda}^{\theta^+} > 0$; but then $\sum_{\theta'} \bar{\lambda}^{\theta'} q^{\theta'} \geq \bar{\lambda}^{\theta^+} q^{\theta^+} > 0$.) It is convenient to let $q_t = \sum_{\theta'} \lambda_t^{\theta'} q^{\theta'}$ and $\bar{q} = \sum_{\theta'} \bar{\lambda}^{\theta'} q^{\theta'} = \lim_{t \to \infty} q_t$, where the second equality follows from Claim 4. Thus, Eq. (A.31) can be written as

$$\lambda_t^{\theta,g} = (1 - q_{t-1}) \lambda_{t-1}^{\theta,g} + \lambda_{t-1}^{\theta} q^{\theta,g}. \tag{A.38}$$

In addition, $\bar{q} \in (0,1)$.

We claim that, for all $T \geq 0$ and $t > T$,

$$\lambda_t^{\theta,g} = \lambda_T^{\theta,g} \prod_{s=T}^{t-1} (1 - q_s) + q^{\theta,g} \sum_{s=T}^{t-1} \lambda_s^{\theta} \prod_{r=s+1}^{t-1} (1 - q_r). \tag{A.39}$$

For $t = T+1$, this follows from Eq. (A.38). Inductively, assume it holds for $t - 1 > T$. Then, by Eq. (A.38) and the inductive hypothesis,

$$\begin{aligned}
\lambda_t^{\theta,g} &= (1 - q_{t-1}) \left[\lambda_T^{\theta,g} \prod_{s=T}^{t-2} (1 - q_s) + q^{\theta,g} \sum_{s=T}^{t-2} \lambda_s^{\theta} \prod_{r=s+1}^{t-2} (1 - q_r)\right] + \lambda_{t-1}^{\theta} q^{\theta,g} = \\
&= \lambda_T^{\theta,g} \prod_{s=T}^{t-1} (1 - q_s) + q^{\theta,g} \sum_{s=T}^{t-1} \lambda_s^{\theta} \prod_{r=s+1}^{t-1} (1 - q_r),
\end{aligned}$$

as claimed.

Fix $\epsilon > 0$ such that $\bar{\lambda}^{\theta} - \epsilon > 0$, $\bar{q} - \epsilon > 0$, $1 - \bar{q} + \epsilon < 1$, and $1 - \bar{q} - \epsilon > 0$. This is possible because $\bar{\lambda}^{\theta} > 0$ and $\bar{q} \in (0,1)$, hence $1 - \bar{q} \in (0,1)$.

Since $\lambda_t^{\theta} \to \bar{\lambda}^{\theta}$ and $q_t \to \bar{q}$, there is $T \geq 0$ such that, for all $t > T$, $\lambda_t^{\theta} < \bar{\lambda}^{\theta} + \epsilon$ and

$q_t > \bar{q} - \epsilon$. Hence, for such $t > T$, Eq. (A.39) implies that

$$
\begin{aligned}
\lambda_t^{\theta,g} \leq & \lambda_T^{\theta,g} \prod_{s=T}^{t-1} (1 - \bar{q} + \epsilon) + q^{\theta,g} \sum_{s=T}^{t-1} (\bar{\lambda}^\theta + \epsilon) \prod_{r=s+1}^{t-1} (1 - \bar{q} + \epsilon) = \\
= & \lambda_T^{\theta,g} (1 - \bar{q} + \epsilon)^{t-T} + q^{\theta,g} (\bar{\lambda}^\theta + \epsilon) \sum_{s=T}^{t-1} (1 - \bar{q} + \epsilon)^{t-1-s} = \\
= & \lambda_T^{\theta,g} (1 - \bar{q} + \epsilon)^{t-T} + q^{\theta,g} (\bar{\lambda}^\theta + \epsilon) \sum_{s=0}^{t-1-T} (1 - \bar{q} + \epsilon)^s = \\
= & \lambda_T^{\theta,g} (1 - \bar{q} + \epsilon)^{t-T} + q^{\theta,g} (\bar{\lambda}^\theta + \epsilon) \frac{1 - (1 - \bar{q} + \epsilon)^{t-T}}{\bar{q} - \epsilon} \to \frac{q^{\theta,g} (\bar{\lambda}^\theta + \epsilon)}{\bar{q} - \epsilon}.
\end{aligned}
$$

This implies that $\limsup_t \lambda_t^{\theta,g} \leq \frac{q^{\theta,g}(\bar{\lambda}^\theta + \epsilon)}{\bar{q} - \epsilon}$. Since this must hold for all $\epsilon > 0$, it must be that $\limsup_t \lambda_t^{\theta,g} \leq \frac{q^{\theta,g} \bar{\lambda}^\theta}{\bar{q}} = \bar{\lambda}^{\theta,g}$.

Similarly, $\lambda_t^\theta \to \bar{\lambda}^\theta$ and $q_t \to \bar{q}$ imply that there is $T \geq 0$ such that, for all $t > T$, $\lambda_t^\theta > \bar{\lambda}^\theta - \epsilon > 0$ and $q_t < \bar{q} + \epsilon < 1$. Then

$$
\begin{aligned}
\lambda_t^{\theta,g} \geq & \lambda_T^{\theta,g} \prod_{s=T}^{t-1} (1 - \bar{q} - \epsilon) + q^{\theta,g} \sum_{s=T}^{t-1} (\bar{\lambda}^\theta - \epsilon) \prod_{r=s+1}^{t-1} (1 - \bar{q} - \epsilon) = \\
= & \lambda_T^{\theta,g} (1 - \bar{q} - \epsilon)^{t-T} + q^{\theta,g} (\bar{\lambda}^\theta - \epsilon) \frac{1 - (1 - \bar{q} - \epsilon)^{t-T}}{\bar{q} + \epsilon} \to \frac{q^{\theta,g} (\bar{\lambda}^\theta - \epsilon)}{\bar{q} + \epsilon},
\end{aligned}
$$

so $\liminf_t \lambda_t^{\theta,g} \geq \frac{q^{\theta,g}(\bar{\lambda}^\theta - \epsilon)}{\bar{q} + \epsilon}$. Again, since this must hold for all $\epsilon > 0$, $\liminf_t \lambda_T^{\theta,g} \geq \frac{q^{\theta,g} \bar{\lambda}^\theta}{\bar{q}} = \bar{\lambda}^{\theta,g}$. Hence, $\lambda_t^{\theta,g} \to \bar{\lambda}^{\theta,g}$. *Q.E.D.*

Next, we establish certain basic properties of the symmetric model considered in the paper. Claims 1 and 3 characterize the set $\Theta^{\max}$ for this specification. Claim 2 ensures that the parameterization satisfies the conditions in Theorem A.1.

Lemma A.1 *Assume that, for every $\theta \in \Theta$, γ^θ, $p^{\theta,m}$ and $p^{\theta,f}$ are as defined in Section 2.. 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 $\rho > \bar{\rho}(\phi, N)$.*

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.*

Recall that $\bar{\rho}(\cdot)$ is defined in Eq. (10).

Proof: Write

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

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 these two extremes:

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

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

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

or equivalently

$$\rho > \frac{1}{4}\left(\left(\frac{1-\phi}{\phi}\right)^{\frac{N}{2}} + \left(\frac{\phi}{1-\phi}\right)^{\frac{N}{2}}\right)^2 = \bar{\rho}(\phi, N), \tag{A.40}$$

which is 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. (A.40) 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.*

We now turn to the proofs of the main Propositions and Corollaries in the text.

Proof of Proposition 3 and Corollary 1: convergence of $(\lambda_t)_{t\geq 0}$, $(\lambda_t^m)_{t\geq 0}$ and $(\lambda_t^f)_{t\geq 0}$ follows from Theorem A.1 and Claim 2 of Lemma A.1. Parts (a) and (b) follow from Claim 1 in Lemma A.1 and Claim 4 in Theorem A.1. Corollary 1 follows from Claim 3 in Lemma A.1. *Q.E.D.*

Proposition 2 follows from Proposition 3.

Proof of Proposition 4: Fix $\theta \in \Theta$, and define θ^{sym} by $\theta^{\text{sym}}_n = \theta_{N+1-n}$ for all $n = 1, \ldots, N$. (Notice that, for some θ, it may be the case that $\theta^{\text{sym}} = \theta$.) We first claim that

$$a_t^{\theta,m} + a_t^{\theta^{\text{sym}},m} \geq a_t^{\theta,f} + a_t^{\theta^{\text{sym}},f}. \tag{A.41}$$

Notice that, if $\theta^{\text{sym}} = \theta$, the above inequality just says that $a_t^{\theta,m} \geq a_t^{\theta,f}$.

Let $m_0 = \sum_{n=1}^{N/2} \theta$ and $m_1 = \sum_{n=N/2+1}^{N} \theta_n$. 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 $p^{\theta^{sym},m} = [\phi(1-\phi)]^{N/2}\left(\frac{1-\phi}{\phi}\right)^{m_0-m_1}$. Moreover, since p_f is defined with the roles of ϕ and $1-\phi$ reversed, $p^{\theta,f} = p^{\theta^{\text{sym}},m}$ and $p^{\theta,m} = p^{\theta^{\text{sym}},f}$, so $p^{\theta,m} + p^{\theta,f} = p^{\theta^{\text{sym}},m} + p^{\theta^{\text{sym}},f}$. Finally, by construction $\gamma^{\theta} = \gamma^{\theta^{sym}}$.

Suppose that $m_0 \geq m_1$. Since $\phi > \frac{1}{2}$, $p^{\theta,m} \geq p^{\theta^{\text{sym}},m}$. At time 0 we thus have $\lambda_0^{\theta} = p^{\theta,m} \geq p^{\theta^{\text{sym}},m} = \lambda_0^{\theta^{\text{sym}}} > 0$. Then, since $q^{\theta} = \gamma^{\theta}(p^{\theta,m} + p^{\theta,f}) + \gamma^{\theta^{sym}}(p^{\theta^{sym},m} + p^{\theta^{sym},f}) = q^{\theta^{sym}}$, by part 3(a) of Theorem A.1, for every $t > 0$, $\frac{\lambda_t^{\theta}}{\lambda_{t-1}^{\theta}} = \frac{\lambda_t^{\theta^{\text{sym}}}}{\lambda_{t-1}^{\theta^{\text{sym}}}}$, and hence $\frac{\lambda_t^{\theta}}{\lambda_t^{\theta^{\text{sym}}}} = \frac{\lambda_{t-1}^{\theta}}{\lambda_{t-1}^{\theta^{\text{sym}}}} = \frac{\lambda_0^{\theta}}{\lambda_0^{\theta^{\text{sym}}}} \geq 1$. Thus, $\lambda_t^{\theta} \geq \lambda_t^{\theta^{\text{sym}}}$ for all $t > 0$ as well. Finally, letting $\bar{\gamma} \equiv \gamma^{\theta^{\text{sym}}} = \gamma^{\theta}$, for every $t \geq 1$,

$$a_t^{\theta} = a_t^{\theta,m} + a_t^{\theta,f} = \bar{\gamma}\lambda_{t-1}^{\theta}(p^{\theta,m} + p^{\theta,f}) \geq \bar{\gamma}\lambda_{t-1}^{\theta^{\text{sym}}}(p^{\theta^{\text{sym}},m} + p^{\theta^{\text{sym}},f}) = a_t^{\theta^{\text{sym}},m} + a_t^{\theta^{\text{sym}},f} = a_t^{\theta^{\text{sym}}}.$$

All the inequalities in the above paragraph are strict if $m_0 > m_1$; they are reversed if $m_0 \leq m_1$; and hold as equalities if $m_0 = m_1$.

Now, regardless of the values of m_0 and m_1,

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

where the last step follows from $p^{\theta,m} = p^{\theta^{\text{sym}},f}$ and $p^{\theta,f} = p^{\theta^{\text{sym}},m}$.

If $m_0 = m_1$, then both terms in square brackets equal zero, so equality obtains; in particular, this is true if $\theta = \theta^{\text{sym}}$. If $m_0 > m_1$, then both terms are positive, if $m_0 < m_1$, then both terms are negative. Thus, in any event, the last inequality, and hence Eq. (A.41), holds; furthermore, if $\theta = \theta^{\text{sym}}$, then $a_t^{\theta,m} = a_t^{\theta,f}$.

Now fix $L \in \{0, \ldots, N\}$. Then

$$\begin{aligned}
\sum_{\theta:\sum_n \theta_n = L} a_t^{\theta,m} &= \sum_{\theta:\sum_n \theta_n = L, \theta = \theta^{\text{sym}}} a_t^{\theta,m} + \sum_{\theta:\sum_n \theta_n = L, \theta \neq \theta^{\text{sym}}} a_t^{\theta,m} = \\
&= \sum_{\theta:\sum_n \theta_n = L, \theta = \theta^{\text{sym}}} a_t^{\theta,m} + \frac{1}{2} \sum_{\theta:\sum_n \theta_n = L, \theta \neq \theta^{\text{sym}}} [a_t^{\theta,m} + a_t^{\theta^{\text{sym}},m}] \geq \\
&\geq \sum_{\theta:\sum_n \theta_n = L, \theta = \theta^{\text{sym}}} a_t^{\theta,f} + \frac{1}{2} \sum_{\theta:\sum_n \theta_n = L, \theta \neq \theta^{\text{sym}}} [a_t^{\theta,f} + a_t^{\theta^{\text{sym}},f}] = \\
&= \sum_{\theta:\sum_n \theta_n = L} a_t^{\theta,f}.
\end{aligned}$$

The second equality follows from the observation that, restricting attention to types θ with $\sum_n \theta_n = L$, also $\sum_n \theta_n^{\text{sym}} = L$, so that adding $a_t^{\theta,m} + a_t^{\theta^{\text{sym}},m}$ over all θ with $\theta \neq \theta^{\text{sym}}$ counts each type twice. The inequality follows from Eq. (A.41), which in particular implies that $a_t^{\theta,m} = a_t^{\theta,f}$ if $\theta = \theta^{\text{sym}}$. This inequality is strict if the second summation is non-empty, i.e., if there is θ with $\sum_n \theta_n = L$ and $\theta_n \neq \theta_{N+1-n}$ for some n, because the latter condition implies $\theta \neq \theta^{\text{sym}}$. Finally, the last equality follows by repeating the first two steps backwards, for F-group researchers. *Q.E.D*

Proof of Proposition 5: We begin with a preliminary result.

Lemma A.2 *For all parameter values and initial conditions, and for all $\theta \in \Theta$ and $t \geq 1$,*

$$\frac{\lambda_t^\theta}{\lambda_{t-1}^\theta} = (1 - a_t) + \gamma^\theta (p^{\theta,m} + p^{\theta,f});$$

and for $t \geq 2$,

$$\frac{a_t^\theta}{a_{t-1}^\theta} = \frac{a_t^{\theta,m}}{a_{t-1}^{\theta,m}} = \frac{a_t^{\theta,f}}{a_{t-1}^{\theta,f}} = \frac{\lambda_{t-1}^\theta}{\lambda_{t-2}^\theta}.$$

Proof: From Eq. (7), $\lambda_t^\theta = \lambda_t^{\theta,m} + \lambda_t^{\theta,f} = (\lambda_{t-1}^{\theta,m} + \lambda_{t-1}^{\theta,f})(1 - a_t) + \gamma^\theta (p^{\theta,m} + p^{\theta,f})$, which yields the first equation because $\lambda_\tau^\theta > 0$ for all θ and τ.

From Eq. (6), for $t \geq 2$,

$$\frac{a_t^{\theta,g}}{a_{t-1}^{\theta,g}} = \frac{\lambda_{t-1}^\theta \gamma^\theta p^{\theta,g}}{\lambda_{t-2}^\theta \gamma^\theta p^{\theta,g}} = \frac{\lambda_{t-1}^\theta}{\lambda_{t-2}^\theta};$$

similarly,

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

Q.E.D.

We now prove Proposition 5. For $N=2$ we only have 4 types, $\Theta=\{(0,0),(1,0),(0,1),(1,1)\}$. Let $a^{L,g}=\sum_{sum_{n=1}^{2}\theta_n=L} a^{\theta,g}$ and $a_t^g=\sum_{\ell=0}^{2} a_t^{\ell,g}$. From Proposition 4, for all t, $a_t^{1,m}>a_t^{1,f}$, $a_t^{2,m}=a_t^{2,f}$, and $a_t^{0,m}=a_t^{0,f}$. Therefore, $a_t^m>a_t^f$, which implies that the weight on $L=1$ for accepted M researchers is

$$\frac{a_t^{1,m}}{a_t^m}=1-\frac{a_t^{2,m}+a_t^{0,m}}{a_t^m}=1-\frac{a_t^{2,f}+a_t^{0,f}}{a_t^m}>1-\frac{a_t^{2,f}+a_t^{0,f}}{a_t^f}=\frac{a_t^{1,f}}{a_t^f}.$$

Similarly, $a_t^m>a_t^f$ and $a_t^{0,m}=a_t^{0,f}$, $a_t^{2,m}=a_t^{2,f}$ imply

$$\frac{a_t^{0,m}}{a_t^m}<\frac{a_t^{0,f}}{a_t^f},\qquad \frac{a_t^{2,m}}{a_t^m}<\frac{a_t^{2,f}}{a_t^f}.$$

Moreover, we claim that, $a_t^{2,g}>a_t^{0,g}$. For $t=0$, $a_0^{2,g}=a_0^{(1,1),g}=p^{(1,1),m}\gamma^{(1,1)}p^{(1,1),g}>p^{(0,0),m}\gamma^{(0,0)}p^{(0,0),g}=a_0^{(0,0),g}=a_0^{0,g}$, because $p^{(0,0),g}=p^{(1,1),g}$ but $\gamma^{(1,1)}>\gamma^{(0,0)}$. Inductively, from Lemma A.2,

$$\begin{aligned}
a_t^{2,g}=a_t^{(1,1),g}=a_{t-1}^{(1,1),g}\cdot\frac{a_t^{(1,1),g}}{a_{t-1}^{(1,1),g}}=a_{t-1}^{(1,1,g)}\left(1-a_{t-1}+\gamma^{(1,1)}(p^{(1,1),m}+p^{(1,1),f}\right)>\\
>a_{t-1}^{(1,1,g)}\left(1-a_{t-1}+\gamma^{(0,0)}(p^{(0,0),m}+p^{(0,0),f}\right)>a_{t-1}^{(0,0,g)}\left(1-a_{t-1}+\gamma^{(0,0)}(p^{(0,0),m}+p^{(0,0),f}\right)=\\
=a_{t-1}^{(0,0,g)}\frac{a_t^{(0,0),g}}{a_{t-1}^{(0,0,g)}}=a_t^{(0,0),g}=a_t^{0,g}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
0<&\frac{a_t^{0,f}}{a_t^{1,f}+a_t^{2,f}+a_t^{0,f}}-\frac{a_t^{0,m}}{a_t^{1,m}+a_t^{2,m}+a_t^{0,m}}=\frac{a_t^{0,f}}{a_t^{1,f}+a_t^{2,f}+a_t^{0,f}}-\frac{a_t^{0,f}}{a_t^{1,m}+a_t^{2,m}+a_t^{0,m}}<\\
<&\left(\frac{a_t^{2,f}}{a_t^{0,f}}\right)\cdot\left(\frac{a_t^{0,f}}{a_t^{1,f}+a_t^{2,f}+a_t^{0,f}}-\frac{a_t^{0,f}}{a_t^{1,m}+a_t^{2,m}+a_t^{0,m}}\right)=\frac{a_t^{2,f}}{a_t^{1,f}+a_t^{2,f}+a_t^{0,f}}-\frac{a_t^{2,f}}{a_t^{1,m}+a_t^{2,m}+a_t^{0,m}}=\\
=&\frac{a_t^{2,f}}{a_t^{1,f}+a_t^{2,f}+a_t^{0,f}}-\frac{a_t^{2,m}}{a_t^{1,m}+a_t^{2,m}+a_t^{0,m}};
\end{aligned}$$

the first inequality follows from $a_t^{1,f}<a_t^{1,m}$ and $a_t^{0,f}=a_t^{0,m}$ and $a_t^{2,f}=a_t^{2,m}$, the next equality from $a_t^{0,m}=a_t^{0,f}$, the second inequality from $a_t^{2,f}>a_t^{0,f}>0$ and the fact that the difference of fractions is positive, and the last equality from $a_t^{2,m}=a_t^{2,f}$.

The result then follows from a symmetry argument.

$$\begin{aligned}
E[L|F] &= \frac{0\times a_t^{0,f}+a_t^{1,f}+2a_t^{2,f}}{a_t^f}\\
E[L|M] &= \frac{0\times a_t^{0,m}+a_t^{1,m}+2a_t^{2,m}}{a_t^m}
\end{aligned}$$

which, since $a_t^{1,g} = 1 - a_t^{0,g} - a_t^{2,g}$, implies

$$E[L|F] = -1\frac{a_t^{0,f}}{a_t^f} + 1 + \frac{a_t^{2,f}}{a_t^f}$$
$$E[L|M] = -1\frac{a_t^{0,m}}{a_t^m} + 1 + \frac{a_t^{2,m}}{a_t^m}$$

It follows that

$$E[L|F] - E[L|M] = -\left(\frac{a_t^{0,f}}{a_t^f} - \frac{a_t^{0,m}}{a_t^m}\right) + \left(\frac{a_t^{2,f}}{a_t^f} - \frac{a_t^{2,m}}{a_t^m}\right) > 0$$

Q.E.D

Detailed dynamics of the mass of M and F accepted agents. It is useful to rewrite Equation (7) for each group $g = m, f$ as follows:

$$\lambda_t^{\theta,m} - \lambda_{t-1}^{\theta,m} = -\lambda_{t-1}^{\theta,m}\, a_t + \lambda_{t-1}^{\theta,m}\gamma^\theta p^{\theta,m} + \lambda_{t-1}^{\theta,f}\gamma^\theta p^{\theta,m} \tag{A.42}$$
$$\lambda_t^{\theta,f} - \lambda_{t-1}^{\theta,f} = -\lambda_{t-1}^{\theta,f}\, a_t + \lambda_{t-1}^{\theta,f}\gamma^\theta p^{\theta,f} + \lambda_{t-1}^{\theta,m}\gamma^\theta p^{\theta,f} \tag{A.43}$$

Consider the dynamics of F-researchers in (A.43), for instance. The change in the mass of F-researchers of type θ decreases due to replacement at the rate a_t, and it then increases due to the young F-researchers who produce quality research and are matched with referees from the F group who share their type and hence view them positively ($\lambda_{t-1}^{\theta,f}\gamma^\theta p^{\theta,f}$), plus the young F-researchers who produce quality research and are matched with M-referees of their own type ($\lambda_{t-1}^{\theta,m}\gamma^\theta p^{\theta,f}$). The asymmetry between the two dynamics (A.42) and (A.43) is apparent in the last two terms of each. If θ is a type that is more prevalent among M-researchers—for instance, $\theta = \theta^m$—then $p^{\theta,f}$ will be small while $p^{\theta,m}$ will be large. If the current mass of M-researchers of type θ is large, then $\lambda_{t-1}^{\theta}\gamma^\theta p^{\theta,m}$ will act to further increase the mass of M-researchers, while the respective term $\lambda_{t-1}^{\theta}\gamma^\theta p^{\theta,f}$ in the F-group dynamics will lead to a smaller increase in the mass of type-θ F-researchers. In particular, if we start from a situation in which *all* referees of type θ are in M-group, then, while they will accept some F-researchers of type θ, they will accept a much larger mass of M-researchers.

This force is at play regardless of the parameter values, and for all types. However, its implications for the limiting group (im)balance in the population depend upon whether or not we are in a "meritocratic" scenario. If research characteristics have a limited effect on the probability of quality research, as in Part (a) of Proposition 3, then θ^m and θ^f are the only types that survive in the limit. These are also the types for which the difference in proportions among young M- and F-researchers is greatest. Thus, in the scenario of Part (a), the force thus described has the greatest effect, which is further reinforced if initially *all* referees are in M-group. The result is that, in the limit, despite the fact that the mass of young M- and F-researchers appearing at each time t is the same, the referees' self-image bias leads to a limiting population in which the majority of scholars are in M group.

By way of contrast, in the meritocratic scenario of Part (b) in Proposition 3, the type that prevails in the limit is the efficient one, namely θ^*. In our symmetric model, the *same* fraction of young M- and F-researchers are of type θ^*. Therefore, the effect described above becomes more and more muted over time. Consequently, in the limit, the mass of M- and F-scholars is the same.

The following Proposition formalizes the above discussion. We denote by $\Lambda_t^m \equiv \sum_\theta \lambda_t^{\theta,m}$ and $\Lambda_t^f \equiv \sum_\theta \lambda_t^{\theta,f}$ the total mass of M- and F-scholars at date t; $\bar{\Lambda}^m$ and $\bar{\Lambda}^f$ are the corresponding limiting quantities.

Proposition A.2 *Assume that all referees are initially from the M-group, i.e., $\lambda_0 = p^m$.*

(a) If $\rho < \bar{\rho}(\phi, N)$, then the limiting masses are

$$\text{(M-researchers of type θ^m): } \bar{\lambda}^{\theta^m,m} = \frac{(\phi^N)^2}{(\phi^N + (1-\phi)^N)^2}; \tag{A.44}$$

$$\text{(F-researchers of type θ^m): } \bar{\lambda}^{\theta^m,f} = \frac{\phi^N \ (1-\phi)^N}{(\phi^N + (1-\phi)^N)^2}; \tag{A.45}$$

$$\text{(M-researchers of type θ^f): } \bar{\lambda}^{\theta^f,m} = \frac{((1-\phi)^N)^2}{(\phi^N + (1-\phi)^N)^2}; \tag{A.46}$$

$$\text{(F-researchers of type θ^f): } \bar{\lambda}^{\theta^f,f} = \frac{(1-\phi)^N \ \phi^N}{(\phi^N + (1-\phi)^N)^2}; \tag{A.47}$$

with

$$\bar{\lambda}^{\theta^m,m} > \bar{\lambda}^{\theta^m,f} = \bar{\lambda}^{\theta^f,f} > \bar{\lambda}^{\theta^f,m} \tag{A.48}$$

In addition, the total mass of M and F researchers are

$$\bar{\Lambda}^m = 1 - \bar{\Lambda}^f = \frac{1 + \left(\frac{\phi}{1-\phi}\right)^{2N}}{1 + \left(\frac{\phi}{1-\phi}\right)^{2N} + 2\left(\frac{\phi}{1-\phi}\right)^{N}} > 0.5. \tag{A.49}$$

(b) If $\rho > \bar{\rho}(\phi, N)$, then $\bar{\lambda}^{\theta^*,m} = \bar{\lambda}^{\theta^*,f} = \bar{\Lambda}^m = \bar{\Lambda}^f = \frac{1}{2}$.

Proof of Proposition 6, A.2 and Corollary 2. For Part (a), since $\gamma^{\theta^m} = \gamma^{\theta^f} = \gamma_0(\rho)^{N/2}$ and, by Proposition 3, $\Theta^{\max} = \{\theta^m, \theta^f\}$, $\bar{\lambda}^{\tilde{\theta},m} = \frac{\lambda_0^{\tilde{\theta}} p^{\tilde{\theta},m}}{\lambda_0^{\theta^m} p^{\theta^m,m} + \lambda_0^{\theta^f} p^{\theta^f,m}}$ for $\tilde{\theta} \in \Theta^{\max}$, and $\bar{\lambda}^{\tilde{\theta},m} = 0$ otherwise; a similar expression holds for $\bar{\lambda}^{\tilde{\theta},f}$. Equations (A.44) through (A.47) then follow from the specification of p^m and p^f. Eq. (13) follows from $\bar{\Lambda}^g = \bar{\lambda}^{\theta^m,g} + \bar{\lambda}^{\theta^f,g}$.

Part (b) follows from the fact that, by Proposition 3 part (b), $\Theta^{\max} = \{\theta^*\}$ in this scenario. Corollary 2 follows from Lemma A.1 Claim (3).

Proposition 6 consists of (b) and the last claim in (a) of Proposition A.2. *Q.E.D.*

Proof of Proposition 7: 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 $\lambda_{\tau,j}^m, \lambda_{\tau,j}^f, \lambda_{\tau,j} \in \mathbb{R}_+^\Theta$ such that

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

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

(iii) $\lambda_{\tau,j}^\theta < \frac{C}{\gamma^\theta (P-U)}$ for $0 \leq \tau \leq t(j) - t(j-1)$ and all $\theta \in \Theta \setminus \Theta_j$, and $\lambda_{t(j)-t(j-1),j}^{\theta_0} < \frac{C}{\gamma^{\theta_0}(P-U)}$ 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 $\lambda_{\tau,k}^m, \lambda_{\tau,k}^f, \lambda_{\tau,k} \in \mathbb{R}_+^\Theta$ for all $\tau \geq 1$ such that (i) holds for $j = k$, and $\lambda_{\tau,j}^\theta \geq \frac{C}{\gamma^\theta P}$ for all $\theta \in \Theta_k$. In the latter case, if the sequences of such vectors converge, then $\lim_{\tau \to \infty} \lambda_{\tau,k}^m = \lim_{t \to \infty} \lambda_t^m$ and similarly for $\lambda_{\tau,k}^f$ and $\lambda_{\tau,k}$.

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

Define $q_k^g \in \mathbb{R}_+^\Theta \setminus \{0\}$ for $g = f, m$ by $q_k^{\theta,g} = \gamma^\theta p^{\theta,g}$ if $\theta \in \Theta_k$, and $q_k^{\theta,g} = 0$ otherwise. Then $q_k^{\theta,m} + q_k^{\theta,f} \leq 1$ for all θ. Consider the sequences $(\lambda_{\tau,k}^{\theta,g})_{\tau \geq 0}$ for $g = f, m$ and $(\lambda_{\tau,k}^\theta)_{\tau \geq 0}$ defined by Eqs. (A.31)–(A.32) for the vectors q_k^f, q_k^m .

Suppose first that there are $\bar{\tau} > 0$ and $\theta_0 \in \Theta_k$ such that $\lambda_{\bar{\tau},k}^{\theta_0} < \frac{C}{\gamma^{\theta_0}(P-U)}$. Let $t(k) = t(k-1) + \bar{\tau}$. Then, for each group $g = f, m$, the dynamics in Eqs. (A.31)–(A.32) induced by the vectors q_k^f, q_k^m for the subsequence $(\lambda_{\tau,k}^g)_{\tau=0,\ldots,\bar\tau}$ coincide with those in Eq. (18) for the subsequences $(\lambda_t^g)_{t=t(k-1),\ldots,t(k)}$; 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 $q_k^{\theta,m}+q_k^{\theta,f}=0$ for $\theta\in\Theta\setminus\Theta_k$; hence, for all $\theta'\in\Theta$ and all $\theta\in\Theta\setminus\Theta_k$, $q_k^{\theta,m}+q_k^{\theta,f}\le q_{m,k}^{\theta'}+q_{f,k}^{\theta'}$. By part 3(a) in Theorem A.1, it must be the case that $\lambda_{\tau+1,k}^\theta/\lambda_{\tau,k}^\theta\le 1$: otherwise, $\sum_{\theta'\in\Theta}\lambda_{\tau+1,k}^{\theta'}>\sum_{\theta'\in\Theta}\lambda_{\tau,k}^{\theta'}=1$, which contradicts the fact that $\lambda_{\tau+1,k}\in\Delta(\Theta)$ per Theorem A.1. Since by definition $\lambda_{0,k}^\theta<\frac{C}{\gamma^\theta P}$ for $\theta\notin\Theta_k$, it follows that also $\lambda_{\tau,k}^\theta<\frac{C}{\gamma^\theta P}$ for $\tau=0,\ldots,\bar\tau$ and for any such θ. Thus, in this case Conditions $C(k)$ hold.

If instead $\lambda_{\bar\tau,k}^\theta\ge\frac{C}{\gamma^\theta(P-U)}$ for all $\theta\in\Theta_k$, then for each group $g=f,m$, the dynamics in Eqs. (A.31)–(A.32) induced by the vectors $q_{m,k},q_{f,k}$ for the subsequence $(\lambda_{\tau,k}^g)_{\tau\ge 0}$ coincide with those in Eq. (18) for the subsequence $(\lambda_t^g)_{t\ge t(k-1)}$. Again, in this case (i) holds for $j=k$. This completes the proof of the claim.

Since the set Θ is finite, there exists $K\ge 0$ such that the induction stops—that is, $\lambda_{\bar\tau,K}^\theta\ge\frac{C}{\gamma^\theta(P-U)}$ for all $\theta\in\Theta_K$. Let $\Theta_k^{\max}=\arg\max\{q_k^{\theta,m}+q_k^{\theta,f}:\theta\in\Theta\}$. Since $\Theta_0\supsetneq\Theta_1\supsetneq\ldots\supsetneq\Theta_K$, by the definition of the vectors q_k^g for $g=f,m$, also $\Theta_0^{\max}\supseteq\Theta_1^{\max}\supseteq\ldots\supseteq\Theta_K^{\max}$. Moreover, for every $k=0,\ldots,K-1$, and every $\theta\in\Theta_k^{\max}$, $\lambda_{\tau+1,k}^\theta/\lambda_{\tau,k}^\theta\ge 1$ for $0\le\tau<t(k)-t(k)$; otherwise, by part 3(a) in Theorem A.1, $\sum_{\theta\in\Theta}\lambda_{\tau+1,k}^\theta<\sum_{\theta\in\Theta}\lambda_{\tau,k}^\theta=1$, which contradicts the fact that $\lambda_{\tau+1}\in\Delta(\Theta)$ per Theorem A.1.

Now assume that $\Theta_0^{\max}\subseteq\Theta_0$. Then, for every $\theta\in\Theta_0^{\max}$,

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

so $\theta\in\Theta_k$ for all $k=0,\ldots,K$, and thus $\Theta_0^{\max}=\Theta_1^{\max}=\ldots=\Theta_K^{\max}\equiv\Theta^{\max}$. In addition, again by part 3(a) of Theorem A.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 $\tau\ge 0$. 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 A.1 part (4),

$$\bar\lambda^\theta=\bar\lambda_K^\theta=\frac{\lambda_{0,K}^\theta}{\sum_{\theta'\in\Theta^{\max}}\lambda_{0,K}^{\theta'}}=\frac{1}{\sum_{\theta'\in\Theta^{\max}}\frac{\lambda_{0,K}^{\theta'}}{\lambda_{0,K}^\theta}}=\frac{1}{\sum_{\theta'\in\Theta^{\max}}\frac{\lambda_0^{\theta'}}{\lambda_0^\theta}}=\frac{\lambda_0^\theta}{\sum_{\theta'\in\Theta^{\max}}\lambda_0^{\theta'}}.\tag{A.50}$$

Similarly, for $\theta \in \Theta^{\max}$, part (5) in the same Theorem implies that

$$\bar{\lambda}^{\theta,m} = \bar{\lambda}_K^{\theta,m} = \frac{\lambda_{0,K}^{\theta} q_K^{\theta,m}}{\sum_{\theta' \in \Theta^{\max}} \lambda_{0,K}^{\theta'} q_K^{\theta'}} = \frac{q_K^{\theta,m}}{\sum_{\theta' \in \Theta^{\max}} \frac{\lambda_{0,K}^{\theta'}}{\lambda_{0,K}^{\theta}} q_K^{\theta'}} = \frac{q_K^{\theta,m}}{\sum_{\theta' \in \Theta^{\max}} \frac{\lambda_0^{\theta'}}{\lambda_0^{\theta}} q_K^{\theta'}} = \frac{\lambda_0^{\theta} q_K^{\theta,m}}{\sum_{\theta' \in \Theta^{\max}} \lambda_0^{\theta'} q_K^{\theta'}}, \tag{A.51}$$

and analogously for $\bar{\lambda}^{\theta,f}$.

Statements (a.1)–(b) now follow. Recall that $\lambda_0 = p^m$. In (a.1), by assumption $\Theta^{\max} = \Theta_0^{\max} = \{\theta^m, \theta^f\} \subseteq \Theta_0$. Substituting $\lambda_0^{\theta^m} = \phi^N$ and $\lambda_0^{\theta^f} = (1-\phi)^N$ in Eq. (A.50) yields $\bar{\lambda}^{\theta^m} = \frac{\phi^N}{\phi^N + (1-\phi)^N}$. Similarly, substituting for q_K^g, $g = f, m$, and $q_K = q_K^f + q_K^m$ in Eq. (A.51) yields the same expression for $\bar{\lambda}^{\theta^m,m}$ as in Proposition 3, because $\theta \in \Theta^{\max}$ implies that $q_K^{\theta,g} = \gamma^\theta p^{\theta,g}$; ditto for $\bar{\lambda}^{\theta^m,f}$, $\bar{\lambda}^{\theta^f,m}$ and $\bar{\lambda}^{\theta^f,f}$, and hence for $\bar{\Lambda}^m$.

For (a.2), $\Theta^{\max} = \Theta_0^{\max} = \{\theta^m\}$. This immediately implies that $\bar{\lambda}^{\theta^m} = \bar{\lambda}_K^{\theta^m} = 1$. Furthermore, from Eq. (A.51), $\bar{\Lambda}^m = \bar{\lambda}^{m,\theta^m} = \bar{\lambda}_K^{m,\theta^m} = \frac{\gamma^{\theta^m} p^{\theta^m,m}}{\gamma^{\theta^m}(p^{\theta^m,m} + p^{\theta^m,f})} = \frac{p^{\theta^m,m}}{p^{\theta^m,m} + p^{\theta^m,f}} = \frac{\phi^N}{\phi^N + (1-\phi)^N}$, as asserted. Finally, we compare this quantity with its counterpart in Eq. (13):

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

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 $t(0), \ldots, t(K)$. *Q.E.D.*

Proof of Proposition 8. For part 1, the key step is analogous to the proof of Proposition 4, 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 Proposition 7, for $t < \min(t^\theta, t^{\theta^{\text{sym}}})$, both θ and θ^{sym} apply, and applying part 3(a) of Theorem A.1 to the relevant subsequence of $(\lambda_t)_{t\geq 0}$ as in the proof of Proposition 7, $\frac{\lambda_t^\theta}{\lambda_{t-1}^\theta} = \frac{\lambda_t^{\theta^{\text{sym}}}}{\lambda_{t-1}^{\theta^{\text{sym}}}}$, 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 8, $\lambda_t^\theta = \lambda_t^{\theta^{\text{sym}}}$ 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 8, and in addition $p^{\theta > p_m^{\theta^{\text{sym}}}, m}$. *Q.E.D.*