Underrepresentation of Women CEOs

Li He and Toni M. Whited*

January 2023

ABSTRACT

Why do so few women become CEOs? We answer the question by estimating a dynamic model of the CEO gender decision that contains three sources of gender-based differences: unobserved productivity, search frictions that reflect limited female labor supply, and employer disutility arising from discrimination against women. We find that the most crucial factor in explaining the apparent glass ceiling is the shortage of suitable candidates. Net of the availability of suitable candidates, boards are in favor of hiring female CEOs.

*He is from Rotterdam School of Management, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands. he@rsm.nl. Whited is from Ross School of Business and the NBER, University of Michigan, 701 Tappan St., Ann Arbor, MI 48109. twhited@umich.edu.

1. Introduction

Women are underrepresented in U.S. companies, holding only 33 or 6.6% of CEO positions at S&P 500 companies.[1] Why do so few women become CEOs? Three main explanations have been proposed for this phenomenon. First, there might be unobserved differences in productivity, human capital, or intrinsic traits between male and female CEOs. For example, Schubert, Brown, Gysler, and Brachinger (1999) suggests that women have high risk aversion and that low risk aversion might be necessary for a firm's success. The second explanation is based on career paths and posits that limited female labor supply might attenuate women's representation in management (Adams and Kirchmaier 2012). In this case, family considerations play an important role in constraining the career advancement of women (Bertrand, Goldin, and Katz 2010), and a limited pool of qualified women increases the cost to a firm of searching for a CEO. The third explanation is based on simple distaste, in which corporate and financial sectors might be particularly prone to discriminate against women. Shareholders and boards might be willing to sacrifice profits to avoid promoting women into top positions (Wolfers 2006).

Despite substantial media attention and academic discussion, the relative quantitative importance of these explanations remains unclear. The goal of this paper is to gauge the contribution of each of these factors in explaining the apparent glass ceiling.

However, disentangling and evaluating these channels present several challenges. First, it is difficult to measure variables that could render credence to each explanation, as gender differences in productivity, CEO searching costs, and distaste for hiring female CEOs are all imperfectly observable to econometricians. Second, the decision to hire a female CEO is determined endogenously along with firm characteristics, firm performance, and CEO personal traits. Third, it is challenging to measure the relative importance of each factor without clear benchmarks.

[1] Catalyst, Women CEOs of the S&P 500 (November 5, 2022).

To address these challenges, we develop a dynamic model of the CEO gender decision that contains three sources of gender-based differences: productivity, search frictions that reflect limited female labor supply, and disutility arising from discrimination against women. In the model, boards choose the gender of the CEO to maximize their own utilities, which consist of profits generated by the CEO, the payment of search costs, and additional disutility in the case of discrimination against women. Using the model as the backbone, we estimate the gender differences by fitting our model to the observed dynamics of CEO gender decisions.

We find that genuine board disutility from hiring women as CEOs plays a small role in the gender gap. In addition, we estimate that actual productivity differences are tiny. The main factor that drives the low fraction of female CEOs is the small number of women in the applicant pools. This paucity of women implies that search costs are high, and these high search costs are the dominant factor behind the gender gap in CEOs.

Our work is related to the few empirical studies that analyze the gender gap in compensation and promotion rates among CEOs. Bertrand and Hallock (2001) find that the gender gap in executive pay disappears after controlling for firm size, occupation, and age. Bell (2005) finds that the promotion chances of female executives are significantly higher in women-led firms. Gayle, Golan, and Miller (2012) and Smith, Smith, and Verner (2013) track the career paths of executives and find that the gender differential of becoming a CEO is associated with initial ranks of the job hierarchy, exit rates, and the area of specialization as a top executive. Our paper adds to this work by quantifying the importance of different reasons behind the gender gap.

Our paper is also related to the empirical literature on discrimination, which attempts to document the effects of discrimination on differences in economic outcomes between groups. The studies in this literature typically measure differences in economic outcomes between blacks and whites, or between women and men, that remain after statistically controlling for observable characteristics of workers. See, for example, Guryan and Charles (2013) for a survey.

Methodologically, our paper is related to Bowlus and Eckstein (2002), which estimates an equilibrium search model using aggregate data on wages and unemployment. The model features two types of workers who differ in three ways: unobserved productivities, search intensities, and discrimination. Our paper differs in two ways. Our analysis features a discrete-choice partial equilibrium model in which the CEO gender decision is endogenously determined by the board, and our estimation uses microeconometric panel data at the individual firm level.

2. Model

In this section, we present a parsimonious dynamic model of the CEO gender decision in discrete time. A firm is characterized by a risk-neutral board that makes a choice of whether to appoint a female or male CEO. The decision is denoted by d_t, which is equal to one if the appointed CEO is female and zero otherwise.

In each period, the firm's profit π_t is

$$\pi_t = (a - \lambda d_t)k_t + \epsilon_t(d_t), \tag{1}$$

where k_t is firm size at time t, a is a male CEO's productivity, and λ is the productivity difference between the male and female CEOs. Both a and λ are parameters that are constant over time. The multiplicative specification between CEO ability and firm size implies that marginal CEO productivity rises with firm size. We opt for this assumption as opposed to an additive model in which effort and ability have a constant effect on firm value for two reasons. First, the multiplicative production function is commonly used in macroeconomic models and makes intuitive sense because, as advocated by Edmans, Gabaix, and Landier (2009), "a majority of CEO moves are rolled out across the entire firm and therefore have a greater influence in a larger firm." Second, as discussed in more detail below, this assumption aids identification of the average difference in gender productivity. If there exists any productivity difference between male and female CEOs, then this difference is amplified by firm size.

The component ϵ_t represents the quality of the match between the CEO and the firm. It is the board's choice-specific private information, which is an independent and identically distributed extreme value type I random variable, with zero mean and dispersion σ_ϵ. This distributional assumption is standard in dynamic discrete choice frameworks (Arcidiacono and Ellickson 2011).

The characterization of firm profitability has several implications. First, CEOs' productivity differs in gender only and is constant over time. Thus, it captures a gender fixed effect in firm productivity. Second, firm profit is net of CEO pay as in Taylor (2010). We justify this assumption given evidence that wage differentials are unlikely to drive the gender gap in appointments. For example, the gender gap in CEO pay falls substantially after controlling for firm size (Bertrand and Hallock 2001; Adams, Gupta, Haughton, and Leeth 2007).

In addition to differing in productivity, male and female CEOs differ in two additional ways. First, once the board decides to replace a male CEO with a female CEO, it has to pay an additional search cost, ϕs_t, that reflects limited female labor supply, with s_t representing the proportion of males in the CEO candidate pool. Naturally, the more available male candidates, the higher the costs of searching for a female CEO.

We motivate a limited female CEO candidate pool in two ways. First, women might self-select away from advancement, finding it less attractive to apply for CEO positions because of prevailing social norms, a trade-off between career and family, or a dislike of the responsibilities and competitive environment associated with a job as a CEO. Similarly, if women are aware of differences in discriminatory treatment in the positions of top management, they may focus their careers on the fields less likely to discriminate against them. Second, there might be impediments preventing women from reaching the top. For example, women might wait longer to get promoted. This "sticky floor" in career ladders occurs because male-dominated firms are less precise at decoding skill signals from women due to, for example, communication style, and/or female workers have fewer opportunities for signaling their skills due to, for example, maternity leave (Bjerk 2008). Note that the setup of an additional search cost implies that the utility from either choice depends on the board's choice, d_{t-1}, in the previous period.

We define the CEO candidate pools as all the past and current executives and board members of the firm. We pursue this definition following the recent findings by Cziraki and Jenter (2020), who show that S&P 500 firms primarily promote insiders (80%) and rarely seek candidates who are not former or current employees or board members, and this pattern persists over their sample period of 1993 to 2012.

Second, the board is prejudiced against women leaders; that is, the board treats a female CEO less favorably than a male one with identical productive characteristics. We model this discrimination by assuming that the board has a disutility κ when hiring a female CEO, as in Becker (1971) who defines the "taste for discrimination" as being willing to pay something, either directly or in the form of reduced income, to be associated with a certain type. In the context of employer discrimination, the taste for discrimination is commonly modeled as differentiated wages for the same productivity. In our model, discrimination is specified as a one-time cost at the moment of a switch but can be viewed as the expected present value of all the future wage differentials between female and male CEOs of identical productivity upon hiring.

The interpretation of this modeling assumption requires discussion, as the theoretical literature on discrimination offers two forms: taste-based and statistical discrimination. In taste-based discrimination models, discrimination results from individual prejudicial tastes that take the form of a willingness to pay a price for this privilege (Becker 1971). In statistical discrimination models, discrimination stems from imperfect information so that employers make an educated guess based on observable characteristics (Phelps 1972;

Arrow 1973).

Although our model cannot technically cannot distinguish between these two forms, we believe statistical discrimination plays a limited role in CEO gender decisions. As noted above Cziraki and Jenter (2020) show that S&P 500 firms hire from a surprisingly small pool of highly familiar candidates. As such, these boards are unlikely to rely exclusively on gender information to assess candidates' abilities, as they have been acquainted with the candidates for several years. Moreover, Fryer (2007) and Bjerk (2008) make the argument that in a dynamic setting, information uncertainty resolves over time, so statistical discrimination is likely to play a small role over time.

These considerations motivate our specification of the board's one-period utility function as:

$$u_t(x_t, d_t) + \epsilon_t(d_t) = (a - \lambda d_t)k_t + \epsilon_t(d_t) - (\phi s_t + \kappa)d_t(1 - d_{t-1}), \tag{2}$$

In which $x_t \equiv \{k_t, s_t, d_{t-1}\}$ is a vector of state variables observed by both the board and the econometrician, and ϵ_t is choice-dependent and is observed by the board only.

The board makes the CEO gender decision, d_t, each period to maximize the expected present value of its utility. The Bellman equation for the problem is

$$V_t(x_t, \epsilon_t) = \max_{d_t} \Big\{ u_t(x_t, d_t) + \epsilon_t(d_t) + \beta \mathbb{E}[V_{t+1}(x_{t+1}, \epsilon_{t+1} \mid x_t, \epsilon_t, d_t)] \Big\}, \tag{3}$$

where $V_t(x_t, \epsilon_t)$ is the expected discounted utility of the board when it is in state (x_t, ϵ_t).

The third term represents the conditional continuation value of choosing d_t and can be rewritten as:

$$\mathbb{E}[V_{t+1}(x_{t+1}, \epsilon_{t+1} \mid x_t, d_t)] = \int \int \Big[V_{t+1}(x_t, \epsilon_t) g(\epsilon_{t+1}) d\epsilon_{t+1} \Big] f(x_{t+1} \mid x_t, d_t) dx_{t+1},$$

where $g(\epsilon)$ is the probability density function of the unobserved shock, and $f(x_{t+1} \mid x_t, d_t)$ is the transition probability density that represents the board's subjective beliefs about uncertain future events. This decomposition of the transition probability density of observable and unobservable state variables is based on the conditional independence assumption that is widely used in the discrete-choice literature (Hotz and Miller 1993). The best response probability that the individual chooses d given x is found by integrating the decision rule over the private information shock ϵ.

We briefly outline the steps to recast the fixed-point problem value function space with one in probability space, as in Aguirregabiria and Mira (2002). As the first step, we integrate the unobservable state variable ϵ out and rewrite the Bellman equation (3) as:

$$\begin{aligned} V_t(x_t) &= \int V_t(x_t, \epsilon_t) g(\epsilon_t) d\epsilon_t \\ &= \int \max_{d_t} \Big[u_t(x_t, d_t) + \epsilon_t(d_t) + \beta \mathbb{E}[V_{t+1}(x_{t+1} \mid x_t, d_t)] \Big] g(\epsilon_t) d\epsilon_t. \end{aligned}$$

The optimal decision δ_t at time t solves

$$\delta_t(x_t, \epsilon_t) = \arg\max_{d_t} \Big[u_t(x_t, d_t) + \epsilon_t(d_t) + \beta \mathbb{E}[V_{t+1}(x_{t+1} \mid x_t, d_t)] \Big],$$

and the conditional choice probability $p(d_t \mid x_t)$ can be obtained by integrating out ϵ_t

$$p(d_t \mid x_t) = \int \mathbb{1} \left\{ \delta_t(x_t, \epsilon_t) = d_t \right\} g(\epsilon_t) d\epsilon_t.$$

This Bellman equation can be further expressed in terms of the conditional choice probability $p(d_t \mid x_t)$ as:

$$V_t(x_t) = \sum_{d_t} p(d_t \mid x_t) \left\{ u_t(x_t, d_t) + \mathbb{E}[\epsilon(d_t) \mid x_t, d_t] + \beta \int V_{t+1}(x_{t+1}) f(x_{t+1} \mid x_t, d_t) dx_{t+1} \right\}, \tag{4}$$

where $\mathbb{E}[\epsilon(d_t) \mid x_t, d_t]$ is the conditional expectation of the unobservable state variable $\epsilon(d_t)$. Under the assumption that ϵ follows an extreme value type I distribution, this conditional expectation can be expressed in closed-form as a function of $p(d_t \mid x_t)$:

$$\mathbb{E}[\epsilon(d_t) \mid x_t, d_t] = \gamma - \sigma_\epsilon \ln(p(d_t \mid x_t)),$$

where γ is Euler's constant. A given probability, $p(d_t \mid x_t)$, corresponds to value function $V^p(x)$. Thus we can recast the equilibrium as a fixed-point problem in probability space:

$$p(d_t \mid x_t) = \int \mathbb{1}\left\{d_t = \arg\max_{d_t}\left[u_t(x_t, d_t) + \epsilon_t(d_t) + \beta \int V^p_{t+1}(x_{t+1}) f(x_{t+1} \mid x_t, d_t) dx_{t+1}\right]\right\} g(\epsilon_t) d\epsilon_t.$$

2.1 Policy functions

Figure 1 depicts the optimal policies of the board. In each panel, on the x- and y-axes are the fraction of men in the application pool and the capital stock. On the z-axis is the conditional choice probability of hiring a female CEO. Panel A corresponds to the case in which the current CEO is a man, and Panel B corresponds to the case in which the current CEO is a woman.

Several patterns are noteworthy. First, the model exhibits a great deal of inertia. In Panel A, in which the current CEO is a man, the probability of a woman being chosen as a new CEO is less than 0.06, while in Panel B, in which the current CEO is a woman, the probability is always above 0.9. This inertia stems from the large estimated search costs for a female CEO. Second, in neither Panel A nor Panel B does firm size matter much for the choice of the gender of the CEO, although smaller firms are slightly more likely to hire a woman if the current CEO is also a woman. Third, in Panel A, the probability of choosing a female CEO is decreasing in the fraction of men in the candidate pool. This result stems from total search costs being proportional to this fraction. Finally, in Panel B, we find that the probability of choosing a woman as a CEO is largely insensitive to the fraction of women in the candidate pool.

3. Data

Our data come from multiple sources. We start by retrieving a sample of executives from the annual 2021 ExecuComp files. We recognize a CEO per year using the annual CEO flag, and manually verify this information using EDGAR and the data provided by Jenter and Kanaan (2015). We exclude co-CEO cases because it is uncertain whether CEO gender decisions are influenced by the adoption of co-CEO models. Arena, Ferris, and Unlu (2011) show that most co-CEOs employment agreements (63%) are structured so that executive assignments are complementary in nature. This complementary task assignment makes it impossible to identify the gender difference in talent when it comes to co-gender leadership. We also exclude from the analysis CEO turnovers associated with acquisitions, mergers, and spin-offs. One goal of the model is to infer talent-based gender decisions from the relationship between CEO gender turnovers and firm size. Mergers, acquisitions, and spin-off triggered CEO turnovers obscure this connection because the simultaneous changes in firm size and CEOs are a mechanical result of the ownership changes.

Following the time convention adopted in ExecuComp, we assign successions to a fiscal year in which the CEO spent the greater part of the time. We exclude interim CEOs who are no longer in office after 12 months following Cremers and Grinstein (2014). This is because interim CEOs are usually appointed under duress and are employed temporarily to lead the firm until the board finds a suitable successor for the empty position. The nature of the decisions is different from those of non-interim CEOs for at least two reasons. First, interim CEOs serve for a short period (less than one year), and the choice of which can hardly reveal the talent difference in gender. Second, interim CEOs are typically selected from inside and usually also serve as the chairman of the board (Ballinger and Marcel 2010).

We collect director information from BoardEx. We merge BoardEx using the linking table provided by WRDS. The linking table builds on the CRSP-Compustat Mmerged database and makes use of variables such as CUSIPs, tickers, CIKs, and company names. It also provides a score indicating the matching quality. We opt for the "preferred" match, that is, with the lowest matching score whenever multiple GVKEYs in Compustat are linked to a single CompanyID in BoardEx. Note that because the linking process utilizes security-level identifiers such as CUSIP, it is possible that multiple CompanyIDs in BoardEx are linked to a single GVKEY in Compustat.

Finally, we obtain firm-level accounting data from Compustat. We require observations to have positive total assets (AT) and sales (SALE), and we remove all financial firms (SIC 6000-6999). We also require that a firm have at least two consecutive years of data because we need to lag some of the variables. Our final sample contains 4,786 CEOs of 2,245 firms with 26,635 firm-year observations from 2001 to 2019. We start from 2000 because BoardEx's coverage of U.S. public firms before 2000 is limited (Engelberg, Gao, and Parsons 2012). Since we will focus on transitions over two-year intervals, the sample for estimation eventually starts from 2001.

Table 1 contains summary statistics. On average, 3.5% of the sample firms are led by female CEOs in a given year. This ratio increases steadily over the sample period, growing from 1.9% during the 2001-2005 period to 5.5% during the 2016-2019 period. Cross-sectionally, the annual fraction of female-led firms (female representation) varies substantially across industries. Female CEOs are more present in Consumer (5.5%) and less visible in High Tech (2.5%) and Health (2.5%).

3.1 CEO Candidate Pools

Constructing the CEO candidate pool presents a challenge. This is because, on the one hand, a firm could potentially hire from a long list of candidates who all possess desired leadership traits. It is difficult to delineate the boundary of the candidate pool. On the other hand, the boards' selection criteria might vary drastically depending on the fastmoving business situations. It is hard to observe the short list.

We address this issue by considering two definitions of CEO candidate pool inspired by Cziraki and Jenter (2020). We also validate our construction of CEO candidate pool by tracking and tracing the new CEO's early employment upon a turnover. Our sample contains 2,625 CEO appointments from 2001 to 2019.

As a first cut, we focus on internal hires and define a CEO candidate pool as all the past and current executives and board members of the firm. We pursue this definition following the findings by Cziraki and Jenter (2020) who show that S&P 500 firms primarily promote insiders (80%) and rarely pouch outsiders, and this pattern persists over the sample period of 1993 to 2012.

Panel A of Table 2 presents the fraction of CEOs chosen from inside or outside of hiring firms in our sample. Similar to the classification used in Cziraki and Jenter (2020), we label insiders as "executives or board members who have been with the firm for at least one year before becoming CEO". This definition takes into account "staged successions," where an external successor is first appointed to the president or COO position as part of the succession plan. Insiders can be broken down into three groups based on their job titles, including (1) current executives that are promoted, (2) former executives (e.g., returning CEOs), and (3) current or former board members. Comparable to the findings reported in Cziraki and Jenter (2020), insiders dominate in CEO successions. A total of 84.6% of new CEOs have previously worked for the hiring firm, either as an executive or a board member. This highly similar result corroborates the validity of our first definition of the CEO candidate pool.

While this first definition of the CEO candidate pool applies to the majority of firms, it excludes the possibilities of external hires. If the main reason for hiring externally is a lack of internal female candidates, precluding external candidates may exaggerate the significance of the career path explanation. As such, we extend the first definition of the CEO candidate pool by including the current executives of peer firms in the same industry defined using 3-digit SIC codes.

To validate this second definition, we zoom in on the 405 external hires in our sample, examining these outside CEOs' prior work experiences. We define an outside CEO's prior working firm as the most recent employer that he or she worked for at least one year. We impose this time limit to avoid counting employment gaps. Ertimur, Rawson, Rogers, and Zechman (2018) find 84% of outside CEOs at S&P1500 firms experience employment gaps when moving from their prior executive positions. The gaps usually last for less than two years and can be attributed to labor market frictions. The gap activities include board membership, consulting, investing, and private firm employment.

Panel B of Table 2 reports statistics describing the prior working firms and associated working positions of outside CEOs. First, we report the distribution of job titles and roles that are grouped into three categories: executives, nonexecutives, and others. Specifically, executives consist of presidents, CEOs, and other C-suit positions. Moving down to the next level of the hierarchy, nonexecutives cover positions such as division heads and vice presidents. Examples of other titles include "owner", "founder" and "partner". Executives, in particular CEOs and COOs, constitute the largest portion (46.9%).[2] Notably, division heads, at 34.8%, are the second-largest source of CEO candidates. Division CEOs typically are not recognized as executives in ExecuComp. They are typically responsible for heterogeneous business sectors in a conglomerate. For example, General Electronic had at least 9 division CEOs in 2004 that operated businesses across 8 industries defined using 3-digit SIC codes.

Second, we report the types of prior working firms. Among 405 external hires, approximately 70% are from U.S. public firms, with the majority of firms in the S&P 1500 composite (53.8%). Similar to Cziraki and Jenter (2020), 22.5% of outside CEOs used to work in U.S. private firms, and the rest, 7.4%, are from foreign firms. This finding sug-

[2] Cziraki and Jenter (2020) distinguish between on-the-job CEOs being poached and unemployed ex-CEOs got hired. They find that, although CEOs are rarely being poached (14.2% of external hires), it is quite common to hire ex-CEOs who are in their gap years.

gests that the CEO candidate pool is bounded, depending on the nature of a hiring firm. Large U.S. public firms naturally lean towards candidates with work experience in similar firms perhaps because these firms expect their CEOs to have knowledge of U.S. markets and skills in dealing with a diverse set of shareholders.

Panel C of Table 2 presents the fraction of outside CEOs whose prior firms are in the same industry as the hiring firms. We do so for the 284 external hires from U.S. public firms because these firms have readily available SIC codes. In an initial attempt, we examine this industry connection using the SIC codes assigned by Compustat. Compustat assigns a single representative SIC code for each firm annually and overwrites the historical ones in case of industry change. Only 30% outside CEOs move from a firm that is in the same industry defined by a three-digit SIC code. This result is consistent with the findings in Ertimur et al. (2018), which shows that "36.1% of the executives move to a firm in the same two-digit SIC code while 20% of the executives move to a firm in the same four-digit SIC code." They interpret this finding as a result of non-compete agreements.

Although supported by similar empirical findings, this result is surprising given the anecdotal evidence that boards vouch for candidates with extensive industry experience. For example, Compass Minerals, a leading provider of minerals and nutrition products, appointed Kevin Crutchfield as CEO in May 2019. In the proxy statement, it states:"Mr. Crutchfield was selected by our Board following an extensive internal and external CEO search conducted by a CEO Search Committee of independent directors, which retained an executive search firm. Mr. Crutchfield brings us more than 30 years of mining experience, as well as his broad executive leadership capabilities, decision-making experience and operations expertise, having held chief executive roles at publicly traded mining companies."

As such, we break a firm down into its business segments and retrieve the segment information from Compustat Segment. Compustat Segment defines segments using four-digit SIC codes and assigns primary and secondary SIC codes to a segment if it involves more than one economic activity. We aggregate the reported segments using three-digit SIC codes and allow each firm to be present in up to three distinct segments (hence up to six three-digit SIC codes) ranked by sales in a given year.[3]

Our approach is motivated by two findings. First, many S&P 1500 firms are remarkably diversified. Among 405 external hires in our sample, half of the turnovers involve a hiring firm that operates in at least two distinct industries defined using three-digit SIC codes. Second, most of the external CEO candidates were division heads before serving as a CEO.

Our modification of the Compustat industry definition offers at least two advantages. First, it reflects the time-varying dynamics of business transformation, a feature that is highlighted in the text-based network industry classifications (TNIC) by Hoberg and Phillips (2016).[4] Second, it treats divisions of the conglomerates separately and therefore embraces the possibility that division heads serve as CEO candidates. This common practice, however, cannot be captured by the TNIC scheme, which identifies a peer firm based on the similarity of the entire business descriptions. For example, before joining Medtronic in 2011, Omar Ishrak served as the president and CEO of the GE Healthcare division. Yet GE and Medtronic do not belong to the same TNIC industry.

With the extension of the Compustat industry definition, the fraction of outside CEOs whose prior working firms are in the same industry as the hiring firms increases by two-fold, to 57.7%. Taking together, our second definition of CEO candidate pool covers 40.5% (= 164/405) of external hires and 90.8% (= 2,384/2,625) of the total hires.

[3]Note that segments are self-reported and are subject to managerial discretion. As such, the inconsistency in segment reporting may occur across firms and over time. To alleviate the impact of this problem, we aggregate the reported segments using three-digit SIC codes.

[4]The main difference between the TNIC by Hoberg and Phillips (2016) and the traditional SIC industry classification is: the Hoberg and Phillips (2016) classifications are based on the products that firms supply to the market rather than production processes (as is the case for existing industry classification schemes). Since CEOs are hired to oversee all areas of the corporation, it is not crucial to discern whether the industry boundary is determined by the production processes or the ultimate outputs.

4. Estimation

4.1 Beliefs

We set the annual discount factor, β, to be 0.99.We estimate this subjective discount factor outside of the model via grid search following, for example, Rust and Phelan (1997). We do so because the likelihood is not very sensitive to changes in β.

The success of model estimation depends crucially on an accurate estimation of the Markov transition matrix $f(k_{t+1}, s_{t+1} \mid k_t, s_t, d_t)$, which represents a firm's one-step-ahead beliefs about firm size and CEO labor supply. Following, for example, Rust and Phelan (1997), we decompose the transition matrix into a product of conditional densities:

$$f(k_{t+1}, s_{t+1} \mid k_t, s_t, d_t) = f_1(k_{t+1} \mid k_t) f_2(s_{t+1} \mid k_t, s_t, d_t) \tag{5}$$

The decomposition builds on the assumption that firm size evolves exogenously. We validate this assumption by running a regression of firm size, k_{t+1}, on all the state variables k_t, s_t, and d_t. We estimate the regression using the dynamic panel method in Arellano and Bover (1995) and Blundell and Bond (1998) that accounts for unobserved heterogeneity across firms.

Column (1) of Table 3 reports the regression results. The results are encouraging. As expected, we find that firm size is highly persistent. The CEO gender decision is statistically insignificant in predicting future firm size conditioning on current firm size. Interestingly, male fraction of the CEO candidate pool positively predicts future firm size. Despite its significance, we exclude s_t as a conditional variable for f_1 on a priori grounds. In the context of our dynamic model, this positive relationship would imply that firms believe that the current male fraction of the CEO candidate pool would be beneficial to its immediate future size. This implication is implausible considering the short timeframe. It is hard to imagine that a firm would expect a positive impact from a new (male) CEO within one year. Under the assumption of being exogenous, we model the stochastic evolution of firm size as an AR(1) process as follows:

$$k_{t+1} = \alpha_0 + \alpha_1 k_t + \epsilon^k_{t+1},$$

where α_0 is the drift, α_1 is the autoregressive coefficient, ϵ^k is a standard normal i.i.d. innovation with standard deviation σ^k. We use the method in Rouwenhorst (1995) to discretize this AR(1) process into a transition matrix with seven points of support. We choose the number of grid points to match the range of the observed firm size, and the autoregressive coefficient α_1 to match the conditional variance of the process. We use Rouwenhorst (1995) because Galindev and Lkhagvasuren (2010) and Kopecky and Suen (2010) find that it is more accurate than other alternatives when persistence is high.

While firm size evolves exogenously following a first-order Markov model, the evolution of the male fraction of the CEO candidate pool is more complicated, as it depends on the other state variables. Column (2) of Table 3 reports the regression results of s_{t+1} on all the state variables k_t, s_t, and d_t, using the same regression method as in Column (1). The estimation results appear plausible. The distribution of the male fraction of the CEO candidate pool shifts downwards with a female CEO, and is affected negatively by firm size. Indeed, female CEOs might lean towards female promotion. Moreover, firm size can predict the gender distribution of the future CEO candidate pool in at least two ways. On the one hand, larger firms have more female executives or board members internally because of focused media attention and public pressure. On the other hand, larger firms are more diversified and therefore are exposed to more female candidates externally across industries.

We let the male fraction of the CEO candidate pool have 25 evenly-spaced points of support. We approximate the conditional transition probabilities using kernel density estimates given that the distribution of s is left-skewed and capped by one. Furthermore, we capture the dependence between the male fraction of the CEO candidate pool and firm size through the use of subtransition matrices. Specifically, we first construct the transition matrix of s for the center grid of k and shift this matrix diagonally across different values of firm size k so as to reflect the negative association between s and k. We do so separately for the male and female CEOs.

4.2 Preferences

We then estimate the remaining parameters $\{a, \sigma_\epsilon, \lambda, \phi, \kappa\}$ in the dynamic model. However, these structural parameters are identified only up to scale. This is because the probability of a given alternative being chosen is determined only by the relative utilities from observed choices, and this ordering of the choice-specific utilities remains whenever we add or multiply all utilities by a constant. As such, we require normalizations of the level and scale of productivity. Specifically, we set the male CEO's productivity a to zero and scale all the rest of the parameters by the dispersion of the error term σ_ϵ.[5] We denote the vector of the normalized parameter $\boldsymbol{\theta} = \{\lambda, \phi, \kappa\}/\sigma_\epsilon$ for brevity. Therefore, the scaled coefficients $\boldsymbol{\theta}$ characterize the impact of the observed variables relative to the dispersion of the unobserved factors.

We estimate $\boldsymbol{\theta}$ using the Nested Pseudo Likelihood algorithm (NPL) proposed by Aguirregabiria and Mira (2002). The algorithm is built on the well-known Nested Fixed Point algorithm (NFXP) by Rust (1987, 1994) but swaps the order of the two-step procedure. First, starting with arbitrary conditional choice probabilities, we obtain parameter estimates $(\boldsymbol{\theta})^K$ by maximizing the following likelihood:

$$\boldsymbol{\theta}^K = \arg\max \sum \ln(\boldsymbol{\Psi}(p^{K-1}(\boldsymbol{\theta}))).$$

[5]The scale of utility and the dispersion of the error term are linked by definition. Multiplying utility by a constant $1/\sigma$ increases the variance of each $\epsilon(d)$ by the square of the constant: $Var(\epsilon(d)/\sigma) = (1/\sigma)^2 Var(\epsilon(d))$. Moreover, since the observed part of utility is linear in parameters $\{\lambda, \phi, \gamma\}$, normalizing the scale of utility is equivalent to normalize these parameters by the dispersion of the error term σ_ϵ. See Train (2009) for more discussion.

Second, we obtain the conditional choice probabilities p by solving a fixed point problem:

$$p^K(\boldsymbol{\theta}) = \boldsymbol{\Psi}(p^{K-1}(\boldsymbol{\theta})),$$

where $\boldsymbol{\Psi}$ denotes the policy iteration operator. We then iterate by repeating the first step, with the conditional choice probabilities obtained in the second step until we reach convergence of p and $\boldsymbol{\theta}$.

4.3 Identification

Just like with any other estimation algorithm, the success of NPL depends on whether the data-generating process is sufficiently informative about the parameters of the model. In our case, parameter identification is achieved by ensuring that these parameters affect the gender choices of the firm. First, the difference between the productivity of male and female CEOs, λ, is identified along two dimensions. Mechanically, the likelihood of the choice of a female CEO declines in λ. A more subtle effect occurs via firm size, as the downward-sloping relation between the probability of hiring a female CEO and firm size steepens as λ rises. The latter effect occurs because total CEO productivity is multiplicative in firm size.

Next, both the search cost parameter, ϕ, and the board disutility due to prejudice against female CEOs, κ, induce a fall in the probability of hiring a CEO. Because they govern the change in the gender of a CEO, they are jointly identified from the differences in the choice probabilities between firms with incumbent male and female CEOs. However, the downward-sloping relation between the probability of choosing a female CEO and the fraction of male job candidates steepens with ϕ, while this slope is unaffected by κ. Therefore, the two parameters are separately identified.

5. Results

5.1 Full-Sample Results

Panel A of Table 4 presents the parameter estimates. All three parameters are statistically significant from zero. Consistent with the human capital explanation, the positive estimate of the gender productivity difference, λ, indicates that the utility difference between having a female and a male CEO increases in firm size.

While the coefficient λ captures the board's utility difference between having a male and a female CEO, the remaining two coefficients capture change in board utility when switching from a male to a female CEO. The large positive estimate of the additional search cost, ϕ, conforms to the career path explanation, suggesting that the gender switching cost increases sharply with the male share of the CEO candidate pool. Next, we find a negative estimate of κ for the the board disutility from switching from a male to a female CEO. In conjunction with the interaction term on the male share of the CEO candidate pool, the negative coefficient implies that when all the CEO candidates are female, replacing a male CEO with a female one increases utility. In other words, net of the availability of suitable candidates, the board is actually in favor of hiring a female CEO. This finding echoes anecdotal evidence that board members are increasingly vocal about a commitment to promoting gender parity.

To assess the goodness of fit of the model, we directly compare the conditional choice probabilities in the data with the ones predicted by the model. These conditional choice probabilities $p(d \mid x)$ in the data are simply the frequencies of choices made by the subsample of boards at state x and thus depict the board decision patterns.

The probabilities are calculated for both the actual and simulated data. Following Kennan and Walker (2011), we simulate data conditioning on the initial observed states for each individual firm, and we use the model solution to generate a history of this firm covering the observed sample period. We repeat this procedure 100 times so as to alleviate simulation bias following Michaelides and Ng (2000).

We aggregate and compare the conditional choice probabilities along three dimensions, which correspond to the three state variables in our model. The small and large firms are the ones that lie below and above the median firm size. Similarly, the firms with fewer and more males in their CEO candidate pools are the ones that lie below and above the median of the fraction of men in the CEO candidate pool.

The results in Panel B of Table 4 show that the model closely reproduces the heterogeneous CEO gender decisions in the data. We first compare the actual versus predicted CEO gender choices made by male-led firms. Overall, approximately 0.5% of male-led firms in the data choose to appoint a female CEO, compared to 0.7% predicted by the model. These decisions vary little by firm size, corroborating the small coefficient estimates on the gender productivity difference. However, in the data, larger firms are more likely to appoint female CEOs, in contrast to the model prediction. This result is not surprising because of a positive correlation in the data between firm size and the fraction of female candidates in the pool. However, conditioning on the gender composition of the CEO candidate pools, we see the negative correlation between firm size and the propensity to hire female CEOs.

In sharp contrast, the tendency to have female CEOs varies dramatically with the gender composition of CEO candidate pools. The firms with relatively low male representation in their candidate pools have a 1.6% (1.7%) chance of replacing their current CEOs with female ones in the data (model), whereas the firms with relatively high male representation in their candidate pools have virtually no chance (0.3%) of doing so.

We then compare the actual versus predicted CEO gender choices made by female-led firms. As expected, 93.3% (93.7%) of female-led firms in the data (model) choose to continue with female leadership, despite the presence of the estimated difference in gender productivity. This result occurs because, when making CEO gender decisions, female-led firms compare the expected discounted future profits under CEOs of different genders. The present value of a male-led firm is affected by the substantial costs when switching from a male to a female CEO. This switching cost not only restrains male-led firms from recruiting women CEOs but also incentivizes female-led firms to continue with their women leaders.

5.2 Counterfactuals

While the parameter estimates in Panel A of Table 4 are indicative of the direction of the various effects in the model, they are not informative about the relative importance of the human capital, career path, and distaste factors. We next conduct counterfactuals to disentangle and gauge the contribution of each of the three channels in explaining firms' propensities to appoint female CEOs.

We implement three experiments in which we shut down the human capital ($\lambda = 0$), career path ($\phi = 0$), or distaste ($\kappa = 0$) channels. We report the counterfactual results in Panel C of Table 4. The human capital and distaste channels explain only a small portion of the low fraction of female CEOs. The career path channel, on the other hand, explains a significant part. In our sample estimation, eliminating the career path factor increases the representation of female CEOs from 4.9% to 52.8%.

5.3 Subsample Results

We next conduct subsample analysis over time and across industries. For each set of subsamples we investigate, we test the equality of estimates using the utility homogeneity test proposed by Swait and Louviere (1993). The null hypothesis of this test is that the difference in parameter estimates from different samples is not due to heterogeneous scale factors. Details regarding the test are in Appendix A.

Female CEO representation grows incrementally, as shown in Table 1, from 1.9% in the 2001-2005 period to 5.5% in the 2016–2019 period. We break our sample into two sub-periods, 2001–2009 and 2010–2019, and present the estimation results in Panel A of Table 5. The model does a good job of matching the upward trend in female CEO representation. The two sets of parameter estimates are similar in magnitude, with the former slightly larger than the latter. This difference, however, can be attributed to the difference in scale factors between the two subsamples based on the Swait and Louviere (1993) equality test. This result suggests that, although progress has been made to propel more women into CEO positions, the relative importance of the main drivers remains the same.

As shown in Table 1, there exists substantial heterogeneity in female CEO representation across industries, ranging from 2.5% in High Tech and Health to 5.5% in Consumer Wholesale, Retail, and Services. Next, we seek to understand whether the relative importance of the three explanations varies across industries.

Panel B of Table 5 presents the estimation results from Fama-French five industries. The model does a good job of matching the variation in female CEO representation across industries. The model-predicted female CEO percentages closely track their data counterparts.

One main theme running through the estimation results is that the parameter estimates of ϕ are highly significant across industries. We conclude that having insufficient female candidates in the pools exerts a statistically significant influence on the underrepresentation of female CEOs. Despite its consistent importance, the marginal impact of this career path channel varies across industries. After adjusting for the possible scale differences, we find that the impact is the strongest for firms in Manufacturing (Energy and Utilities), which is traditionally male-dominated, and less so for firms in Consumer (Wholesale and Services).

This finding is in accord with the fact that the U.S. labor market is highly gender-segregated by industry and occupation. For example, roughly a fifth of jobs in the oil and gas industry are filled by women, according to a report by the World Petroleum Council and The Boston Consulting Group in 2017. These women also work disproportionately in office jobs as opposed to technical ones, and therefore fail to accumulate the critical experiences that are often considered prerequisites for career advancement. As a result, it is unsurprising that 94.9% of the CEO candidates in Energy are male, and identifying an equally competitive female CEO in this pool becomes extremely challenging. Gender segregation is not unique to the oil and gas industry. It also applies to High-Tech, Health, and the rest of Manufacturing.

Interestingly, the decision to have a female CEO is only marginally sensitive to firm size, suggesting that there exists few differences in gender productivity. The parameter estimates of λ are statistically significant in two out of five industries. Finally, all the industries appear to be in favor of appointing female CEOs to counteract the negative effect of the career path factor.

6. Heterogeneity

6.1 CEO productivity, firm performance, and the glass cliff

We don't include firm performance as one of the state variables. Yet firm performance is considered to be relevant to our study for at least two reasons. First, it is commonly used to infer CEO productivity. For example, Taylor (2010) assumes that firm-specific profitability, measured by returns on assets, mean-reverts around the current CEO's ability level. Following this model assumption, the CEO mean ability is inferred by the average level of profitability. However, using firm profitability to identify CEO ability is complicated. As discussed by Taylor (2010), besides differences in CEO productivity, average profitability might vary because of variations in (1) industry profit margins and accounting rules, (2) persistence in profitability, and (3) idiosyncratic shocks that are beyond the CEO's control. To address these concerns, Taylor (2010) therefore tries numerous ways to adjust profitability so as to separate these factors. We opt for a much simpler approach by utilizing directly the multiplicative relation between firm size and CEO productivity, an assumption also adopted by Taylor (2010). We then infer gender productivity difference from the variation of gender choice probabilities in response to firm size. This approach is advantageous to serve our purpose as it allows us to identify neatly the average gender gap in CEO productivity while maintaining a parsimonious model setup.

6.2 Corporate governance

Taylor (2010) estimates a dynamic model and shows that the scarcity of forced CEO turnovers can be mainly attributed to entrenchment and weak corporate governance. If weak governance influences all firms the same way, reducing the forced CEO turnover and therefore lowering the incidence of CEO gender switches. Our estimation results serve as an upper bound because more gender switches in the CEO imply fewer gender-related differences.

Alternatively, weak governance can impact firms differently depending on the state variables. For example, weak governance can be systematically associated with male-led firms with few female candidates available. In this case, we would overestimate the importance of the career path channel since we misattribute weak governance to it. However, it is important to note that the CEO candidate pool for a firm includes suitable contestants both inside and outside of the firm. As shown in Panel A of Table 6, the CEO candidate pool, on average, consists of 353 contestants, with 324 (91.8%) coming from outside of the firm. Treating insiders and outsiders equally, the influence of a firm's corporate governance on the gender composition of its candidate pool is limited. One may also argue that weak governance causes a firm to mainly promote inside as well. In the data, we indeed observe more inside promotions (84.6%) as compared to external hires. Yet, as shown in Table 6, gender composition is fairly comparable for the candidate pools inside and outside of the firm.

Nevertheless, we conduct subsample analysis to examine the effect of corporate governance. Our (first) governance measure is board independence, which is defined by the representation of outside directors.[6] Outside directors have incentives to enhance board decisions because of career concerns, that is, concerns about the effects of their current decisions on their reputations (Fama and Jensen 1983).

We sort firm-year observations into terciles based on board independence and contrast the top and bottom subsamples in Table 6. At the first glance, more outsiders are associated with less male representation in the CEO candidate pools, more so for contestants inside. This is not surprising in light of Adams and Ferreira (2009) who show that (1) women serve as independent directors in 84.07% of positions, and (2) female directors are more independent and are likely to be tougher monitors of CEOs as compared to their male counterparts.

More importantly, this wedge in the gender composition of the CEO candidate pools does not explain the different female CEO representation between the high and low board independence groups. Instead, the estimation results in Panel B of Table 6 show that the lower number of women CEOs in the firms with weak governance is due to the enlarged gender difference in productivity. Conditional on the same switch costs, women CEOs appear to underperform more in firms with poor governance.

7. Conclusion

This study has sought to understand the quantitative importance of three central factors that contribute to the low representation of women among the CEOs of the largest U.S. corporations. We consider actual productivity differences, search costs, and genuine board distaste for female CEOs. To understand the relative importance of these factors,

[6] We refer outside directors as the non-executive ones in BoardEx. Alternatively, ISS (formerly RiskMetrics) classifies directors into three categories: inside directors, affiliated outside directors, and independent outsiders. Unlike BoardEx that retrieves data directly from firms' SEC filings and therefore contains self-reported director classification, ISS independently classifies directors as independent according to its own standards. We opt for BoardEx to preserve the consistency of the classification. RiskMetrics acquired ISS in 2007 and since then significantly relaxed the director independence criteria. See, for example Houston, Lee, and Shan (2016), for a discussion.

we estimate a dynamic discrete choice of the gender of a CEO. We find that search costs matter most. After we shut down this channel in our model, we find that boards actually prefer women, and the fraction of women CEOs rises from the low single digits to over 50%. In our model, high search costs arise because of a high fraction of men in the CEO applicant pool. This finding suggests that the problem of low female representation occurs much further down the job ladder, as few women are in the applicant pool in the first place.

References

Adams, Renée B., and Daniel Ferreira, 2009, Women in the boardroom and their impact on governance and performance, *Journal of Financial Economics* 94, 291–309.

Adams, Renee B., and Tom Kirchmaier, 2012, From female labor force participation to boardroom gender diversity, Working paper, Oxford University.

Adams, Susan M., Atul Gupta, Dominique M. Haughton, and John D. Leeth, 2007, Gender differences in ceo compensation: evidence from the usa, *Women in Management Review* 22, 208–224.

Aguirregabiria, Victor, and Pedro Mira, 2002, Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models, *Econometrica* 70, 1519–1543.

Arcidiacono, Peter, and Paul B. Ellickson, 2011, Practical methods for estimation of dynamic discrete choice models, *Annual Review of Economics* 3, 363–394.

Arellano, Manuel, and Olympia Bover, 1995, Another look at the instrumental variable estimation of error-components models, *Journal of Econometrics* 68, 29–51.

Arena, Matteo P., Stephen P. Ferris, and Emre Unlu, 2011, It takes two: The incidence and effectiveness of co-ceos, *Financial Review* 46, 385–412.

Arrow, Kenneth J., 1973, The theory of discrimination, in *Discrimination in Labor Markets*, volume 3, 3–33 (Princeton University Press).

Ballinger, Gary A., and Jeremy J. Marcel, 2010, The use of an interim CEO during succession episodes and firm performance, *Strategic Management Journal* 31, 262–283.

Becker, Gary S., 1971, *The Economics of Discrimination*, second edition (University of Chicago Press).

Bell, Linda A., 2005, Women-Led firms and the gender gap in top executive jobs, Working paper, Haverford College and IZA.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz, 2010, Dynamics of the gender gap for young professionals in the financial and corporate sectors, *American Economic Journal: Applied Economics* 2, 228–255.

Bertrand, Marianne, and Kevin F. Hallock, 2001, The gender gap in top corporate jobs, *Industrial & Labor Relations Review* 55, 3–21.

Bjerk, David, 2008, Glass ceilings or sticky floors? statistical discrimination in a dynamic model of hiring and promotion, *The Economic Journal* 118, 961–982.

Blundell, Richard, and Stephen Bond, 1998, Initial conditions and moment restrictions in dynamic panel data models, *Journal of econometrics* 87, 115–143.

Bowlus, Audra J., and Zvi Eckstein, 2002, Discrimination and skill differences in an equilibrium search model, *International Economic Review* 43, 1309–1345.

Cremers, K. J. Martijn, and Yaniv Grinstein, 2014, Does the market for CEO talent explain controversial CEO pay practices?, *Review of Finance* 18, 921–960.

Cziraki, Peter, and Dirk Jenter, 2020, The market for CEOs, Working paper, University of Toronto, and London School of Economics and Political Science.

Edmans, Alex, Xavier Gabaix, and Augustin Landier, 2009, A multiplicative model of optimal CEO incentives in market equilibrium, *Review of Financial Studies* 22, 4881–4917.

Engelberg, Joseph, Pengjie Gao, and Christopher A. Parsons, 2012, The price of a ceo's rolodex, *Review of Financial Studies* 26, 79–114.

Ertimur, Yonca, Caleb Rawson, Jonathan L. Rogers, and Sarah L. C. Zechman, 2018, Bridging the gap: Evidence from externally hired ceos, *Journal of Accounting Research* 56.

Fama, Eugene F., and Michael C. Jensen, 1983, Separation of ownership and control, *The Journal of Law and Economics* 26, 301–325.

Fryer, Roland G., 2007, Belief flipping in a dynamic model of statistical discrimination, *Journal of Public Economics* 91, 1151–1166.

Galindev, Ragchaasuren, and Damba Lkhagvasuren, 2010, Discretization of highly persistent correlated AR(1) shocks, *Journal of Economic Dynamics and Control* 34, 1260–1276.

Gayle, George-Levi, Limor Golan, and Robert A Miller, 2012, Gender differences in executive compensation and job mobility, *Journal of Labor Economics* 30, 829–872.

Guryan, Jonathan, and Kerwin Kofi Charles, 2013, Taste-based or statistical discrimination: The economics of discrimination returns to its roots, *The Economic Journal* 123, F417–F432.

Hoberg, Gerard, and Gordon Phillips, 2016, Text-based network industries and endogenous product differentiation, *Journal of Political Economy* 124, 1423–1465.

Hotz, Joseph V., and Robert A. Miller, 1993, Conditional choice probabilities and the estimation of dynamic models, *The Review of Economic Studies* 60, 497–529.

Houston, Joel F., Jongsub Lee, and Hongyu Shan, 2016, In search of board independence: Former employees, shades of gray and director classifications revisited, Working paper, University of Florida.

Jenter, Dirk, and Fadi Kanaan, 2015, CEO Turnover and Relative Performance Evaluation, *The Journal of Finance* 70.

Kennan, John, and James R. Walker, 2011, The effect of expected income on individual migration decisions, *Econometrica* 79, 211–251.

Kopecky, Karen A., and Richard M. H. Suen, 2010, Finite state Markov-Chain approximations to highly persistent processes, *Review of Economic Dynamics* 13, 701 714.

Michaelides, Alexander, and Serena Ng, 2000, Estimating the rational expectations model of speculative storage: A monte carlo comparison of three simulation estimators, *Journal of econometrics* 96, 231–266.

Phelps, Edmund S., 1972, The statistical theory of racism and sexism, *American Economic Review* 62, 659–661.

Rouwenhorst, Geert, 1995, Asset pricing implications of equilibrium business cycle models, in Thomas F. Cooley, ed., *Frontiers of Business Cycle Research* (Princeton University Press).

Rust, John, 1987, Optimal replacement of gmc bus engines: An empirical model of harold zurcher, *Econometrica* 55.

Rust, John, 1994, Structural estimation of markov decision processes, in Robert Engel, and Daniel McFadden, eds., *Handbook of Econometrics*, volume 4, 3081–3143 (Elsevier).

Rust, John, and Christopher Phelan, 1997, How social security and medicare affect retirement behavior in a world of incomplete markets, *Econometrica* 65, 781.

Schubert, Renate, Martin Brown, Matthias Gysler, and Hans Brachinger, 1999, Financial Decision-Making: are women really more risk averse?, *American Economic Review* 89, 381–385.

Smith, Nina, Valdemar Smith, and Mette Verner, 2013, Why are so few females promoted into CEO and vice president positions? danish empirical evidence, 1997–2007, *Industrial & Labor Relations Review* 66, 380–408.

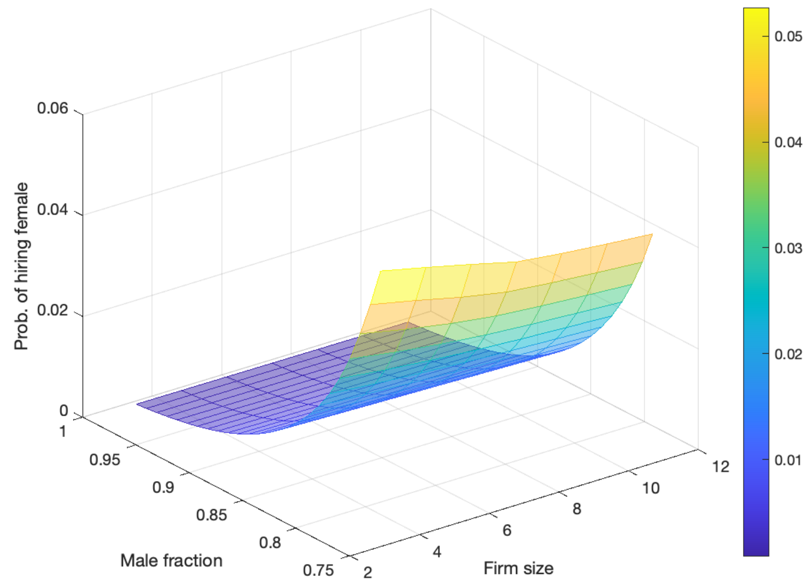
Swait, Joffre, and Jordan Louviere, 1993, The role of the scale parameter in the estimation and comparison of multinomial logit models, *Journal of Marketing Research* 30, 305–314.

Taylor, Lucian A., 2010, Why are CEOs rarely fired? evidence from structural estimation, *The Journal of Finance* 65, 2051–2087.

Train, Kenneth E., 2009, *Discrete Choice Methods with Simulation*, second edition (Cambridge University Press).

Wolfers, Justin, 2006, Diagnosing discrimination: stock returns and CEO gender, *Journal of the European Economic Association* 4, 531–541.

Panel A:



Panel B:

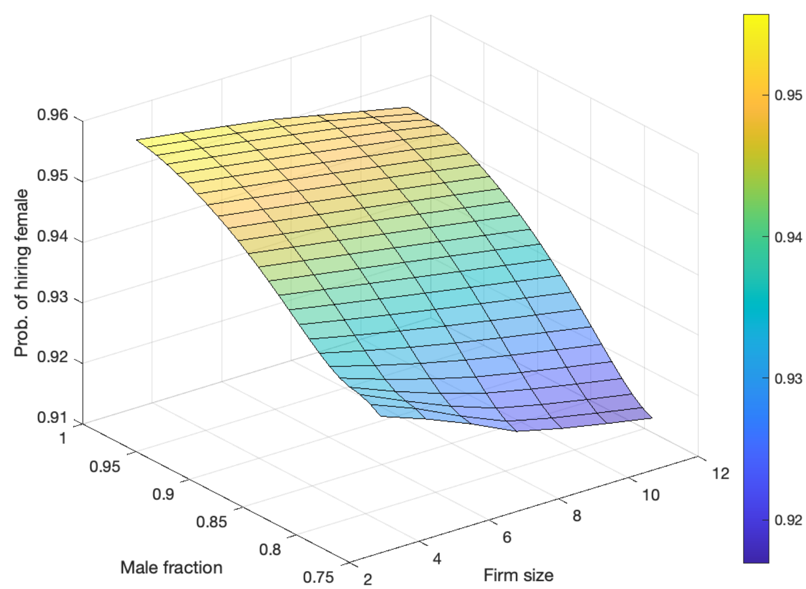


Figure 1. Policy functions. This figure depicts the optimal conditional choice probabilities as a function of the three model state variables. Panel A corresponds to the case in which the current CEO is a man, and Panel B corresponds to the case in which the current CEO is a woman.

Table 1. Summary Statistics

The table presents summary statistics on representation of female CEOs over the years and across industries. The sample contains 4,786 CEOs of 2,245 firms with 26,635 firm-year observations from 2001 to 2019. The industry is defined using Fama-French 5 industry classification.

	Firm-Year Obs	Female CEOs	Male CEOs	% Female
Full sample	26,635	938	256,972	3.5%
2001-2005	6,873	134	6,739	1.9%
2006-2010	7,388	221	7,167	3.0%
2011-2015	7,215	300	6,915	4.2%
2016-2019	5,159	283	4,876	5.5%
Consumer	6,185	340	5,845	5.5%
Manufacturing	7,413	248	7,165	3.3%
High Tech	6,535	165	6,370	2.5%
Health	2,762	69	2,693	2.5%
Other	3,740	116	3,624	3.1%

Table 2. Early Employment of CEO Hires

The table presents statistics describing early employment of the new CEOs upon turnovers. The sample contains 2,625 CEO turnovers from 2001 to 2019. Panel A presents the fraction of CEOs chosen from inside or outside of hiring firms. Panel B reports statistics describing the prior working firms and associated working positions of outside CEOs. Panel C presents the fraction of outside CEOs whose prior working firms are in the same industry as the hiring firms. We do so for the 284 external hires from US public firms.

Panel A: All CEO hires (N = 2,625)

Insiders			Outsiders
Current executives	Former executives	Current or former board members	
2,049	39	132	405
78.1%	1.5%	5.0%	15.4%

Panel B: External CEO hires (N = 405)

US Public firms		US Private firms	Foreign
S&P 1500	Non S&P 1500		
218	66	91	30
53.8%	16.3%	22.5%	7.4%

Executives	Non Executives		Others
	Division Heads	VP	
190	141	42	32
46.9%	34.8%	10.4%	7.9%

Panel C: External Public CEO hires (N = 284)

3-digit SIC Compustat	3-digit SIC Segment
84	164
29.6%	57.7%

Table 3. Regression Estimates for the Transition Matrices Construction

The sample contains 4,786 CEOs of 2,245 firms with 26,635 firm-year observations from 2001 to 2019. The table presents coefficient estimates and standard errors in parentheses from the two panel regressions. Columns (1) and (2) are with the dependent variable firm size k and male representation of the CEO candidate pool s at time $t+1$ respectively. Each of the regression is estimated using the dynamic panel method in Arellano and Bover (1995) and Blundell and Bond (1998) that accounts for unobserved heterogeneity across firms.

	(1) Firm size k_{t+1}	(2) Male representation s_{t+1}
Firm size k_t	0.88	-0.003
	(0.017)	(0.001)
Male representation s_t	0.694	0.888
	(0.119)	(0.011)
CEO gender decision d_t	-0.019	-0.006
	(0.019)	(0.002)

Table 4. Estimation Results: Full Sample

The table reports the estimation results from the full sample. Panel A reports the structural parameter estimates from the dynamic choice model in Section 2 with their corresponding standard errors in parentheses. λ capture the productivity difference between the male and female CEOs. ϕ capture the additional search cost demanded for getting a female CEO. κ stands for the distaste for the board to hire a female CEO. Panel B compares the conditional choice probabilities calculated from the actual data with the ones predicted by the model. Small and large firms refer to the ones lie below and above the median of the discretized distribution respectively. Similarly, firms with less and more male candidates refer to the ones lie below and above the median of the discretized male representation of the CEO candidate pool. Panel C reports the results of three counterfactual experiments. In each of the experiment, we shut down the human capital ($\lambda = 0$), career path ($\phi = 0$), and distaste ($\kappa = 0$) channels respectively.

Panel A: Parameter estimates			
	λ	ϕ	κ
	0.003	19.280	-9.199
	(0.002)	(1.882)	(1.620)

Panel B: Data and model predicted choice probabilities		
	Data	Model
All firms	3.51%	4.92%
Male-led firms	0.51%	0.71%
Small firms	0.41%	0.69%
Large firms	0.56%	0.71%
Less male candidates	1.56%	1.68%
More male candidates	0.29%	0.34%
Female-led firms	93.34%	93.71%
Small firms	94.44%	93.96%
Large firms	91.21%	93.28%
Less male candidates	92.77%	93.13%
More male candidates	93.83%	94.67%

Panel C: Counterfactual experiments			
	$\lambda = 0$	$\phi = 0$	$\kappa = 0$
	5.82%	52.76%	1.45%

Table 5. Estimation Results: Subsamples

The table reports the estimation results from several subsamples. Panel A reports the estimation results from two sub-periods: 2000-2009 and 2010-2019. Panel B reports the estimation results from Fama-French five industries. For each subsample, the first three columns contain the structural parameter estimates from the dynamic choice model in Section 2 with their corresponding standard errors in parentheses. λ capture the productivity difference between the male and female CEOs. ϕ capture the additional search cost demanded for getting a female CEO. κ stands for the distaste for the board to hire a female CEO. The last two columns contain the choice probabilities calculated from the actual data with the ones predicted by the model.

Panel A: Time

	Estimates			% Female	
	λ	ϕ	κ	Model	Data
2001-2009	0.004	20.695	-10.610	3.2%	2.4%
	(0.003)	(3.181)	(2.775)		
2010-2019	0.001	19.016	-8.904	5.8%	4.6%
	(0.002)	(2.398)	(2.039)		

Panel B: Industry

	Estimates			% Female	
	λ	ϕ	κ	Model	Data
Consumer	0.004	14.513	-5.269	7.2%	5.5%
	(0.003)	(2.459)	(2.027)		
Manufacturing	0.000	23.997	-13.574	4.7%	3.3%
	(0.003)	(4.404)	(3.838)		
High Tech	0.004	26.654	-15.905	3.5%	2.5%
	(0.004)	(8.625)	(7.747)		
Health	0.002	10.527	-0.580	4.1%	2.5%
	(0.004)	(16.917)	(15.226)		
Other	0.008	18.993	-9.397	4.5%	3.1%
	(0.005)	(4.947)	(4.303)		

Table 6. Governance

The table reports results on governance splits. Panel A reports the gender composition of the CEO candidate pools. Panel B reports the estimation results from top and bottom thirds of the sample ranked by board independence. For each subsample, the first three columns contain the structural parameter estimates from the dynamic choice model in Section 2 with their corresponding standard errors in parentheses. λ capture the productivity difference between the male and female CEOs. ϕ capture the additional search cost demanded for getting a female CEO. κ stands for the distaste for the board to hire a female CEO. The last two columns contain the choice probabilities calculated from the actual data with the ones predicted by the model.

Panel A: Gender composition of CEO candidate pools

	Total		Insiders		Outsiders		Diff.	
	No.	% Male	No.	% Male	No.	% Male	Mean	t-Stat.
Full sample	353	91.1%	29	90.7%	324	91.3%	-0.6%	-13.7
Less outsiders	344	91.0%	26	91.0%	318	91.2%	-0.1%	-1.5
More outsiders	300	90.9%	34	89.7%	266	91.5%	-1.7%	-20.5

Panel B: Estimates

	Estimates			% Female	
	λ	ϕ	κ	Model	Data
Less outsiders	0.004	16.584	-6.800	4.4%	2.5%
	(0.004)	(2.969)	(2.537)		
More outsiders	0.003	16.401	-6.894	6.3%	4.2%
	(0.003)	(3.833)	(3.337)		

Appendix A. Test of the equality of parameter vectors from subsamples with heterogenous scale factors

In this section, we briefly outline the utility homogeneity test proposed by Swait and Louviere (1993) for assessing whether the parameters estimated from two different data sets are equal while controlling for scale differences between these data sets.

As discussed in Section, the parameters of logit models are identified only up to scale, and this utility scale is inversely related to the error variance that is sample specific. Consequently, it is necessary to consider the scale factors when comparing the model parameters estimated from distinct data subsamples.

Specifically, let β be the true underlying parameters and let μ be the scale factor for a particular data set. Since μ cannot be identified separately, we estimate the parameter vectors $\theta = \mu\beta$. To test the equality of θ from the two subsamples, we conduct a two stage Chow-test with the null hypothesis that $\beta_1 = \beta_2$ and $\mu_1 = \mu_2$.

In the first stage, we test whether $\beta_1 = \beta_2 = \beta$ while allowing heterogeneous scale factors between two subsamples. The likelihood ratio statistic is

$$\tau_A = -2[L_\mu - (L_1 + L_2)],$$

where L_μ is the log-likelihood of the model estimated on the combined sample that is with the same β but allows for scale differences. L_1 and L_2 are the log-likelihoods of the two separate models estimated on the two subsamples. This test statistic is asymptoticly Chi-squared distributed with $(K+1)$ degrees of freedom. K is given by the number of restrictions imposed on the parameter vector, and the additional degree of freedom is to allow the scale to vary under the alternative hypothesis.

Because that the scaling factors of the two subsamples μ_1 and μ_2 cannot be identified separately in any particular set of empirical data, we form the relative scale factor $\bar{\mu}_2 = \mu_2/\mu_1$ by normalizing μ_1 to unity, and create the combined sample by concatenating the two subsamples scaled by 1 and $\bar{\mu}_2$ respectively. We then conduct a grid search to obtain the point estimate of $\bar{\mu}_2$ that maximizes the log likelihood L_μ.

Once we fail to reject the null hypothesis in the first stage, we proceed to the second stage and test whether $\mu_1 = \mu_2 = \mu$. The likelihood ratio statistic is

$$\tau_B = -2[L_p - L_\mu],$$

where L_p is the log-likelihood of the model estimated on the combined sample of a direct concatenation. This test statistic is asymptotically Chi-squared distributed with one degree of freedom.