RESEARCH ARTICLE | FEBRUARY 26 2019

Mathematical model of gender bias and homophily in professional hierarchies FREE

Sara M. Clifton ; Kaitlin Hill ; Avinash J. Karamchandani ; Eric A. Autry; Patrick McMahon; Grace Sun

Chaos 29, 023135 (2019)
https://doi.org/10.1063/1.5066450

Chaos ARTICLE scitation.org/journal/cha

Mathematical model of gender bias and homophily in professional hierarchies

Cite as: Chaos **29**, 023135 (2019); doi: 10.1063/1.5066450
Submitted: 15 October 2018 · Accepted: 29 January 2019 ·
Published Online: 26 February 2019

View Online Export Citation CrossMark

Sara M. Clifton,[1,a)] Kaitlin Hill,[2] Avinash J. Karamchandani,[3] Eric A. Autry,[4] Patrick McMahon,[5] and Grace Sun[1]

AFFILIATIONS

[1] Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA
[2] School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455, USA
[3] Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, Illinois 60208, USA
[4] Department of Mathematics, Duke University, Durham, North Carolina 27708, USA
[5] Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA
[6] Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

[a)] **E-mail:** smc567@illinois.edu

ABSTRACT

Women have become better represented in business, academia, and government over time, yet a dearth of women at the highest levels of leadership remains. Sociologists have attributed the leaky progression of women through professional hierarchies to various cultural and psychological factors, such as self-segregation and bias. Here, we present a minimal mathematical model that reveals the relative role that bias and homophily (self-seeking) may play in the ascension of women through professional hierarchies. Unlike previous models, our novel model predicts that gender parity is not inevitable, and deliberate intervention may be required to achieve gender balance in several fields. To validate the model, we analyze a new database of gender fractionation over time for 16 professional hierarchies. We quantify the degree of homophily and bias in each professional hierarchy, and we propose specific interventions to achieve gender parity more quickly.

Women constitute approximately 50% of the population and have been an active part of the U.S. workforce for over half a century. Yet women continue to be poorly represented in leadership positions within business, government, medical, and academic hierarchies. As of 2018, less than 5% of Fortune 500 chief executive officers are female, 20% of the U.S. congress is female, and 34% of practicing physicians are female. The decreasing representation of women at increasing levels of power within hierarchical professions has been called the "leaky pipeline" effect, but the main cause of this phenomenon remains contentious. Using a mathematical model of gender dynamics within professional hierarchies and a new database of gender fractionation over time, we quantify the impact of the two major decision-makers in the ascension of people through hierarchies: those applying for promotion and those who grant promotion. The model is the first to demonstrate that intervention may be required to reach gender parity in some fields.

I. INTRODUCTION

A professional hierarchy is a field in which an employee enters at a designated low level and gradually moves up the ranks. For instance, large businesses have interns through CEOs, hospitals have residents through head physicians, and academic institutions have undergraduates through full professors. Over time, women have generally become better represented in many industries (e.g., Refs. 1–4), but women are still poorly represented at the highest levels of most professional hierarchies (e.g., Refs. 5–7). This has been called the "leaky pipeline" effect.

Countless factors have been proposed to explain this so-called "leaky pipeline" effect: family responsibilities,[8] different professional interests between the genders,[9–11] biological differences,[12] unconscious bias in the workplace,[13,14] laws restricting gender discrimination,[15,16] societal gender roles,[17–19] and other entrenched cultural or psychological factors. Many of these qualitative theories require an implicit assumption that men and women make fundamentally different decisions, either as a result of biological differences or social indoctrination.

Some quantitative models have attempted to study the ascension of women through certain fields without relying on intrinsic differences between the sexes. Shaw and Stanton[20] calculate the "inertia" of women through several academic hierarchies and find that gender differences play a diminishing role in promotion over time. Holman *et al.*[21] present the first quantitative model to our knowledge that attempts to predict the time required to reach gender parity in academic STEM fields, with estimates as high as several centuries in some disciplines. Their model assumes logistic growth to gender parity of the proportion of women in senior and junior academic roles (as estimated by last and first authorship on research papers, respectively). Although logistic growth and eventual gender parity is a reasonable assumption for their phenomenological model, we create a mechanistic model to examine the relative impact of two major sociological factors, homophilic (self-seeking) instincts and gender bias, on the progression of women through professional hierarchies. We find that gender parity is not guaranteed, and gender fractionation may never settle to an equilibrium.

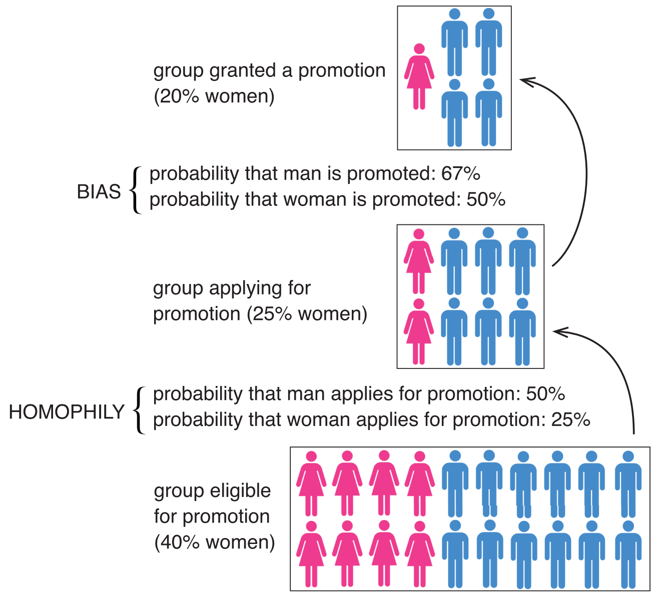

FIG. 1. Example of a potential decision process between two levels in a professional hierarchy.

II. MODEL

Broadly speaking, two classes of people influence the ascension of individuals through a professional hierarchy: people at lower levels choose to apply for higher positions, and people at higher levels choose to promote applicants into the next level. People at higher levels affect the promotion of individuals through their hiring biases, while the decisions to apply for promotion made by those at lower levels are affected by their own homophilic tendencies.

Women in hierarchical professions tend to be promoted more slowly than men, even when accounting for differing productivity and attrition, indicating that gender is a salient factor in the hiring process.[22–26] If gender is the determining factor when deciding between equally qualified candidates, we will say that a gender bias exists. We define gender bias as all conscious or unconscious decisions made by the employer during the hiring process that are affected by the gender of the applicant. For simplicity, we will assume that gender-based hiring bias is constant across all hierarchy levels (i.e., employers will uniformly reduce or enhance female candidates' relative chance of promotion at all levels).

Gendered differences in promotion also depend on gender differences in the applicant pool due to individuals self-selecting, consciously or unconsciously, whether or not to submit an application. When gender is a salient factor in deciding whether or not to submit an application, we will assume that such decisions are based largely on a homophilic instinct. In other words, when an individual considers whether or not to apply for a promotion, he or she looks at the demographics of those working at the above level and evaluates whether or not they "belong" in that higher level. While this assumption may seem simplistic, many studies show that people unconsciously self-segregate based on gender from a early age.[27–30] In fact, perceptions of gendered jobs perpetuate much of the occupational gender segregation we see today.[31–35]

With the goal of understanding the relative roles that bias and homophily (self-segregation) play in the ascension of women through professional hierarchies, we derive a minimal mathematical model that incorporates both forces. To introduce the model, we begin with a simple example.

A. Example

Consider the decision process that occurs during the transition between two levels in a professional hierarchy (Fig. 1). Suppose the lower level is 40% women, and gender is not a factor in eligibility for promotion; then the group eligible for promotion is also 40% women. If women are not well-represented in the higher level, then women may not feel as comfortable applying for promotion as men. To be clear, we do not suppose that women are intrinsically less likely to apply for promotion; rather, we assume that the gender demographics in the upper level affect both men's and women's feeling of belonging (homophily) in the upper level.

Say men are twice as likely to apply for promotion due to these homophilic instincts. Then, the applicant pool will shrink to 25% women. If no bias toward or against women exists in hiring, then 25% of those granted promotion will be women. However, if women are slightly less likely than men to be granted promotion due to bias, then the fraction of women

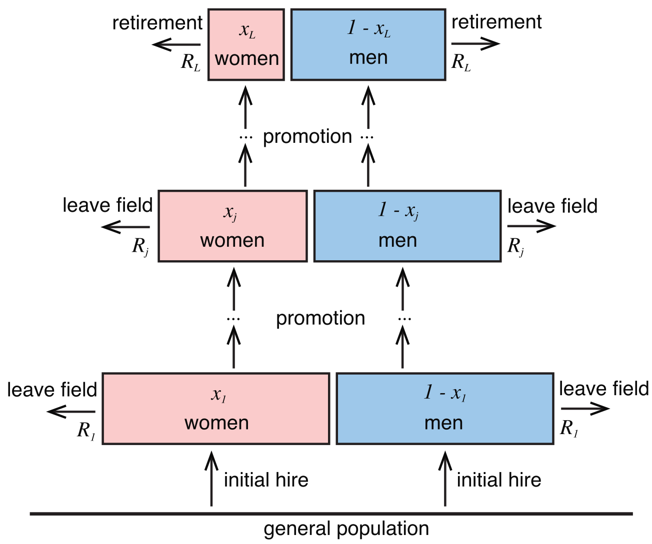

FIG. 2. Schematic of an L-level hierarchy. The jth level in the hierarchy has a certain fraction women x_j, and people retire or leave the field from each level at a rate R_j. The general population is assumed to be $1/2$ women at all times.

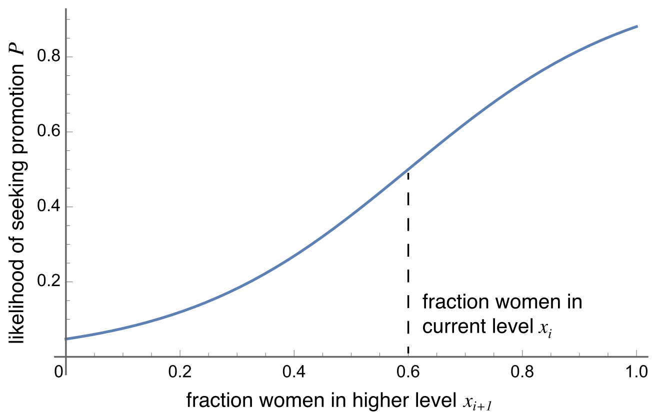

FIG. 3. An example of the probability that a woman seeks promotion, dependent on the demographics of the level to which she is applying. In this example, a woman is more likely to apply for promotion if there are more women in the level above her. The probability changes most rapidly around the demographic split she is most accustomed to, the gender split in her current position.

hired will shrink again. We assume that this decision process occurs between all levels in a professional hierarchy. The schematic in Fig. 2 is a visualization of a generic hierarchy.

B. Model derivation

We begin by assuming that the probability $P(u, v)$ of seeking promotion to the next level is a function of the fraction of people at the upper level who share the applicant's gender, u, and the fraction of like-gendered individuals in the applicant's current level, v. There exists a "one-third hypothesis" that supports the anecdotal evidence that an individual feels comfortable in a group environment when at least 30% of the members share the individual's demographic status.[36,37] To our knowledge, this hypothesis has not been rigorously tested in the real world, so we allow P to take a more flexible form. Specifically, we suppose that the threshold of comfort may depend on the environment in which a person currently resides. We also assume that the threshold does not delineate an instantaneous switch from 0% comfort to 100% comfort; instead, the comfort level may gradually change around that threshold. One simple function that captures this behavior is the sigmoid

$$P(u, v) = \frac{1}{1 + e^{-\lambda(u-v)}}, \tag{1}$$

where u is the fraction of like-gendered individuals in the level above, v is the fraction of like-gendered individuals in the current level, and λ is the strength of the homophilic tendency. This function need not be a literal probability because only the relative likelihood of applying for promotion is relevant. Because we choose to not include inherent gender difference in the model, we assume that this function applies to both men and women. See Fig. 3 for a sketch of this homophily function.

Given this probability $P(u, v)$ of seeking promotion, the fraction of women in the applicant pool is

$$f_0(u, v) = \frac{vP(u, v)}{vP(u, v) + (1 - v)P(1 - u, 1 - v)}, \tag{2}$$

where u is the fraction of women in the higher level and v is the fraction of women in the current level.

In addition to self-segregation dynamics, hiring bias toward or against women will change the proportion of female applicants who are promoted. We incorporate this constant bias b as the female fraction of those promoted if the applicant pool has an equal number of men and women. For instance, a bias b exceeding $\frac{1}{2}$ would imply that women are favored disproportionately, and a bias less than $\frac{1}{2}$ suggests that men are favored. The fraction of women promoted to the next level is then

$$f(u, v; b) = \frac{b\, v\, P(u, v)}{b\, v\, P(u, v) + (1 - b)(1 - v)P(1 - u, 1 - v)}. \tag{3}$$

This is not the only way to incorporate bias, but it is a simple way to ensure that bias does not leave vacancies or induce the promotion of those who have not applied. As an example, a naive choice to incorporate bias would be $f(u, v; b) = bf_0(u, v)$, where $b < 1$ indicates bias against women. However, this choice permits $f > 1$ if b or f_0 is sufficiently large.

Because professional hierarchies are frequently competitive, with each level smaller than the level below it, we assume that all vacancies will be filled. The vacancies are created by individuals who are promoted to the next level, those leaving the field at a particular level, or those retiring from the top level. The change in the number of women at each level, $x_j N_j$, is

$$\begin{aligned}\frac{d}{dt}(x_L N_L) &= R_L N_L f(x_L, x_{L-1}; b) - R_L N_L x_L,\\ \frac{d}{dt}(x_j N_j) &= \left(\sum_{k=j}^{L} R_k N_k\right) f(x_j, x_{j-1}; b) - R_j N_j x_j\\ &\quad - \left(\sum_{k=j+1}^{L} R_k N_k\right) f(x_{j+1}, x_j; b) \quad \text{for} \quad 1 < j < L,\\ \frac{d}{dt}(x_1 N_1) &= \left(\sum_{k=1}^{L} R_k N_k\right) f(x_1, \tfrac{1}{2}; b) - R_1 N_1 x_1\\ &\quad - \left(\sum_{k=2}^{L} R_k N_k\right) f(x_2, x_1; b),\end{aligned} \tag{4}$$

where L is the number of levels in the hierarchy, x_j is the fraction of people in level j who are women, N_j is the number of people in the jth level, R_j is the retirement/leave rate at the jth level, $\sum_{k=j+1}^{L} R_k N_k$ is the total number of retiring people above the jth level, b is the bias parameter, and $f(\cdot)$ is the fraction of people promoted to the next level who are women. Because it may not be intuitive that the change in the number of women at lower levels depends on the total number of retiring people above the level, we provide a simple example to illustrate this feature in the supplementary material.

We normalize system (4) by dividing each equation by the number of people retiring/leaving the level (i.e., $R_j N_j$)

$$\begin{aligned}\frac{1}{R_L}\frac{dx_L}{dt} &= \overbrace{f(x_L, x_{L-1}; b)}^{\text{promoted from lower level}} - \overbrace{x_L}^{\text{retire out of level}},\\ \frac{1}{R_j}\frac{dx_j}{dt} &= (1 + r_j) f(x_j, x_{j-1}; b) - x_j - r_j f(x_{j+1}, x_j; b) \quad \text{for } 1 < j < L,\\ \frac{1}{R_1}\frac{dx_1}{dt} &= \underbrace{(1 + r_1) f(x_1, \tfrac{1}{2}; b)}_{\text{hired from general pool}} - \underbrace{x_1}_{\text{leave field}} - \underbrace{r_1 f(x_2, x_1; b)}_{\text{promoted to next level}},\end{aligned} \tag{5}$$

where r_j is the ratio of the total retiring people above the jth level to the retiring people in the jth level. Algebraically, this ratio is $r_j = \sum_{k=j+1}^{L} R_k N_k / R_j N_j$. Note that this system can be condensed into one line by taking $r_L = 0$ and $x_0 = 1/2$. Refer to Table I for descriptions of the model variables and parameters.

TABLE I. Model variables and parameters for (5).

Variable	Meaning
x_j	Fraction of people in the jth level who are women
L	Number of levels in hierarchy
R_j	Retirement/leave rate at the jth level
N_j	Number of people in the jth level
r_j	Ratio of the total retiring people above the jth level to the retiring people in the jth level $\left(\sum_{k=j+1}^{L} R_k N_k / R_j N_j\right)$
$P(\cdot)$	Likelihood of seeking promotion
$f(\cdot)$	Fraction of people promoted to next level who are women
b	Bias toward or against women ($b = 1/2$ is no bias)
λ	Strength of homophilic tendency

C. Null model

Consider a null model with no hiring bias or homophily. In model (5), this would imply that bias $b = \frac{1}{2}$ and the likelihood of seeking promotion P is a constant ($\lambda = 0$). The model then reduces to the linear system

$$\frac{1}{R_j}\frac{dx_j}{dt} = (1 + r_j)(x_{j-1} - x_j) \quad \text{for } 1 \le j \le L. \tag{6}$$

The only steady state is $\{x_j^*\} = \{\frac{1}{2}\}$. The Jacobian of the system evaluated at this state yields all real, negative eigenvalues. Therefore, $\{x_j\} = \{\frac{1}{2}\}$ is a stable sink of the null model. In other words, without bias or homophily, each level in the hierarchy will directly converge to equal gender representation, as seen in the model by Holman *et al.*[21] The rate of convergence to parity for each level depends on the eigenvalues of the system: $\lambda_j = -R_j(1 + r_j)$, for $j = 1, \ldots, L$. The eigenvalues depend only on the level sizes N_j and leave rates R_j. The convergence time to parity for the whole system is then given by $1/\min_j\{R_j(1 + r_j)\}$, the characteristic timescale of the system.

See the supplementary material for a more complete discussion of analysis of the null model. Figure 4 shows convergence to gender parity in a hypothetical academic hierarchy.

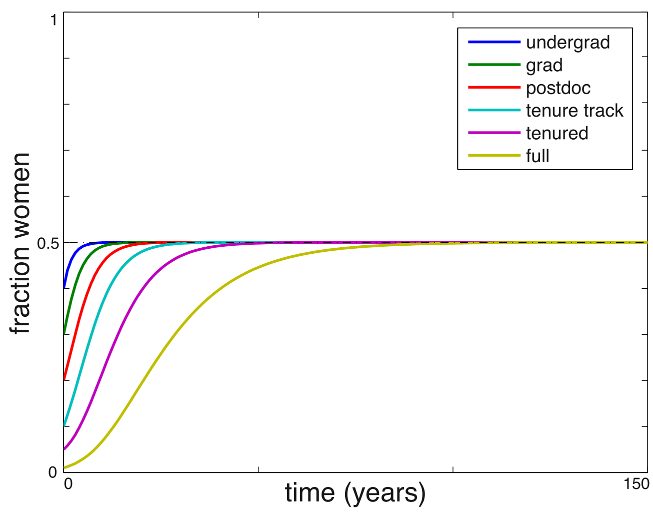

FIG. 4. Example of direct convergence to 50/50 gender split under null model (6). In this example, we consider a hypothetical academic hierarchy with six levels, $\{R_j\} = \{1/4, 1/5, 1/6, 1/7, 1/9, 1/15\}$, and $\{N_j\} = \{13, 8, 5, 3, 2, 1\}$.

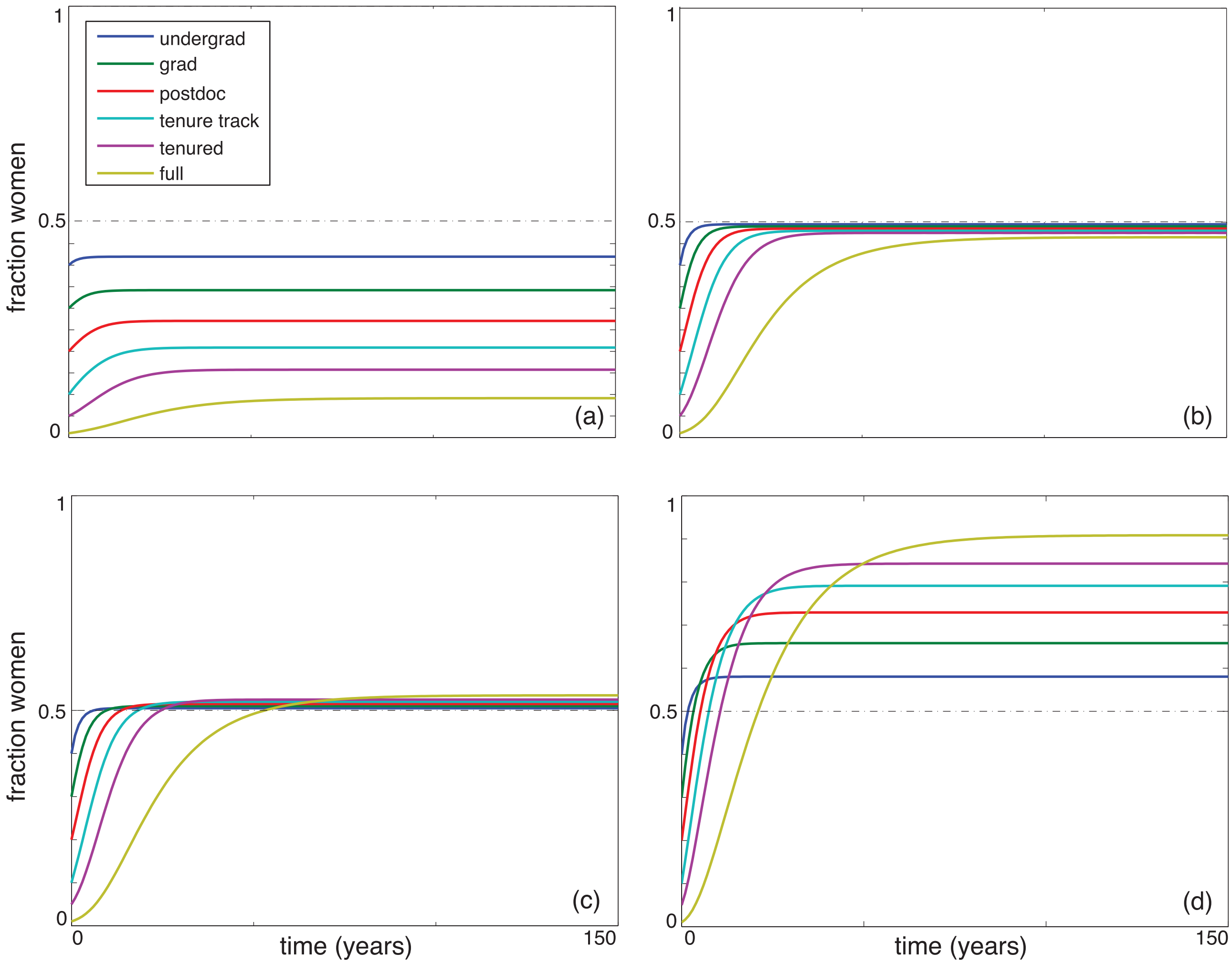

FIG. 5. Examples of transient behavior of 6-level homophily-free model (7). (a) For strong bias against women ($b = 0.35$), all levels directly converge to male majority, with the strongest majority in the highest levels of leadership. (b) For weak bias against women ($b = 0.49$), the fraction of women in each level directly converges to a value near 50/50, though there are still more men in each level. (c) For weak bias favoring women ($b = 0.51$), the fraction of women in each level directly converges to a value near 50/50, though there are more women in each level. (d) For strong bias favoring women ($b = 0.65$), all levels directly converge to female majority, with the strongest majority in the highest levels of leadership.

D. Homophily-free model

Now consider a model in which people do not use gender to decide whether to apply for a promotion (i.e., $\lambda = 0$), but employers are biased toward or against women (i.e., $b \neq \frac{1}{2}$). In this case, model (5) reduces to

$$\frac{1}{R_j}\frac{dx_j}{dt} = (1+r_j)\frac{bx_{j-1}}{bx_{j-1}+(1-b)(1-x_{j-1})} - x_j - r_j\frac{bx_j}{bx_j+(1-b)(1-x_j)} \quad \text{for } 1 \le j \le L. \tag{7}$$

As in null model (6), the homophily-free model has a single, attracting fixed point. The presence of bias, however, pushes the steady-state gender fractionation away from the gender parity. This effect is more extreme in higher levels than in lower ones. In particular, if the bias is against women ($b < \frac{1}{2}$),

$$x_L^* < \cdots < x_j^* < \cdots < x_1^* < \frac{1}{2}.$$

See the supplementary material for details, and see Fig. 5 for transient model behavior for a hypothetical academic hierarchy.

E. Bias-free model

Consider an alternative model in which people self-segregate by gender, but employers are not biased toward or

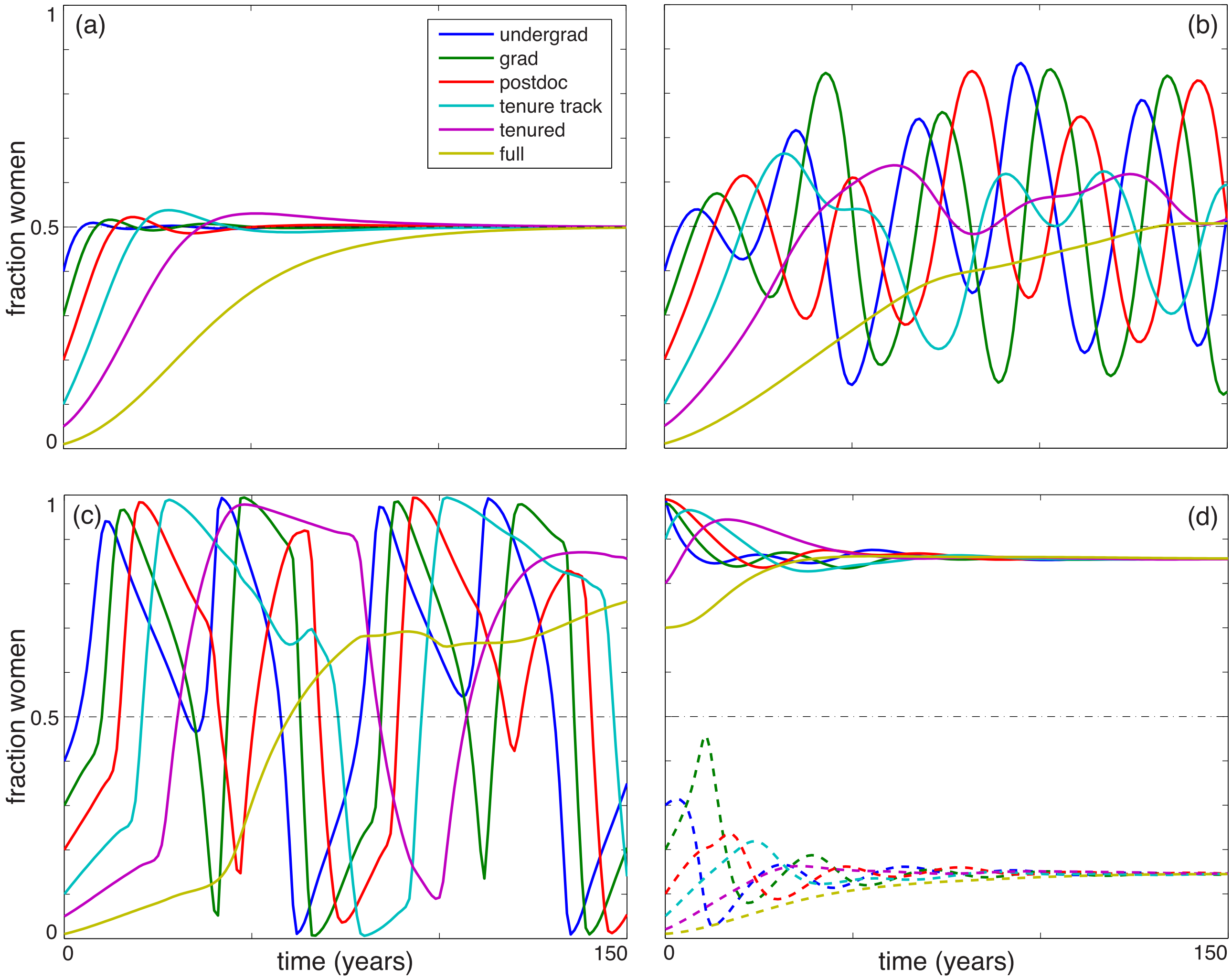

FIG. 6. Examples of transient behavior of a hypothetical academic 6-level bias-free model (8). (a) For mild homophily ($\lambda = 2$), all levels converge to gender equity after oscillating above and below a 50/50 split. (b) For stronger homophily ($\lambda = 3$), the fraction of women in each level oscillates about the 50/50 split without converging. (c) For yet stronger homophily ($\lambda = 4.5$), limit cycles appear to behave like those of a relaxation oscillator. (d) For strong homophily ($\lambda = 5$), each level equilibrates to nearly all women (solid lines) or nearly all men (dashed lines), depending on the initial condition. For all examples, $\{R_j\} = \{1/4, 1/5, 1/6, 1/7, 1/9, 1/15\}$, and $\{N_j\} = \{13, 8, 5, 3, 2, 1\}$.

against women (i.e., $b = \frac{1}{2}$). Then, model (5) reduces to

$$\frac{1}{R_j}\frac{dx_j}{dt} = (1 + r_j)f_0(x_j, x_{j-1}) - x_j - r_jf_0(x_{j+1}, x_j) \quad \text{for } 1 \leq j \leq L. \tag{8}$$

We observe three qualitatively different model behaviors for (8): for mild homophilic tendencies, the system converges to gender parity; for moderate homophily, the fraction of women oscillates in all levels; and for strong homophily, the system converges to either male or female dominance depending on the initial state. The emergence of oscillations in such a system may not seem intuitively obvious. We explain the onset of oscillations in the supplementary material.

Figure 6 shows the range of model behavior for a hypothetical academic hierarchy. See Fig. 7 for an example of a bifurcation diagram for the bias-free system. Although this diagram is representative of typical model behavior, the location of bifurcation points may shift as parameters vary.

For the parameter values listed in the caption of Fig. 6, we see that as homophily increases from a small value, a supercritical Hopf bifurcation occurs, which initiates the onset of stable oscillations in all hierarchy levels. Although these oscillations are not identical, they have the same period at steady state, as suggested by the transient behavior in Figs. 6(b) and 6(c). At $\lambda \approx 3.9$, the limit cycle in each hierarchy level undergoes a pitchfork bifurcation of limit cycles, after which

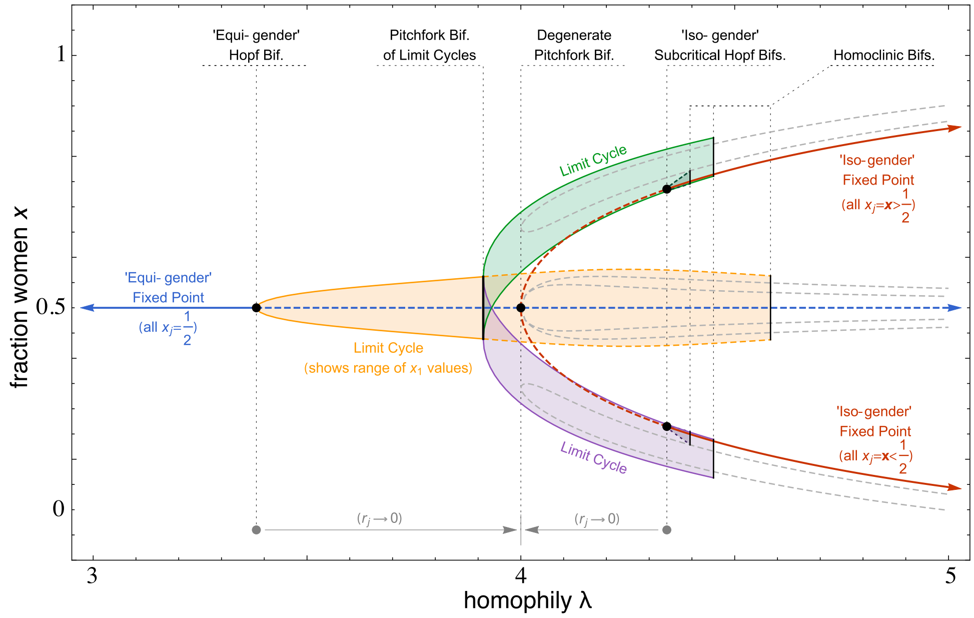

FIG. 7. Numerical bifurcation diagram for homophily parameter λ in a 3-level bias-free system. Solid lines are stable equilibria/cycles, dashed lines are unstable equilibria/cycles, black dots are bifurcations of equilibria, and black lines are bifurcations of limit cycles. All limit cycles show the gender fractionation for the lowest level, x_1. Generated using AUTO[38,39] with $N_1 = 70, N_2 = 2, N_3 = 1, R_1 = 1/4, R_2 = 1/5, R_3 = 1/6$. Convergence to a degenerate pitchfork bifurcation at $\lambda = 4$ as $r_j \to 0$ is shown in the supplementary material.

no stable equilibria at or steady oscillations about gender parity occur.

At $\lambda = 4$, a degenerate pitchfork bifurcation occurs for all parameter values. At this point, $2L + 1$ equilibria, several of which are unstable, emanate from the pitchfork as determined by a center manifold reduction. In Fig. 7, we focus on the equilibrium at gender parity and a pair of equilibria which eventually become stable, through subcritical Hopf bifurcations at $\lambda \approx 4.35$. All limit cycles eventually end at homoclinic bifurcations: the periodic orbit spends more and more time near a saddle point (not shown) as the period diverges.

As $r_j \to 0$ for each level, the Hopf bifurcations converge at the pitchfork bifurcation. In that limit the pitchfork has a greater degeneracy, producing 3^L equilbria. Loosely speaking, hierarchies with small r_j have very few people retiring relative to the number of people who would like to be promoted, making the hierarchies competitive. The limit $r_j \to 0$ is not realistic for any real-world hierarchy to our knowledge, but analysis near this limit aids numerical continuation; see the supplementary material for details.

F. Model with homophily and bias

Finally, we explore full model (5) with bias $b \neq \frac{1}{2}$ and homophily $\lambda \neq 0$. The long-term dynamics are similar to those of the bias-free model (8). For small homophily, regardless of initial state, the hierarchy tends toward a "biased" fractionation profile. For large homophily, the gender fraction polarizes with bistable equilibria at both large and small fractions of women at each level. Figures 8(a), 8(d), and 8(e) show examples of transient behavior at these high and low homophily values.

Figure 8 shows a slight perturbation of the system from the bias-free case, highlighting the degeneracy of the pitchfork bifurcation in Fig. 7; branch colors correspond with the colors of related branches in Fig. 7. Generically, for moderate levels of homophily, the limit cycles that emanate from the bifurcation "bend" in the direction of bias (e.g., for $b < \frac{1}{2}$, toward fewer women in each hierarchy level) as homophily increases through the supercritical Hopf bifurcation. The degenerate pitchfork bifurcation unfolds into several saddle-node bifurcations and a continuous fixed point curve. Similarly, the pitchfork bifurcation of limit cycles unfolds into a saddle-node bifurcation of limit cycles and a continuous limit cycle curve. As in Fig. 7, all limit cycles end in homoclinic bifurcations.

For lower values of bias b, the Hopf bifurcation from the equigender fixed point shifts along the branch of equilibria it emanates from, corresponding to a decrease in x_1 and an increase in homophily. At the same time, the length of the limit cycle branches emanating from the Hopf point decreases, and the Hopf point is eliminated in a Takens-Bogdanov bifurcation. For stronger bias ($b \approx 0.45$), long-term behavior manifests solely as equilibria, which includes the possibility of

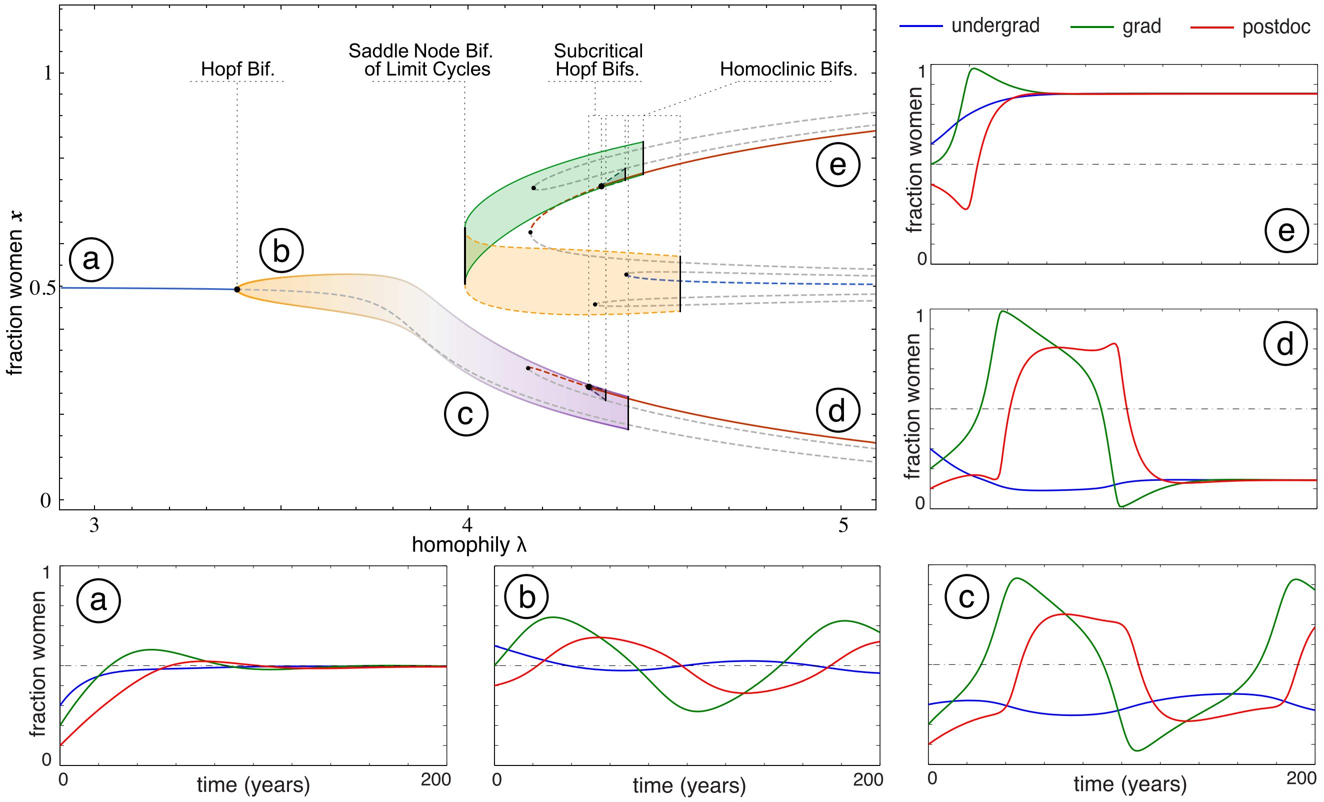

FIG. 8. Numerical bifurcation diagram for homophily parameter λ in a 3-level system with slight bias against women ($b = 0.499$). Solid lines are stable equilibria/cycles, dashed lines are unstable equilibria/cycles, black dots are bifurcations of equilibria, and black lines are bifurcations of limit cycles. All curves show the gender fractionation for the lowest level, x_1. Generated using AUTO[38,39] with $N_1 = 70$, $N_2 = 2$, $N_3 = 1$, $R_1 = 1/4$, $R_2 = 1/5$, $R_3 = 1/6$. Examples of transient behavior for several positions within the bifurcation diagram are on the margins: (a) $\lambda = 3$, (b) $\lambda = 3.5$, (c) $\lambda = 4$, (d) $\lambda = 5$ with lower initial condition, and (e) $\lambda = 5$ with higher initial condition.

decaying oscillations. Limit cycles are no longer possible. See the supplementary material for the co-dimension 2 bifurcation diagram, where both bias b and homophily λ are varied.

III. MODEL VALIDATION

With this simple model, we aim to extract useful information from real-world hierarchies without claiming to fully explain their dynamics. For instance, we wish to predict when (or if) fields will reach gender parity, what sociological or psychological factors may be the main drivers of gender fractionation dynamics, and what interventions may help various fields reach gender parity more quickly.

A. Data

We collect time series data on the fraction of women in each level of many professional hierarchies.[40–66] Although most studies of this nature have focused on academia,[20,21] the generality of our model allows us to examine a larger variety of hierarchies: medicine, law, politics, business, education, journalism, entertainment, and fine arts/music. Of the 23 hierarchy datasets we assembled, 16 are sufficiently comprehensive to attempt model fitting. Each dataset comprises the following components:

- A hierarchy structure (e.g., undergraduate → graduate → postdoctoral → assistant professor → associate professor → professor, in a typical academic hierarchy). In the real world, the hierarchical structure is not perfectly rigid, but we take the structure to be the "typical" route through the ranks. This structure determines the hierarchy size L and the ordering of levels in our model (5).
- The fraction of each level of the hierarchy that are women over time. We include datasets with at least a decade's worth of continuous yearly data for all levels. If there are missing years, we use linear interpolation to fill the gaps. Some datasets were available in a table, but others were extracted from graphical representations using WebPlotDigitizer.[67] This determines the exact $x_j(t)$ for a range of discrete times.
- The approximate relative sizes of each level. Although fields may grow (e.g., medicine) or shrink (e.g., journalism) over time, we find that the relative level sizes generally stay approximately the same. Where data on the relative level

sizes were not available, we made educated guesses. This information estimates N_j in our model if we normalize the top level to 1 ($N_L = 1$).
- The approximate yearly "leave" or "retirement" rates for each level. These statistics are not available for any hierarchies, to our knowledge. We made educated guesses for these parameters based on the expected amount of time spent in each level. For instance, the vast majority of undergraduate degrees are completed in approximately four years, and relatively few graduates continue on to doctoral study. Therefore, our initial estimate for the undergraduate leave rate is 0.25 (i.e., approximately a quarter of undergraduates leave college each year without moving up the academic hierarchy). We take these proxies to exit rates as estimates for R_j in our model.

All compiled data, including datasets not sufficient for model fitting, are available at Northwestern's ARCH repository: https://doi.org/10.21985/N2QF28.

B. Model fitting

We wish to fit the model to each dataset in order to quantify the degree of bias and homophily in each field; with this information, we may predict the long-term fraction of women in each level of the hierarchies without any intervention, and we can suggest targeted interventions to reach gender parity more quickly. Theoretically, distinguishing between bias and homophily in the data should be straightforward because the qualitative effects of each parameter are different. Bias is the only parameter that independently "separates" levels (i.e., bias causes the female fractionation x_j to differ among levels), while homophily is the only parameter that independently causes oscillations.

There are many possible ways to fit the model to each dataset. One qualitative way to measure the degree of bias and homophily in each dataset is to look for separation between levels and indications of oscillations. Roughly speaking, datasets with strong bias either toward or against women will have large changes in the proportion of women as one ascends the hierarchy [e.g., see Figs. 5(a) and 5(d)].

On the other hand, datasets with weak bias and moderate homophily will show signs of oscillations in each level [e.g., see Figs. 6(b) and 6(c)], although real datasets may not include enough time points to resolve a full period of the oscillations. Datasets with weak bias and strong homophily will appear male- or female-dominated without much separation between levels [e.g., see Fig. 6(d)]. If both bias and homophily are strong, then the impact of each phenomenon will be difficult to deduce visually (see the supplementary material for phase diagram), and quantitative methods will be needed.

As a quantitative attempt at fitting, we perform a global minimization of error between the model and data. We first find a best fit of the model to each dataset by minimizing the sum of squared error between the model gender fractionation $\hat{x}_j$ and the data x_j over time using the Nelder-Mead minimization algorithm.[68] The fitting parameters are b, λ, R_j, N_j, and the initial conditions. We include R_j and N_j as fitting parameters because we do not have exact values for these parameters, but we heuristically verify that the model fit does not select values far from our initial guesses. The initial condition is a fitting parameter to ensure that the first data point does not contribute more weight to the fitting process than the subsequent data points in the time series.

We seed the Nelder-Mead algorithm with 20 initial guesses for the fitting parameters b and λ, selected uniformly from $b \in (0.2, 0.8)$ and $\lambda \in (1.5, 6.5)$. All other parameter guesses are taken to be our best estimates from available data. After finding the best fit parameters $(\tilde{b}, \tilde{\lambda})$ from among the 20 seeded searches, we run a second search in the parameter space near the best fit. In this next step, we seed the Nelder-Mead algorithm with 10 new initial guesses for b and λ, selected from normal distributions $b \in \mathcal{N}(\tilde{b}, 0.05)$ and $\lambda \in \mathcal{N}(\tilde{\lambda}, 0.1)$. We take, as our final fit, the best fit parameters after this second search. See the supplementary material for a visual representation of this algorithm.

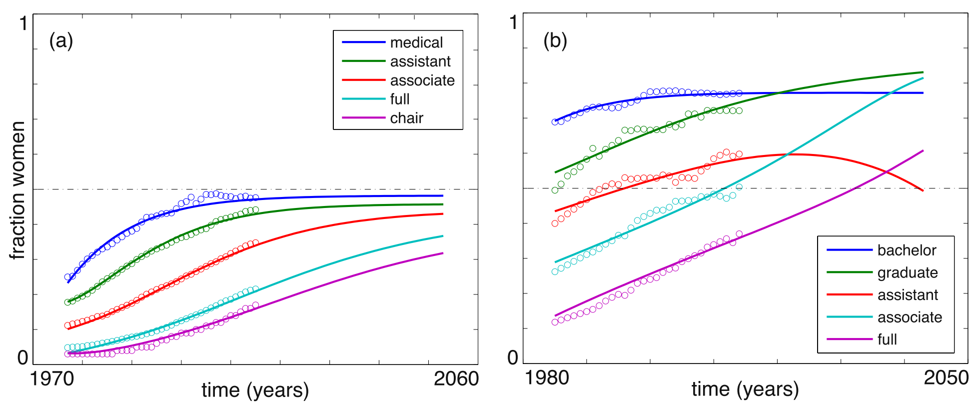

FIG. 9. Model fit to data from (a) clinical academic medicine[46–48] and (b) academic psychology.[41,42,45]

Chaos 29, 023135 (2019); doi: 10.1063/1.5066450

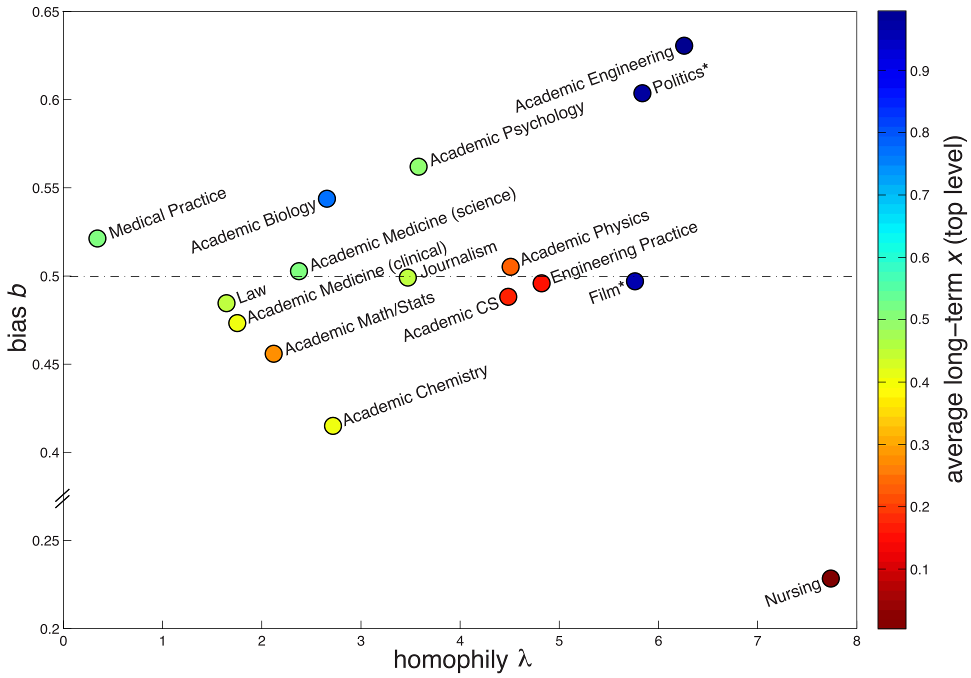

FIG. 10. Bias and homophily best fit parameters for each hierarchy. Colors indicate the predicted long-term (equilibrium) female fractionation in the highest level of leadership; if the hierarchy is not predicted to reach equilibrium, then a time average over the limit cycle was taken. *May not be a strict hierarchy: although producers hire directors, producers do not typically "promote" directors to producer positions. Likewise for politics.

We present the best fits from two representative hierarchies in Fig. 9. Best fit parameters $\hat{b}$ and $\hat{\lambda}$ from all datasets are shown in Fig. 10. See the supplementary material for fit parameters and additional model predictions for all datasets.

To address the concern of possible overfitting, we verify that the ratio of data points to fitting parameters is large. For each dataset, there are $3L + 1$ fitting parameters and LT data points, where L is the number of levels and T is the number of years in the dataset. The datasets with the fewest number of levels available and fewest years available should prompt greatest concern regarding overfitting. Among our datasets, the smallest ratio of parameters to data points was for the journalism hierarchy, which had 51 data points and 10 parameters. The typical ratio of data points to parameters was about 10:1.

Because our parameter search algorithm is not guaranteed to find the absolute minimum error between the model and data, we verify that our model results are not excessively sensitive to changes in our fitting procedure. To illustrate, we seed our algorithm's random number generator with ten different seeds and verify that the variation in predicted average gender fractionation is small. We select the two most concerning datasets for this computationally intensive test: (1) journalism, due to its risk for overfitting, and (2) academic engineering, due to its unpredictable fitting results during early tests (see the supplementary material).

IV. DISCUSSION

The presented model vastly simplifies the process by which people choose to advance their careers, yet we may exploit the model to extract useful predictions and suggestions for interventions to reach gender parity. By fitting the model to data from over a dozen professional hierarchies, we may predict the time required to reach gender parity if there are no cultural or policy shifts within the fields. Unlike the model by Holman *et al.*,[21] we predict that many fields may never reach gender parity without intervention (see the supplementary material). For instance, fields that indicate especially strong homophily (e.g., engineering and nursing) are expected to become male- or female-dominated. Fields with apparently strong bias against women (e.g., academic chemistry, math, and computer science) are predicted to never reach sustained gender parity, at least in the highest levels of leadership.

Fields with bias near 1/2 and weak homophily (e.g., medicine and law) are predicted to eventually reach gender parity as fast as inertia allows, as modeled by Shaw and Stanton.[20] Effective affirmative action programs could artificially speed the process, but resources may be better spent in fields where gender parity is not inevitable. One benefit of our modeling approach is that we can extract the relative impact of two major decision-makers in a professional hierarchy: those who apply for promotion and those who grant promotion. For fields with strong bias against women ($b < 1/2$), the decision-makers that should be targeted are hiring committees. For instance, hiring committees could be trained in unconscious bias, or policies could mandate that the number of promotions offered to women match the applicant pool. For fields with strong homophily, the decision-makers that should be targeted are women eligible for promotion. Knowing that fewer women than are eligible are applying for promotion in male-dominated fields, hiring committees could actively recruit women to apply for promotion or make the under-represented gender more visible within the field.

A. Limitations

Of course, the predictions and interventions suggested by this simple model are subject to limitations. We assume that hierarchical structures remain constant over time, but this is not always the case. For instance, some fields that now require a college degree were once accessible to those with a high school education. We also assume that individuals must pass through each level linearly, but many academic fields may or may not include a postdoc, and political or business leaders may come from outside their field entirely.

To avoid overfitting, we assume that bias and homophily are constant both across time and across the hierarchy structure. Naturally, the cultures and policies that shape these sociological properties are not constant; perhaps bias against women has diminished over time, but maybe bias is stronger at higher levels of leadership. Also, gender may be more salient to a young person deciding on a major than on an associate professor up for promotion. Therefore, we think of the fitting parameters $\hat{b}$ and $\hat{\lambda}$ as an average bias and homophily over time and the hierarchy structure.

Finally, we have ignored the different decisions that men and women may make. Our model assumes that men and women on hiring committees are equally biased against a certain gender, that gender is equally salient to men and women, and that men and women are equally qualified for advancement. A more sophisticated model may break the symmetry between men and women.

B. Future steps

Allowing bias and homophily to change over time and across the hierarchy structure is a natural model extension. In addition to making the model more realistic, it would also permit interventions to be incorporated directly into the model. If the effect of an intervention is to change bias and/or homophily, then the model could serve as the basis of a control problem to find an optimal time-dependent intervention.

Due to the generality of the model, it could also be extended to study the progression of under-represented minorities through professional hierarchies. A few complications are introduced in this case: our model assumes that the gender distribution of the general population is constant in both space and time, but for racial minorities this is not true. Also, data collection may prove to be more complicated due to the evolving and sometimes overlapping definitions of various racial and ethnic groups.

Finally, the model could be generalized to include a spectrum of gender identities, income levels, or socioeconomic privilege. Two major challenges are introduced with this model extension. First, the current system of L ordinary differential equations may become a system of L partial integro-differential equations, which will make model analysis more difficult. Second, data required to validate such a model will be more challenging to obtain.

V. CONCLUSION

We have developed a simple model of the progression of people through professional hierarchies like academia, medicine, and business. The model assumes that gender is a salient factor in both the decision to apply for promotion and the decision to grant promotion, but that men and women do not make fundamentally different decisions. Unlike previous models of the phenomenon, our model predicts that gender parity is not inevitable in many fields. Without intervention, a few fields may even become male- or female-dominated in the long term.

By fitting our model to available data, we extract the relative impact of the major decision-makers in the progression of women through 16 professional hierarchies. In some fields, like academic chemistry, bias of promotion and hiring committees may be the dominant reason that women are poorly represented. In other fields, like engineering, women not applying for promotion may be the dominant reason for the so-called leaky pipeline. With this information, we may suggest effective interventions to reach gender parity.

SUPPLEMENTARY MATERIAL

See the supplementary material for additional discussion, analysis, and figures.

ACKNOWLEDGMENTS

The authors wish to thank Danny Abrams, Yuxin Chen, Stephanie Ger, Joseph Johnson, and Rebecca Menssen for valuable conversations during the model development stage. Thanks are also due to Elizabeth Field, Alan Zhou, and the Illinois Geometry Lab for contributions to and support of model analysis and data collection. The authors additionally thank Chad Topaz for offering comments that greatly improved the manuscript.

The authors also wish to thank João Moreira (Amaral Lab, Northwestern University), Peter Buerhaus and Dave Auerbach (Center for Interdisciplinary Health Workforce Studies, Montana State University), Roxanna Edwards (Bureau of Labor Statistics), and Karen Stamm (Center for Workforce Studies, American Psychological Association) for sharing unpublished data.

This work was funded in part by the National Science Foundation Graduate Research Fellowship No. DGE-1324585 and Mathways Grant No. DMS-1449269 (S.M.C.), Royal E.

Cabell Terminal Year Fellowship (K.H. and E.A.A.), and the National Science Foundation Research Training Grant No. DMS-1547394 (A.J.K.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

All data (Excel .xlsx file) and software (Matlab .m files and XPPAUT .ode files) are publicly available from the Northwestern ARCH repository at https://doi.org/10.21985/N2QF28.

REFERENCES

[1] L. Luckenbill-Edds, "The educational pipeline for women in biology: No longer leaking?," AIBS Bull. **52**(6), 513–521 (2002).
[2] J. S. Kaye and A. C. Reddy, "The progress of women lawyers at big firms: Steadied or simply studied," Fordham L. Rev. **76**, 1941 (2008).
[3] S. Terjesen, R. Sealy, and V. Singh, "Women directors on corporate boards: A review and research agenda," Corp. Govern. Int. Rev. **17**(3), 320–337 (2009).
[4] K. A. Farrell and P. L. Hersch, "Additions to corporate boards: The effect of gender," J. Corp. Fin. **11**(1), 85–106 (2005).
[5] D. J. Nelson and D. C. Rogers, *A National Analysis of Diversity in Science and Engineering Faculties at Research Universities* (National Organization for Women, Washington DC, 2003).
[6] P. L. Carr, C. M. Gunn, S. A. Kaplan, A. Raj, and K. M. Freund, "Inadequate progress for women in academic medicine: Findings from the national faculty study," J. Women. Health **24**(3), 190–199 (2015).
[7] M. Shapiro, D. Grossman, S. Carter, K. Martin, P. Deyton, and D. Hammer, "Middle school girls and the leaky pipeline to leadership: An examination of how socialized gendered roles influences the college and career aspirations of girls is shared as well as the role of middle level professionals in disrupting the influence of social gendered messages and stigmas," Middle Sch. J. **46**(5), 3–13 (2015).
[8] A. H. Eagly and L. L. Carli, "Women and the labyrinth of leadership," Contemp. Issues Leadersh. **16**(1), 147–162 (2012).
[9] A. M. Konrad, J. E. Ritchie, Jr, P. Lieb, and E. Corrigall, "Sex differences and similarities in job attribute preferences: A meta-analysis.," Psychol. Bull. **126**(4), 593 (2000).
[10] S. J. Ceci and W. M. Williams, "Sex differences in math-intensive fields," Curr. Dir. Psychol. Sci. **19**(5), 275–279 (2010).
[11] S. M. VanAnders, "Why the academic pipeline leaks: Fewer men than women perceive barriers to becoming professors," Sex Roles **51**(9-10), 511–521 (2004).
[12] P. Sapienza, L. Zingales, and D. Maestripieri, "Gender differences in financial risk aversion and career choices are affected by testosterone," Proc. Natl. Acad. Sci. **106**(36), 15268–15273 (2009). pnas–0907352106.
[13] D. M. Easterly and C. S. Ricard, "Conscious efforts to end unconscious bias: Why women leave academic research," J. Res. Admin. **42**(1), 61–73 (2011).
[14] A. J. Lee, "Unconscious bias theory in employment discrimination litigation," Harv. CR-CLL Rev. **40**, 481–503 (2005).
[15] T. I. Chacko, "Women and equal employment opportunity: Some unintended effects," J. Appl. Psychol. **67**(1), 119 (1982).
[16] G. P. Sape and T. J. Hart, "Title vii reconsidered: The equal employment opportunity act of 1972," Geo. Wash. L. Rev. **40**, 824 (1971).
[17] D. M. Britton, "The epistemology of the gendered organization," Gend. Soc. **14**(3), 418–434 (2000).
[18] R. M. Blackburn, J. Siltanen, and J. Jarman, "The measurement of occupational gender segregation: Current problems and a new approach," J. R. Stat. Soc. Ser. A **158**(2), 319–331 (1995).
[19] M. Dean, R. Gill, and J. B. Barbour, "Work space, gendered occupations, and the organization of health: Redesigning emergency department communication," in *Organizations, Communication, and Health* (Routledge, 2015), pp. 125–142.
[20] A. K. Shaw and D. E. Stanton, "Leaks in the pipeline: Separating demographic inertia from ongoing gender differences in academia," Proc. R. Soc. Lond. B **279**(1743), 3736–3741 (2012).
[21] L. Holman, D. Stuart-Fox, and C. E. Hauser, "The gender gap in science: How long until women are equally represented?," PLoS Biol. **16**(4), e2004956 (2018).
[22] B. J. Tesch, H. M. Wood, A. L. Helwig, and A. B. Nattinger, "Promotion of women physicians in academic medicine: Glass ceiling or sticky floor?," J. Am. Med. Assoc. **273**(13), 1022–1025 (1995).
[23] M. E. Heilman, "Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder," J. Soc. Issues **57**(4), 657–674 (2001).
[24] A. H. Eagly and L. L. Carli, "Women and the labyrinth of leadership," Harv. Bus. Rev. **85**(9), 62 (2007).
[25] S. Kumra, and S. Vinnicombe, "A study of the promotion to partner process in a professional services firm: How women are disadvantaged," Br. J. Manage. **19**(s1), S65–S74 (2008).
[26] J. A. Winkler, "Faculty reappointment, tenure, and promotion: Barriers for women," Prof. Geogr. **52**(4), 737–750 (2000).
[27] E. E. Maccoby and C. N. Jacklin, "Gender segregation in childhood," Adv. Child. Dev. Behav. **20**, 239–287 (1987).
[28] G. M. Alexander and M. Hines, "Gender labels and play styles: Their relative contribution to children's selection of playmates," Child. Dev. **65**(3), 869–879 (1994).
[29] L. C. Moller and L. A. Serbin, "Antecedents of toddler gender segregation: Cognitive consonance, gender-typed toy preferences and behavioral compatibility," Sex Roles. **35**(7), 445–460 (1996).
[30] K. K. Powlishta, "Gender segregation among children: Understanding the 'cootie phenomenon'," Young Child. **50**(4), 61–69 (1995).
[31] L. Miller, F. Neathey, E. Pollard, and D. Hill, "Occupational segregation, gender gaps and skill gaps," in *Institute of Employment Studies EOC Working Paper* (2004), Vol. 15.
[32] J. Pan, "Gender segregation in occupations: The role of tipping and social interactions," J. Labor. Econ. **33**(2), 365–408 (2015).
[33] J. Ludsteck, "The impact of segregation and sorting on the gender wage gap: Evidence from german linked longitudinal employer-employee data," Ind. Labor. Relat. Rev. **67**(2), 362–394 (2014).
[34] O. Alonso-Villar, C. Del Rio, and C. Gradín, "The extent of occupational segregation in the united states: Differences by race, ethnicity, and gender," Ind. Relat. J. Econ. Soc. **51**(2), 179–212 (2012).
[35] K. A. Bender, S. M. Donohue, and J. S. Heywood, "Job satisfaction and gender segregation," Oxf. Econ. Pap. **57**(3), 479–496 (2005).
[36] H. O. Engelmann, "Communication to the editor," Am. Sociol. **2**(4), 21 (1967).
[37] K. Srikantan, "A curious mathematical property," Am. Sociol. **3**(2), 154–155 (1968).
[38] E. J. Doedel, T. F. Fairgrieve, B. Sandstede, A. R. Champneys, Y. A. Kuznetsov, and X. Wang, "Auto-07p: Continuation and bifurcation software for ordinary differential equations," 2007.
[39] B. Ermentrout, *Simulating, Analyzing, and Animating Dynamical Systems: A Guide to XPPAUT for Researchers and Students* (SIAM, 2002), Vol. 14.
[40] National Science Foundation, National Center for Science and Engineering Statistics, "Women, minorities, and persons with disabilities in science and engineering," 2017, see https://www.nsf.gov/statistics/2017/nsf17310/data.cfm.
[41] National Science Foundation, "Survey of graduate students and postdoctorates in science and engineering (gss)," 2018, see https://ncsesdata.nsf.gov/ids/gss.
[42] Institute of Education Science, National Center for Education Statistics, "Integrated postsecondary education data system (ipeds)," 2018, see https://nces.ed.gov/ipeds/use-the-data.
[43] American Chemical Society, Chemical & Engineering News, "Women faculty by institution," 2015, see https://cen.acs.org/articles/92/i14/Women-Faculty-By-Institution.html.
[44] The National Academies of Science, Engineering, and Medicine, "Seeking solutions: Maximizing american talent by advancing women of color in academia," 2013, see https://www.nap.edu/read/18556/chapter/24.
[45] American Psychological Association, "Graduate study in psychology: Student demographics," 2017, see http://www.apa.org/education/grad/survey-data/2017-student-demographics.aspx.

[46]Henry J Kaiser Family Foundation, "Distribution of medical school graduates by gender," 2017, see https://www.kff.org/other/state-indicator/medical-school-graduates-by-gender/?dataView=1&activeTab=graph&currentTimeframe=0&startTimeframe=15&selectedDistributions=female&selectedRows=%7B%22wrapups%22:%7B%22united-states%22:%7B%7D%7D%7D&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D.
[47]Association of American Medical Colleges, "Faculty diversity in U.S. medical schools: Progress and gaps coexist," 2016, see https://www.aamc.org/data/aib/474174/december2016facultydiversityinu.s.medicalschoolsprogressandgaps.html.
[48]American Board of Internal Medicine, "Resident and fellow workforce data," 2017, see https://www.abim.org/about/statistics-data/resident-fellow-workforce-data.aspx.
[49]S. E. Brotherton and S. I. Etzel, "Graduate medical education, 2016-2017," J. Am. Med. Assoc. **318**(23), 2368–2387 (2017).
[50]S. E. Brotherton and S. I. Etzel, "Graduate medical education, 2015-2016," J. Am. Med. Assoc. **316**(21), 2291–2310 (2016).
[51]S. E. Brotherton and S. I. Etzel, "Graduate medical education, 2014-2015," J. Am. Med. Assoc. **314**(22), 2436–2454 (2015).
[52]C. M. Topaz and S. Sen, "Gender representation on journal editorial boards in the mathematical sciences," PLoS ONE **11**(8), e0161357 (2016).
[53]M. Wasserman, X. Zeng, and L. Amaral, "U.S. movies with gender-disambiguated actors, directors, and producers," see https://figshare.com/articles/U˙S˙movies˙with˙gender-disambiguated˙acto rs˙directors˙and˙producers/4967876/1.
[54]American Society of News Editors, "Employment of men and women by job category," 2015, see https://www.asne.org/content.asp?contentid=144.
[55]J. M. Blount, *Destined to Rule the Schools: Women and the Superintendency, 1873-1995* (SUNY Press, 1998).
[56]J. A. Dana and D. M. Bourisaw, *Women in the superintendency: Discarded leadership* (Rowman & Littlefield Education, 2006).
[57]Catalyst: Workplaces that Work for Women, "Statistical overview of women in the workforce," 2018, see https://www.catalyst.org/knowledge/statistical-overview-women-workforce.
[58]American Bar Association, "Enrollment and degrees awarded," 2013, see https://www.americanbar.org/content/dam/aba/administrative/legal˙education˙and˙admissions˙to˙the˙bar/statistics/enrollment˙degrees˙awarded.authcheckd am.pdf.
[59]National Association for Law Placement, "Women and minorities at law firms," 2018, see https://www.nalp.org/0218research.
[60]National Association for Women Lawyers, "Survey on promotion and retention of women in law firms," 2017, see https://www.nawl.org/page/2017.
[61]Rutgers University, "Center for american women and politics," 2018, see http://www.cawp.rutgers.edu/current-numbers.
[62]Pew Research Center, "The data on women leaders," 2018, see http://www.pewsocialtrends.org/fact-sheet/the-data-on-women-leaders/.
[63]Montana State University, Center for Interdisciplinary Health Workforce Studies, "Current trends of men in nursing," 2017, see http://healthworkforcestudies.com/publications-data/data˙brief˙update˙c urrent˙trends˙of˙men˙in˙nursing.html.
[64]U.S. Census Bureau, Men in Nursing Occupations, "American community survey highlight report," 2013, see https://www.census.gov/content/dam/Census/library/working-papers/2013/acs/2013˙Landivar˙02.pdf.
[65]U.S. Department of Health and Human Services, National Center for Health Workforce Analysis (NCHWA), "Nursing workforce survey data," 2008, see https://data.hrsa.gov/topics/health-workforce/nursing-workforce-survey-data?tab=RegisteredNurses.
[66]League of American Orchestras, "Racial/ethnic and gender diversity in the orchestra field," 2016, see http://americanorchestras.org/images/stories/diversity/Racial-Ethnic-an d-Gender-Diversity-in-the-Orchestra-Field-Final-92116.pdf.
[67]A. Rohatgi, "Webplotdigitizer," 2018.
[68]J. A. Nelder and R. Mead, "A simplex method for function minimization," Comput. J. **7**(4), 308–313 (1965).