Gender Bias in Promotions: Evidence from Financial Institutions

RUIDI HUANG[a]

ERIK J. MAYER[b]

DARIUS MILLER[c]

November 8, 2021

Abstract

We test for gender bias in promotions at financial institutions using two central predictions of Becker's (1957, 1993) model: firms with bias will (1) raise the promotion bar for marginally promoted female workers, and (2) incur costs from forgoing efficient employment practices. We find support for both of these predictions using a new nationwide panel of mortgage loan officers and their branch managers, encompassing approximately 72,000 workers from over 1,000 shadow banks from 2014 to 2019. Overall, our findings provide evidence that gender bias is an important factor in gender disparities at financial institutions.

JEL Classification: G21, M51, J71

Keywords: Promotion, gender bias, financial institutions, shadow bank

Acknowledgments:
For helpful comments we thank Swaminathan Kalpathy, Stacey Jacobsen, Pab Jotikasthira, and seminar and conference participants at Southern Methodist University (finance group and economics group) and Aggie Longhorn Innovation and Policy Conference. We are grateful to the Conference of State Bank Supervisors (CSBS) for granting us access to data from NMLS Consumer Access[SM], please see https://nmlsconsumeraccess.org/. We also thank Zillow for providing the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the authors and do not reflect the position of the CSBS or Zillow Group.

[a] Cox School of Business at Southern Methodist University, ruidih@smu.edu.
[b] Cox School of Business at Southern Methodist University, emayer@smu.edu.
[c] Cox School of Business at Southern Methodist University, dpmiller@smu.edu.

The U.S. financial sector employs over 9 million workers, often in high paying jobs that influence capital allocation in the economy. However, women hold less than a quarter of the industry's top leadership roles, despite making up the majority of its entry-level workforce.[1] If financial firms are biased against women because of tastes (Becker, 1957) or inaccurate stereotypes (Bordalo, Coffman, Gennaioli, and Shleifer, 2016), they will raise the promotion bar for female workers. Although the media, policymakers, and surveys all point to discrimination as a factor inhibiting women's climb up the corporate ladder, it is well known to researchers that omitted variable and infra-marginality problems make it difficult to find convincing evidence of bias using standard regression methods (Bertrand and Duflo, 2017).

In this paper, we combine novel large-sample microdata on workers and their managers with recent advances in Becker's (1993) "outcome test" to provide new evidence on gender bias at U.S. financial institutions. We first test whether male workers are more likely to be promoted than their female counterparts, controlling for performance, experience, and other characteristics. Next, we use outcome tests to investigate whether gender disparities are the result of bias. We then study how gender bias in promotions is affected by the gender of the decision maker ("in-group" tests) as well as past sales performance (the "Peter Principle"). Finally, we investigate Becker's (1957) prediction that biased employment practices are costly to firms.

A significant barrier to conducting tests for biased promotion practices is obtaining large-sample demographic and employment data on rank-and-file employees, since such data are typically not required disclosures. To overcome this challenge, we use licensing and registration information from the Nationwide Mortgage Licensing System (NMLS) to construct, to our knowledge, the first nationwide panel of mortgage loan officers and branch managers at mortgage companies.[2] We then link these loan officers directly to the loans they originate in a national mortgage database provided by CoreLogic, and information on subsequent foreclosures from

[1] For industry employment statistics, see data from the Bureau of Labor Statistics here: https://www.bls.gov/emp/tables/emp-by-detailed-occupation.htm. For information on gender gaps in senior management, see for example, Deloitte's "Within Reach? Achieving gender equality in financial services leadership." Women hold less than 22% of top leadership roles, less than 25% of board seats, and it wasn't until the year 2020 that a woman was first appointed as CEO of a major U.S. bank.

[2] The NMLS is a registry of all residential mortgage loan originators in the United States that was created as part of the Secure and Fair Enforcement for Mortgage Licensing Act of 2008 (SAFE Act). For non-depository institutions (mortgage companies), the NMLS also contains the loan officers' supervisors, denoted as branch managers. We follow the NMLS convention and use the term branch manager to denote loan originators' direct supervisor. Section 3.1 discusses the NMLS in more detail.

Zillow. Our sample covers approximately 72,000 workers from over 1,000 mortgage companies from 2014 to 2019.

Financial institutions and their loan officers provide a natural laboratory to test for gender bias in promotions. First, mortgage origination is economically important, with new originations totaling over $2 Trillion annually. Second, mortgage loan officers are attractive workers to study because their compensation is primarily based on loan origination volume, which gives the researcher a relatively complete measure of job performance. Finally, our setting allows us to study the bias that women face at the beginning of their careers, potentially revealing a "Broken Rung" of the corporate ladder for women which could perpetuate gender disparities in senior management.[3]

We start our analysis by documenting that only 27% of branch managers are women, despite the fact that women make up 35% of loan officers. We then conduct multivariate tests that control for sales performance (lending volume), experience, and other observable characteristics, and find that male loan officers are 15% more likely to be promoted to manager than their female counterparts. These findings document a "gender promotion gap" in which female workers begin to lose out to male workers at the very first step of the corporate ladder.

While gender bias could indeed be the cause of this promotion gap, it is well known that disparities alone do not prove bias because they can also result from differences in preferences or innate characteristics that the researcher cannot observe. For example, women can be more averse to risk (Barber and Odean, 2001) and competition (Niederle and Vesterlund, 2007). Women are also more likely to choose jobs with family-friendly hours, potentially forgoing some career advancement opportunities (Bertrand, Goldin, and Katz, 2010). Given the difficulty of controlling for all such factors, standard regression methods cannot ensure that gender gaps are due to bias rather than unobserved heterogeneity (Guryan and Charles, 2013).

To overcome this problem, Becker's (1993) Nobel Prize lecture proposed the "outcome test" as a direct approach to detect bias by comparing the success or failure of decisions at the margin.[4] For promotion decisions, the relevant outcome is how well the promoted manager

[3] The term "Broken Rung" was coined by McKinsey & Company and Sheryl Sandberg through LeanIn.Org. They argue that because it determines the pipeline of women available to promote, discrimination that begins at the first promotion is even more detrimental to women than the well-known "glass ceiling" at the top of the corporate ladder.
[4] We use the term "gender bias" to refer to biased decisions, whether they are driven by preferences (Becker, 1957), or by inaccurate beliefs such as exaggerated stereotypes (Bordalo et al., 2016). An important strength of outcome tests is that they can clearly distinguish bias from statistical discrimination, which does not lead to differences at the margin.

performs. If firms are unbiased, they will equate expected manager performance at the margin, creating the testable prediction that managerial performance will be identical for marginally promoted men and women. Unfortunately, Becker's outcome test has been challenging to bring to the data because of the difficulty in identifying marginal cases, which introduces the infra-marginality problem when researchers use averages instead: one can observe gender differences in *average* outcomes even if firms equate expected outcomes at the *margin* (Ayres, 2002). In other words, gender differences in average managerial performance may say very little about gender bias in promotions.

Several recent studies, however, have made important advances in implementing Becker's outcome test (see Arnold, Dobbie, and Yang (2018) for bail decisions and Dobbie, Liberman, Paravisini, and Pathania (2021) for consumer lending). These studies utilize the fact that instrumental variables (IV) tests estimate the local average treatment effect (LATE) on compliers to the instrument in order to estimate outcomes that are local to (at) the margin. Most relevant to our setting is Benson, Li, and Shue (2019), who use this IV approach to implement a Becker outcome test for the "Peter Principle" in promotions of sales workers to managerial roles—finding that firms are biased toward promoting high sales workers with lower managerial potential relative to low sales workers with higher managerial potential. We use an IV approach similar to Benson, Li, and Shue (2019) to circumvent the infra-marginality problem and test for gender bias based on the performance of marginally promoted male and female managers.

Our outcome tests require both a measure of managers' performance and a valid instrument for promotions. We measure managerial performance by taking our loan officer-year panel and calculating managers' impact on their subordinate officers' lending volume via manager fixed effects. These specifications include subordinate fixed effects, firm-year fixed effects, and other factors, to ensure that our measure of manager performance is not influenced by the matching of managers to subordinates, or by firm or industry trends. Our method closely follows advances in computing "value added" by workers (e.g., Abowd, Kramarz, and Margolis, 1999), teachers (e.g., Chetty, Friedman, and Rockoff, 2014a, 2014b), and sales managers (Benson, Li, and Shue, 2019).

In our setting, statistical discrimination would occur if employers with imperfect information were rational and non-prejudiced, yet optimized promotion decisions by using gender to proxy for aspects of managerial potential that were unobservable to them (Phelps, 1972, and Arrow, 1973).

We instrument for each worker's promotion using the firm-level average promotion rate in each year, leaving out the focal worker and her teammates (following the approach in Benson, Li, and Shue (2019)). We find that this instrument has a strong positive effect on promotions for both male and female loan officers. Yet, as we discuss in more detail in Section 5, the instrument remains plausibly exogenous to managers' ex post performance due in large part to the inclusion of firm-year fixed effects when computing performance. The compliers to this instrument can be thought of as marginal managers because they are promoted only if overall promotion rates are high but would not have been promoted had promotion rates been slightly lower. Therefore, by estimating the LATE of promotion on managerial performance for this set of instrument-compliers, we are able to estimate the performance of marginally promoted male/female managers.

Our main IV results show that marginally promoted male managers reduce the productivity of their subordinates by roughly 3% compared to an average manager. In contrast, marginally promoted female managers increase subordinates' productivity by 2% compared to an average manager. These results at the margin are significantly different from zero, and from each other, showing that firms set the bar for promotion higher for female workers. Taken together with the gender gap in promotions, these findings provide strong evidence of gender bias at the financial institutions we study.

We extend our analysis to examine "in-group" favoritism by testing whether women face more bias when they work for male managers (e.g., Egan, Matvos, and Seru, 2020). These tests can offer insight into whether biased preferences or inaccurate stereotypes are more likely to be driving our results. If preferences are the root cause, women should face greater bias under male decision makers, whose preferences likely tilt in favor of in-group members (men). Instead, we find that female workers face similar levels of bias under male and female decision makers, consistent with inaccurate gender stereotypes about managerial potential being held by both men and women.[5] An important implication of these results is that policies focused solely on increasing female representation in upper management will likely not be sufficient to create equal opportunity for the next generation of women in finance.

Our next tests explore whether an interplay between gender and the Peter Principle contributes to the gender bias we find. We are motivated by the incongruity between the notion

[5] Studies find similar evidence of biased stereotypes holding within groups in other settings. For example, estimates of racial bias against black defendants are similar under white and black judges (see, e.g., Anwar and Fang, 2006).

that high-performers are promoted past their level of competence, and recent studies showing positive effects of having women in top leadership roles.[6] For men, we find evidence consistent with the Peter Principle—when high volume loan officers are promoted, they make worse managers than low volume officers (at the margin). In sharp contrast, we find that when high volume female loan officers are promoted, if anything, they make *better* managers than low volume female officers (at the margin). Hence, the Peter Principle appears to be gender specific: a bias toward promoting the best salesman at the margin results in lower managerial performance (the Peter Principle), while in contrast, promoting more high-performing saleswomen at the margin does not sacrifice managerial performance (a finding we denote as the "Penelope Principle").

We next rule out alternative explanations for the differences we document at the margin, leaving gender bias as the most compelling explanation for our findings. For example, we show that the superior sales performance of marginally promoted female managers is not driven by their subordinates originating lower quality loans, nor by a lack of overlap in the underlying distributions of male and female managerial talent. Further, we provide evidence that our results are not driven by non-random selection into our promotion sample, such as female workers turning down promotions.

Our final set of tests examine a second central prediction of Becker's (1957, 1993) model: biased preferences/beliefs unrelated to productivity will be costly to firms. We present results from four tests of this hypothesis. First, at the worker-level we find using a difference-in-differences analysis that female loan officers' productivity increases under female managers, offering additional evidence on the opportunity cost of overpromoting men. Second, at the firm-level we find that less female representation among management (i.e., fewer female managers relative to the share of female workers) is associated with reduced lending volume. This pattern holds both across firms, and within firms over time. Third, we find that firms with more female representation among management experience greater employment growth. Fourth, we find that these firms are also more likely to survive. Although these firm-performance tests document correlations rather than causal effects, the results are consistent with Becker's (1957) prediction that biased employment practices are costly to firms.

[6] For example, studies show that firms benefit from female leadership in the context of investment decisions (Huang and Kisgen, 2013), corporate governance (Adams and Ferreira, 2009), and avoiding corporate litigation (Adhikari, Agrawal and Malm, 2019).

Our paper makes several contributions to the literature. Prior research documents gender gaps in a variety of economic outcomes such as wages (e.g., Bertrand, Goldin, and Katz, 2010), access to credit (Alesina, Lotti, and Mistrulli, 2013), and career trajectories in academic finance (Sherman and Tookes, 2021).[7] Studies of the gender gap in promotions are relatively rare and typically focus on a single firm due to the scarcity of detailed data on rank-and-file employees. Most recently, studies document that subjective performance evaluations (Benson, Li, and Shue 2021) and gender targets (Hospido, Laeven, and Lamo, 2021) help explain the promotion gap at a large retail chain and the European Central Bank, respectively. We contribute to this literature by documenting a significant gender gap in promotions at over 1,000 financial institutions, where "hard" measures of worker performance are available (loan volume), and the women we study have chosen a commission-based job in a male-dominated industry, likely reducing gender differences in risk aversion and competitive behavior (Fang and Huang, 2017, and Kumar, 2010).

Our primary contribution is to then provide direct evidence on the role of bias in the gender promotion gap using Becker's marginal outcome test.[8] Our tests show that women are held to a roughly 5% higher promotion standard, and point to biased stereotypes (rather than tastes) as the primary mechanism. These findings add to the nascent literature identifying bias using marginal outcome tests (e.g., Arnold, Dobbie, and Yang, 2018), and to recent work showing inaccurate beliefs can perpetuate long-standing inequities.[9] We provide additional insight into the mechanisms generating bias by documenting a new link between the Peter Principle and gender bias. Recent studies have found empirical support for the Peter Principle among sales managers (Benson, Li, and Shue, 2019) and in experimental settings (Chan, 2018). We extend this literature by providing novel evidence that the Peter Principle is stronger for men, which contributes to gender bias in promotions.

Our study also adds to the literature testing Becker's (1957) prediction that biased employment practices are costly to firms. Although prior evidence is mixed, our tests suggest that

[7] The recent literature has also studied gender differences in the pricing of arts (Adams, Kraussl, Navone, and Verwijmeren, 2021), the value of connections (Fang and Huang, 2017), the choice of finance career (Adams, Barber, and Odean, 2019), and the differences in finance research productivity during the pandemic (Barber, Jiang, Morse, Puri, Tookes, and Werner, 2021).

[8] The difficulty of distinguishing between gender gaps (disparities) and gender bias (discrimination) has led researchers to use field experiments to uncover evidence of gender bias in employment practices, such as "blind" orchestra auditions (Goldin and Rouse, 2000). See Bertrand and Duflo (2017) for a review of the literature.

[9] Several studies document biased stereotypes in experimental settings (e.g., Bordalo, Coffman, Gennaioli, and Shleifer (2016, 2019), and Coffman, Exley, and Niederle (2020)). Our results add to the "real-world" empirical evidence on the importance of biased stereotypes (see, e.g., Arnold, Dobbie, and Yang, 2018).

a bias toward promoting men can be costly in terms of total sales (lending volume), employment growth, and survival.[10] Our findings also support the hypothesis that having fewer women in managerial roles is costly because it reduces incentives for women who aspire to these positions (Blau and Kahn, 2017).

Finally, our analysis provides some of the first evidence we know of regarding the bias women face in promotions at the very beginning of their career. In contrast to the large literature examining gender differences at the top of the corporate hierarchy, early-career promotion decisions have largely evaded researchers' analysis. Our results reveal that women face a "Broken Rung" long before the "Glass-Ceiling" at the top of the corporate ladder.

2. Institutional setting

The mortgage companies we study, often referred to as "shadow banks," range from small and medium sized lenders with both in-person and online operations, to large FinTech lenders. These financial institutions are playing an increasingly important role in the U.S. mortgage market (e.g., Buchak, Matvos, Piskorski, and Seru, 2018). The top panel of Figure 1 shows that mortgage companies originated over $1.3 Trillion in new mortgages in 2019, outpacing all depository institutions combined (banks, credit unions, etc.). The bottom panel of Figure 1 shows that mortgage companies now account for roughly 53% of mortgage lending in the U.S., and that they employ 38% of all mortgage loan officers.

[Insert Figure 1 Here]

Loan officers serve as the primary point of contact for customers looking to take out a mortgage. An officer's duties include presenting information about mortgage products and pricing to prospective applicants, answering their questions, aiding applicants in filling out applications, and following up with applicants regarding documentation. Loan officers can likely increase their lending volume by developing skills and/or exerting effort along several dimensions. Similar to a salesperson, loan officers can potentially increase their volume through referrals, either via

[10] For example, evidence on CEO gender and firm performance is mixed (see, e.g., Farrell and Hersch, 2005; Adams and Ferreira, 2009; Post and Byron, 2015). A notable exception is Huber, Lindenthal, and Waldinger (2021) who show that firms that lost Jewish managers during the rise of Nazi Germany experienced reductions in stock prices, dividends, and returns on assets.

satisfied customers or through connections to intermediaries like real estate agents and brokers. After the first contact with a prospective applicant, loan officers can likely increase the share of deals they convert by improving their customer service, their diligence in following up with applicants, their aid to marginal applicants, and their application processing speed.

An important advantage of this setting is that loan volume is a straightforward, yet convincing measure of loan officers' productivity. Three factors support this view. First, most mortgage approval and pricing decisions at the companies we study are made by underwriting algorithms, leaving loan officers with a limited role in setting prices, and placing their focus on generating volume. Second, the mortgage companies we study sell close to 100% of the loans they originate on the secondary market, leaving them with minimal exposure to defaults, and reducing the impact of loan quality on their profitability.[11] Finally, industry practices for compensating loan officers support the view, because pay is typically linked to loan volume but not to other factors.[12] Loan officers' incentive pay is usually either a percentage of their dollar volume, or a fixed amount per transaction. Therefore, our main analyses use loan volume in terms of number or dollar amount as our measures of loan officer productivity.

Branch managers at the mortgage companies we study are tasked with monitoring their subordinate loan officers' effort, productivity, and compliance with company policies. The most successful managers likely increase the productivity of their subordinates by creating a culture conducive to success, offering advice, or directly aiding their loan officers in generating business. Fortunately, our data linking managers to their subordinates allow us to compute ex post measures of manager's performance based on their impact on subordinates' productivity (measured via changes in subordinates' productivity under the manager in question relative to other managers). We are then able to use this measure of manager performance to conduct outcome tests for gender bias in promotions.

[11] We collect HMDA data on all mortgages originated by mortgage companies during our 2014-2019 sample period and find that they sell 93.4% of loans by the end of the calendar year of origination (when HMDA reporting on loan sales ends). This statistic understates the true percentage of loans sold, because loans can be sold after this point in time, especially for loans originated toward the end of the calendar year.

[12] In fact, since April 2011, loan officers have been prohibited by law from receiving compensation based on loan terms such as the interest rate or fees charged at origination. The law makes an exception in order to allow compensation based on the loan amount. These restrictions were first put in place by the Loan Originator Final Rule (75 FR 58509) issued by the Board of Governors of the Federal Reserve as part of their historical responsibility to implement the Truth in Lending Act (TILA). The Dodd-Frank Act has since codified the restrictions into law, and transferred authority to implement the TILA to the Consumer Financial Protection Bureau, who has continued to enforce the restrictions.

3. Data

We combine data from several sources to conduct our empirical tests. We start by building the first nationwide panel of mortgage loan officers and branch managers at mortgage companies based on licensing and registration information from the NMLS. Next, we put extensive effort into identifying the gender of loan officers and managers. We then merge on data from CoreLogic with the loans originated by the loan officers, and information on foreclosures from Zillow. We discuss these primary data sources and the key merges in more detail below.

3.1 Individual loan officer data

The Secure and Fair Enforcement for Mortgage Licensing Act of 2008 (SAFE Act) was designed to enhance consumer protection and reduce fraud in the mortgage market. The law requires that all residential mortgage loan originators (i.e., loan officers) who are employed by federally insured depository institutions, credit unions, or their subsidiaries must be federally registered. All other loan officers, e.g., those working at mortgage companies, must be state-licensed. Importantly, the SAFE Act requires that all loan officer licenses and registrations must be recorded in the Nationwide Mortgage Licensing System (NMLS).[13] By 2012, all state and federal regulators had implemented licensing/registration regimes that conform with national standards and integrated into the NMLS, making it a comprehensive registry of mortgage lenders and their loan officers.

We obtain access to data from NMLS Consumer Access[SM] through an agreement with the State Regulatory Registry, a subsidiary of the Conference of State Bank Supervisors (CSBS) tasked with operating the NMLS.[14] Specifically, we obtain historical snapshots of the files with information on licenses, registrations, and other fields for individual loan officers. These snapshots are taken at the end of each calendar year from 2012 to 2019. NMLS assigns each loan officer a unique NMLS ID that stays with them over time and across employment spells, allowing researchers to accurately track people throughout their career in the mortgage industry.

[13] The NMLS was created in 2008 by the Conference of State Bank Supervisors (CSBS) and the American Association of Residential Mortgage Regulators (AARMR), see: https://nationwidelicensingsystem.org.
[14] For additional information on NMLS Consumer Access[SM], see https://nmlsconsumeraccess.org/.

Most important for this study is the information on loan officers' employment histories and the managers at each branch. Whereas banks and credit unions only register their loan officers, state-licensed institutions (i.e., mortgage companies) report the names and NMLS IDs of both their loan officers and the managers at each branch. From these files, we are able to construct a national panel of mortgage loan officers, with information on their employment history, employer information, specific worksite, and for mortgage companies, whether they are managers and subordinate-manager links.

3.2 Loan-level data

3.2.1 Mortgage transactions

We obtain mortgage transaction data from CoreLogic's Mortgage Basic database. CoreLogic is recognized as the premier provider of real estate and mortgage transaction data, and Mortgage Basic covers nearly all residential mortgages in the U.S. starting in the early 2000s. This dataset provides us with borrower information, property information, basic mortgage characteristics, and most importantly, an identifier for the loan officer (NMLS ID) associated with each transaction starting in 2014.

In order to assess loan officers' productivity, we aggregate their mortgage transactions to the yearly level to match the frequency of our NMLS panel. Specifically, we count the total number of loans made and calculate the total dollar amount of mortgages for each loan officer-year. We then merge the CoreLogic information onto our NMLS loan officer panel by NMLS ID and year. The merged dataset includes each loan officer's employment history and sales performance from 2014 to 2019. In Table 1 Panel A, we report the number of loan officers and the number of managers in columns 1 and 4, respectively. The number of loan officers increases from around 20,000 in 2014 to more than 33,000 in 2019. The number of managers also increases by a similar magnitude from 4,300 in 2014 to more than 6,800 in 2019.

3.2.2 Foreclosures

We obtain foreclosure information from Zillow's Transaction and Assessment Database (ZTRAX). ZTRAX collects transaction level data on foreclosures, including property identifiers, delinquency date, unpaid loan balance, etc. (Zillow, 2021). We merge ZTRAX to the CoreLogic data using property identifiers, i.e., state and county FIPS codes and parcel IDs. This procedure

allows us to link the mortgage transactions from CoreLogic for each loan officer to the foreclosure variables from ZTRAX. We again aggregate the foreclosure information to the yearly level by loan officer. Specifically, we count the number of foreclosures and calculate the total dollar value of unpaid balance based on the year of loan origination.

3.3 Identifying loan officer and manager gender

Another important piece of information that is essential for our study is the gender of loan officers and branch managers. The NMLS data do not provide information on gender. To remedy this issue, we identify gender using a combination of name-based methods and extensive hand-searching.

We start by identifying loan officer and manager gender based on their first names as reported in the NMLS. We standardize and clean the first names and then match them with a list of the most popular first names by gender between 1950 and 2010 published by the Social Security Administration, following Niessen-Ruenzi and Ruenzi (2019).[15] This procedure allows us to code gender for around 80% of the people our data. Subsequently, we have multiple researchers and research assistants look at each of the most popular first names that have not yet been assigned a gender, following Aggarwal and Boyson (2016) and Niessen-Ruenzi and Ruenzi (2019). If each researcher/assistant agrees the name can unambiguously be associated with a gender, we code gender accordingly. At this point, we have identified gender for roughly 90% of the people in our sample. For the remaining loan officers and managers, we carry out an extensive web searching effort to determine their gender based on LinkedIn profiles, company websites, and other material available on the Internet, following Barber, Scherbina, and Schulusche (2017).

3.4 Other data

We take several additional steps to prepare our dataset. First, we match the lending institutions in our NMLS database to lenders in the Home Mortgage Disclosure Act (HMDA) data based on company names and addresses. This merge helps us verify the comprehensiveness of the NMLS data—we are able to match over 97% of HMDA lenders to NMLS. This merge also allows

[15] For the list of names, see https://www.ssa.gov/oact/babynames/decades/index.html

us to compute the descriptive statistics in Figure 1 showing mortgage companies' market share from the HMDA data.

To hold other demographics constant while studying gender, we identify loan officers' and managers' race and ethnicity. Following Ambrose, Conklin, and Lopez (2021), we assign race and ethnicity using the Bayesian Improved First Name Surname Geocoding (BIFSG) method, which is based on each individual's first name, last name, and location. Throughout our analyses we control for *White*, an indicator for the person being non-Hispanic and white.

3.5 Summary statistics

Our final dataset covers over 72,000 unique loan officers working at over 1,000 lending institutions between 2014 and 2019. Table 1 Panel A shows that we are able to identify the gender for 99.4% of the loan officers and 99.9% of the managers (see columns 2 and 5). Columns 3 and 6 report the gender distribution by year. For loan officers, the fraction of women increases from 34.4% in 2014 to 36.4% in 2019. For managers, the fraction of women increases from 26.3% to 28.7% during the same period. Two points stand out from these statistics. First, women are underrepresented among loan officers compared to their 50.8% share of the population (2010 Census). And second, there is a non-trivial difference (approximately eight percentage points) between the fraction of female loan officers and female managers. These statistics show that gender disparities appear even in the first tier of management at financial firms.

Our sample includes 2,037 promotions from loan officer to branch manager at mortgage companies between 2014 and 2019. Table 1 Panel B presents the summary statistics for our main loan officer-year panel dataset. The average promotion rate is around 1.2% per year. The average loan officer has worked at their firm for just over 2 years, and originates 37 loans totaling 8.5 million dollars in a year. See Appendix A for the definition of each variable used in the analysis.

[Insert Table 1 Here]

4. The gender promotion gap

Our univariate findings show an approximately eight percentage point difference between the fraction of female loan officers that are subsequently promoted to managers, suggesting that female workers begin to lose out to male workers at the very first step of the corporate ladder.

However, this finding does not control for observable differences in employees that are likely to influence promotion decisions, such as employee performance. Therefore, in our first test, we examine how the gender of mortgage loan officers affects their likelihood of being promoted, using our loan officer-year panel dataset and equation (1):

$$Promoted_{i,t} = \beta_0 + \beta_1 Female_i + \beta_2 \times X_{i,t} + \eta_{j,t} + \eta_k + \varepsilon_{i,t,} \qquad (1)$$

We estimate equation (1) using OLS regressions, in which i indexes individual loan officers, t indexes year, j indexes firm, and k indexes branch. The dependent variable, *Promoted*, is an indicator variable for the loan officer being promoted in the following year. The independent variable of interest, *Female*, is an indicator variable for the loan officer being a woman. We include controls for loan officer performance using both overall and relative measures. The overall performance measure is the loan officer's total number of loans originated (*Log(number of loans)*), which mirrors how loan officers are compensated. The relative performance measure is an indicator variable that equals one if the loan officer is ranked number one in the number of loans originated at their branch, which controls for the possibility that promotions are a tournament between loan officers.[16] We control for worker tenure using the number of years the loan officer has worked at the firm and control for worker race/ethnicity using an indicator variable that takes on the value one if the loan officer is non-Hispanic and white. Combinations of firm-year and branch fixed effects are included.

Table 2 presents results from three estimations of equation (1) that vary by the number of included control variables and fixed effects. Column 1 includes the *female* indicator variable and firm-year fixed effects, while column 2 adds controls for performance and race/ethnicity, and column 3 further adds branch fixed effects. Across all three specifications, we find that mortgage companies are less likely to promote female loan officers, revealing an economically significant gender promotion gap. For example, column 3, which includes our full set of controls and fixed effects, shows that female loan officers are 0.17 percentage points less likely to be promoted each year than comparable male officers. This represents a 15% decrease relative to the base rate of promotion.

[16] We obtain similar results if we use total loan amount to compute performance measures, but do not include them together as the correlation between the number of loans and loan amount is 0.96.

Inspection of the control variables shows that the coefficients on both relative and absolute performance are positive and significant, indicating that more productive loan officers are more likely to be promoted.[17] We also find some evidence that longer tenure is negatively related to being promoted, though the coefficient is only significant in column 2. The coefficient on *White* is also positive and significant, documenting that white loan officers are more likely to be promoted to manager. Taken together, the results of Table 2 provide new large sample evidence of a significant promotion gap for female loan officers at financial institutions: women are approximately 15% less likely to be promoted to managers, even after controlling for observable characteristics such as job performance.

[Insert Table 2 Here]

5. Gender bias in promotions

Our results so far show that women face a significant gender disparity in promotions at financial institutions. However, disparities alone do not prove bias. Even with a rich set of controls, standard OLS regressions cannot completely rule out omitted variables or statistical discrimination as explanations for the gender promotion gap. To isolate bias from these alternatives, we implement a Becker outcome test based on managerial performance. Our test directly evaluates whether firms equate expected performance for marginally promoted men and women, i.e., whether they set the same promotion standard for men and women.

We start this section by outlining our empirical approach for estimating outcomes at the margin of promotion. We then present our main Becker outcome test for gender bias in promotions. Next, we investigate the mechanisms behind gender bias by examining the cross-sectional variation in the bias we document. Finally, we discuss whether any salient alternatives to gender bias could explain our results, and conduct robustness tests.

5.1 Empirical approach

[17] The coefficient on *Top LO at branch* switches to being negative in column 3 when branch fixed effects are included. However, this is due to the fact that *Log(number of loans)* captures relative performance in a continuous fashion once branch fixed effects are included.

To implement Becker's outcome test for gender bias in promotions, we need to be able to (1) accurately measure managers' performance, and (2) estimate performance for marginally promoted male and female managers. To measure managerial performance, we employ a method similar to those used to compute "value added" by workers, teachers, and sales managers. We then take advantage of a new approach that utilizes the local nature of instrumental variables estimators to estimate outcomes at the margin. We discuss each of these aspects of our empirical design in more detail below.

5.1.1 Measuring the quality of newly promoted managers

To create the relevant outcome measure of managerial performance, we need a measure that reflects the responsibility that branch managers face to increase their subordinate's loan originations. It is, however, important to recognize that managerial performance measures can be biased by the non-random assignment of managers and subordinates. For example, if a manager supervises a particularly weak (strong) team of subordinates, then the weak (strong) performance of the team could be incorrectly credited to the manager's performance measure. Fortunately, significant progress has been made in "value added" measures of the performance of workers (e.g., Abowd, Kramarz, and Margolis, 1999), teachers (e.g., Chetty, Friedman, and Rockoff, 2014a, 2014b), and sales managers (Benson, Li, and Shue, 2019). Following these studies, we compute our measure, *Managerial effect*, by explaining each subordinate's loan volume (i.e., their *Log(number of loans)* or *Log(loan amount)*) using the following regression:

$$Loan\ volume_{i,m,j,t} = \beta_0 + \eta_i + \eta_m + \eta_{j,t} + X_{i,t} + \varepsilon_{i,t,} \qquad (2)$$

where the dependent variable is the loan volume of loan officer i under manager m at firm j in year t. We regress the dependent variable on loan officer fixed effects (η_i), manager fixed effects (η_m), firm times year fixed effects ($\eta_{j,t}$), and loan officer tenure ($X_{i,t}$).

Equation (2) estimates *Managerial effect* via the manager fixed effects, η_m. Since the manager's fixed effect is the mean change in loan originations across all loan officers who either join or leave that manager's supervision, a manager's effect on subordinates is determined by deviations from the subordinates' mean loan originations under all managers. This circumvents the potential bias due to non-random assignment of managers and loan officers, since managers are

only evaluated on the changes in the performance of their subordinates. In addition, firm-year fixed effects control for firm-time specific effects, as well as macroeconomic and industry conditions that could impact subordinates' loan origination performance. *Tenure* controls for potential returns to experience for loan officers.

We estimate equation (2) based on all loan officer-years and managers. The estimated manager fixed effects represent the average, time-invariant component of a manager's performance, and have a mean of zero by construction. We then collect the estimated manager effects for the 2,037 newly promoted managers during our sample. Taking managers' effect on their subordinates' *Log(number of loans)* as an example, the 25th percentile is -0.68, implying that when a loan officer is assigned to a 25th percentile manager, her output is 51% ($e^{-0.68} = 0.51$) of what it would have been if assigned to an average manager. Similarly, the 75th percentile is 0.40, implying that the output of a loan officer assigned to a 75th percentile manager is 149% ($e^{0.40} = 1.49$) of their output under an average manager.

5.1.2 Instrumental Variables approach to implement Becker's outcome test

To implement Becker's outcome test, we need to estimate the performance of marginally promoted male and female managers. Recent work by Arnold, Dobbie, and Yang (2018) demonstrates how the local nature of instrumental variables estimates can be used to estimate outcomes at the margin. The approach takes advantage of work by Angrist, Imbens, and Rubin (1996) showing that IV estimates represent the local average treatment effect for a subgroup of units—the compliers to the instrument. In our setting, we instrument for loan officers being promoted to manager. The compliers—loan officers who would not have been promoted except for the instrument—can be considered the set of marginally promoted managers. We can therefore use the LATE property of IV estimates to directly estimate outcomes that are local to (at) the margin.

Our instrumental variables design closely follows the setup in Benson, Li, and Shue (2019), who identify marginally promoted sales managers and show that companies are biased toward promoting salespeople with high prepromotion sales but low managerial ability (the Peter Principle). Although our study focuses on a different research question, in a different industry, the direct parallels between our data structure and measure of managerial performance and theirs make it natural to utilize their IV design. As in Benson, Li, and Shue (2019), we use the "leave-out"

average promotion rate at the firm, which excludes the focal worker and her teammates, to instrument for whether the worker is promoted to manager. Intuitively, the compliers to this instrument can be thought of as marginal managers because they were promoted due to the firm's promotion rate being high, but would not have been if it were lower.

The approach from Benson, Li, and Shue (2019) has several important strengths. First, as we discuss in more detail below, the specific form of variation that the instrument exploits (firm-by-time level variation) pairs well with the methodology used to compute managers' value-added, making the exclusion restriction plausible. For instance, one might be concerned that high promotion rates at a firm coincide with positive demand shocks, which could affect all workers' productivity and thus be correlated with managerial performance. However, the methodology used to compute managers' value-added controls directly for firm-by-time fixed effects, eliminating this concern. Second, the instrument can induce variation in promotion decisions within decision-makers (decisions during high vs. low promotion rate years). This type of variation parallels the experimental ideal for testing for biased decision-making, and obviates the need for additional assumptions about homogeneity/consistency in decision-makers' behavior that are required in settings like Arnold, Dobbie, and Yang (2018), where the IV design uses only across-decision-maker variation and the quasi-random assignment of decision-makers.[18]

We implement our instrumental variables tests using a two-stage least squares (2SLS) approach with the following specifications:

$$\text{Stage 1:} \quad Promoted_{i,t} = \beta_0 + \beta_1 Leave-out\ measure_{i,t} + \beta_2 \times X_{i,t} + \varepsilon_{i,t,} \qquad (3)$$

$$\text{Stage 2:} \quad Managerial\ effect_{i,t} = \beta_0 + \beta_1 \widehat{Promoted}_{\iota,t} + \beta_2 \times X_{i,t} + \varepsilon_{i,t,} \qquad (4)$$

where i indexes individual loan officer and t indexes time. These tests are run using the sample of loan officer-years (which excludes manager-years). For loan-officer years that are not immediately followed by a promotion to manager, the *Managerial effect* is set to zero. For loan officer-years

[18] The ideal experiment in our setting would compel a firm to increase or decrease promotions from a random set of male and female loan officers, therefore enabling a comparison of managerial ability for promoted (treated) versus non-promoted (control) within gender. For small changes in the promotion rate, finding a larger managerial effect for women would indicate bias in favor of male loan officers. Similar to this idealized within-firm experiment, our instrument also exploits within-firm treatment rate variation by inducing firms to increase or decrease the promotion rate among the group of eligible male and female loan officers.

where the person is promoted to manager the following year, *Managerial effect* takes the value of their manager fixed effect estimate from equation (2). Therefore, our estimate of β_1 in the second stage is an estimate of the effect of promoting the loan officer in question to manager on their *Managerial effect*, relative to not promoting them and having it remain at zero. Given that IV estimates represent the LATE, β_1 can be interpreted as an estimate of the treatment effect on the compliers to the instrument, or in this case, the local average *Managerial effect* of the marginally promoted managers during our sample period. With this approach, we estimate the quality of marginally promoted male and female managers separately, and test for differences at the margin (gender bias).

To provide a valid Becker outcome test, our IV design needs to satisfy the standard IV requirements. The instrument must be significantly and monotonically correlated with the endogenous variable (instrument relevance), and the exclusion restriction must hold.

Instrument Relevance

In our setting, the instrument (the leave-out firm promotion rate) must be correlated with the endogenous variable we are instrumenting for (the individual's promotion). Table IA1 in the Internet Appendix reports our first-stage regressions of promotion on the leave-out instrument. The results show that the instrument strongly predicts promotion in the full sample, as well as when we split the sample by gender. Furthermore, the first-stage F-statistics (reported in our second-stage tables for convenience) are well above 10, showing that we are not working with weak instruments (Staiger and Stock, 1997).

The Exclusion Restriction

The second requirement for our IV design is the exclusion restriction, which states that the instrument cannot be correlated with unobserved determinants of the outcome variable. In other words, the instrument must affect the outcome only through the channel of the endogenous variable. In our setting, the leave-out firm promotion rate must affect the manager's value-added (*Managerial effect)* only through the channel of getting the person promoted (*Promoted).* As in all IV designs, this exclusion restriction cannot be tested directly. However, we outline the key factors that make this identifying assumption plausible in our setting (which parallel those in Benson, Li, and Shue, 2019), and we offer indirect evidence supporting the assumption.

Much of the strength of our IV design stems from the relationship between the specific variation that the instrument uses, and the methodology used to compute managers' value-added.

First, we note that we compute managerial performance using firm-year fixed effects. This controls for the most obvious challenge to the exclusion restriction: firm-level factors that lead to changes in the performance of newly promoted managers. For example, if mortgage loan demand is high and the firm responds by promoting more managers, these managers could get credit for high managerial performance, although it was caused by the time trend. Including firm-year fixed effects in calculating managerial performance prevents this issue. Second, the inclusion of loan officer fixed effects when computing managerial performance prevents any correlation between the instrument and the fixed quality/talent of the subordinates that the manager is supervising (we also control directly for loan officer experience, which increases over time). Third, we note that by construction, our instrument is not affected by reverse causality, where firms might decide to promote at a higher rate in order to promote a particularly strong worker. This is because by employing a "leave out" approach that omits a loan officer's own promotion status and her teammates', the correlation between the instrument and any particular worker's strength is zero. Thus, the methodology used to compute our outcome variable, *Managerial effect*, mitigates the most salient challenges to the exclusion restriction.

We also provide evidence, to the extent possible, suggesting the exclusion restriction is valid by showing that the leave-out instrument is uncorrelated with managerial performance, as well as with other factors that could be related to promotions and performance. Table IA2 in the Internet Appendix shows that the instrument is uncorrelated with *Managerial effect*, as well as with the firm's number of employees and loan volume.

5.2 Becker's outcome test for gender bias in promotions

If firms do not exhibit gender bias, they will equate expected manager performance at the margin, creating the testable prediction that managerial performance will be identical for marginally promoted male and female managers. We report the IV estimation results of equation (4) in Table 3. The dependent variables in columns 1 and 2 and columns 3 and 4 are the managerial effects on subordinates' number of loans and on the total loan amount, respectively. The key independent variable is *Promoted*, an indicator that equals one if the loan officer is promoted in the next year and zero otherwise. In the first stage (equation 3), we instrument for *Promoted* with our leave-out measure of the average promotion rate of loan officers at the same firm in the same year, excluding the focal loan officer and other officers working at the same branch. We also

include a host of control variables, including loan officer tenure, top loan officer at the branch, number of loans, and race, which are defined in Appendix A.

Column 1 of Table 3 shows that the effect of male managers on their subordinates' loan origination is negative and statistically significant. The point estimate, i.e., 0.0308, suggests that marginally promoted male managers decrease the number of loans their subordinates originate by 3% relative to the average manager. In contrast, Column 2 shows marginally promoted female managers have a positive and significant effect on their subordinates' loan originations, representing a 2% increase in the number of loans relative to an average manager. Looking at columns 1 and 2 together, the difference in subordinate's performance between male and female marginally promoted managers is approximately 5%, which is economically large and statistically significant at the 1% level. When we compute managerial quality using total loan amount instead of the number of loans, we find similar results. Column 3 shows the coefficient on *Promoted* is -0.0400 for male managers and column 4 shows the coefficient on *Promoted* is 0.0206 for female managers, both are statistically significant. The difference between those two columns suggests that a wedge in performance between loan officers working under marginally promoted male and female managers is approximately 6%, which is economically nontrivial and statistically significant at the 1% level.[19]

[Insert Table 3 Here]

Overall, the findings in Table 3 provide evidence that marginally promoted men perform significantly worse as managers than their female counterparts. In Becker's (1957, 1993) framework, this test shows that firms set the bar for promotion higher for female workers. Taken together with the "promotion gap" that female employees are promoted at a lower rate than male loan officers, our results here provide strong evidence of gender bias in promotion decisions at mortgage companies.

[19] The Internet Appendix Table IA3 presents OLS results of the average managerial performance between promoted men and women, conditional on observables. In conjunction with Table 3's findings based on marginal effects, these average OLS estimates imply that the marginally promoted male manager is a worse manager than the average promoted male manager, while the marginally promoted female manager is a better manager than the average promoted female manager. These results suggest that financial institutions make substantial errors in predicting managerial quality for female workers.

5.3 In-group tests: The role of managers' gender in promotion equity

Our findings so far are consistent with Becker's (1957, 1993) model of discrimination, that financial firms either knowingly or unknowingly exhibit bias against female workers at the margin of promotion. While this bias can be the result of inaccurate gender stereotypes, it can also result from explicit animus against female workers. Explicit animus against female workers by their male managers is likely to be especially detrimental to women in finance. Given the relative lack of women in leadership roles at financial firms, female workers will be increasingly likely to be supervised by male managers as they move up the corporate ladder. Therefore, it is perhaps not surprising to see the popularity of policies attempting to alleviate gender discrimination by increasing the number of female managers. However, it is important to note that such policies will be most effective if female managers are less biased against female subordinates, which is not obvious. For example, biased stereotypes regarding women's managerial ability can be held by both male and female managers (Bordalo et al., 2016), resulting in all managers being biased against promoting women. Therefore, in this subsection we extend our analysis to "in-group" tests, following previous studies such as Egan, Matvos, and Seru (2020) who argue that in-group tests can provide additional evidence on the mechanisms behind gender bias.[20]

To perform our in-group analysis, we further partition our outcome test by the gender of the manager whom the marginally promoted manager worked under prior to being promoted. The intuition behind these tests is to determine if male managers are more biased against female workers than female managers. If so, this would provide additional evidence of explicit gender animus towards female subordinates, as most models of prejudice would predict such an outcome. Table 4 reports the results of IV regressions. Panel A reports results based on managerial effect on the number of loans their subordinates originate, while Panel B computes the managerial effect based on the dollar volume of loans originated. The key variable of interest is the coefficient on *Promoted*.

Columns 1 and 2 of Panel A report in-group tests for marginally promoted managers who previously worked under male managers. Column 1 of Panel A shows that marginally promoted

[20] In-group tests are also used in many racial bias studies, including settings within a Kenyan flower packing firm (Hjort, 2014), and a French grocery chain (Glover, Pallais, and Pariente, 2017).

male managers who previously worked under a male manager have a negative and significant managerial effect on their subordinates. The point estimate suggests that marginally promoted male managers whose previous managers are also male decrease the number of loans their subordinates originate by almost 4% compared to the average manager, which is economically large and statistically significant at the 5% level. In contrast, column 2 of Panel A shows that the managerial effect of marginally promoted female managers who previously worked under a male manager is non-negative. Taken together, the managerial effect for marginally promoted managers who work for male managers is 4% smaller for men than women, which is statistically significant at the 5% level. This provides evidence that male managers hold female workers to a higher promotion standard than their male counterparts. That is, male managers have an in-group tolerance for male subordinates.

Columns 3 and 4 of Panel A report in-group tests for marginally promoted managers who previously worked under female managers. We find that under female managers, marginally promoted male managers perform worse than marginally promoted female managers. More specifically, subordinate's performance under marginally promoted managers whose previous manager is female is 5% higher for women than men, for which a difference in the last row of Table 4 panel A shows is statistically significant. This provides evidence that female managers also hold female workers to a higher promotion standard than their male counterparts. Panel B of Table 4 replicates Panel A by computing managerial performance using loan amount and shows similar results.

[Insert Table 4 Here]

Taken together, the results contained in Table 4 show that the gender bias against marginally promoted female managers is similar under previous male and female managers. Given that such findings are generally inconsistent with most models of taste-based discrimination, these results suggest either that gender animus is not driving our results, or that male and female managers hold equal levels of gender bias against female underlings. Similar conclusions have been reached in studies of racial bias. For example, studies find that racial bias against black defendants is similar under white and black judges (see e.g., Anwar and Fang, 2006). Given we find that female managers are also biased against female subordinates, an important policy

implication of our results is that policies intended to reduce gender discrimination by increasing the number of women in upper management are unlikely to eliminate the bias women face at financial institutions.

5.4 Gender and the Peter Principle

A consistent finding throughout our tests is that marginally promoted male managers perform relatively poorly compared to their female counterparts. Our next set of tests explore whether an interplay between gender and the Peter Principle contributes to the gender bias we document.

Table 5 reports outcomes tests that examine by gender, the managerial performance between marginally promoted managers with high and low sales prior to promotion. The last row of each panel of Table 5 reports significance levels for the difference between high and low sales managerial performance. For male loan officers, the coefficients on *Promoted* in columns 1 and 2 of Panel A, combined with the test of the differences, show that when high sales male employees (above median number of loan originations) are promoted to manager, they make significantly worse managers than promoted low sales male employees (below median number of loan originations). The difference between the coefficients on *Promoted* for low and high sales male employees is statistically significant and represent a 5% worse managerial effect on subordinate's performance for high sales male employees compared to low sales counterparts. This finding provides evidence that the Peter Principle applies to male workers at mortgage companies, extending to financial institutions the limited outcome test-based evidence on the Peter Principle.

In sharp contrast to the evidence supporting the Peter Principle in male loan officers, columns 3 and 4 of Table 5 show that when high sales female loan officers are marginally promoted to manager, they do not make worse managers than their low sales counterparts. The coefficients on *Promoted* actually suggest that high sales female loan officers make *better* managers than marginally promoted low sales female managers, though the last row of Panel A shows that we cannot reject the hypothesis that they are equal (p-value=0.2148). In Panel B of Table 6, we replace the number of loans outcome measure with loan amount and finds consistent, if not stronger, results that the Peter Principle does not apply to female workers in our sample.

[Insert Table 5 Here]

Taken together, the results in Table 5 show that in financial institutions, promoting the best salesman at the margin results in lower managerial performance (the Peter Principle) while in contrast, promoting the best saleswoman at the margin does not sacrifice managerial performance (a finding we denote as the "Penelope Principle"). Overall, the results provide evidence that the Peter Principle is not monolithic, but instead differs significantly by gender. Therefore, our results suggest another positive aspect of female leadership: unlike male workers, the incentive benefits of basing promotions on sales performance for female workers is not associated with promoting workers with lower managerial potential.

5.5 Discussion of alternative explanations and robustness tests

In this subsection, we discuss the most salient remaining alternatives to gender bias—those alternatives which, if true, could generate the differences in marginal male/female manager performance we document. We then empirically investigate the merits of these alternative explanations in robustness tests. We focus on three alternatives to bias: that female managers underperform on non-sales dimensions of the job offsetting their better sales performance (at the margin), that the distribution of female managerial potential strictly dominates the male distribution, or that women with lower managerial potential systematically turn down promotion offers that their male counterparts accept.

We start by examining whether marginally promoted female managers underperform on non-sales dimensions of the job, such as the quality of the loans their subordinates originate. We expect non-sales dimensions to be second order concerns compared to sales (lending volume), because the firms we study typically focus solely on mortgage origination and they sell nearly all of the loans they originate on the secondary market, which makes volume the primary driver of profitability.[21] However, if loans made under marginal female managers were of substantially worse quality than those under their male counterparts, it could be costly to the firm either through its reputation on the secondary market or other channels such as disputes with borrowers.

[21] Loan quality is not a metric that is typically used to compensate the loan officers we study. Moreover, in untabulated tests, we find that loan quality is not significantly related to promotion when included in Table 2 specifications, providing additional evidence that it is not a point of emphasis at mortgage companies.

Fortunately, we can directly test for differences in loan quality (default rates) under marginal male/female managers.

Table 6 presents IV tests similar to those in Table 3, except that the outcome variables are measures of managers' impact on their subordinate officers' loan quality (rather than quantity). Columns 1 and 2 report results where managerial performance is computed based on subordinates' foreclosure ratio, defined as the percentage of loans originated during the year that end in foreclosure. Columns 3 and 4 report results where managerial performance is calculated using the foreclosure amount, defined as the ratio of the total dollar amount of loans ending in foreclosure over the total dollar amount of loans originated in the year. The results in Table 6 show that loan quality is not worse under marginally promoted female managers, providing evidence that female managers are not underperforming on non-sales dimensions. The combination of higher lending volume and similar loan quality that we document provides strong evidence that the marginally promoted female manager outperforms her male counterpart.

[Insert Table 6 Here]

With a properly defined outcome and a valid IV approach, Becker's outcome test will convincingly isolate bias from alternatives by directly estimating outcomes at the margin (and testing for differences across groups). However, an implicit assumption here is that the margin, or in our case the "promotion bar," is truly being set by firms. The two remaining alternatives to bias that we discuss represent scenarios where this assumption could be violated.

We first investigate an extreme case where the margin is not set by the firm, but by the unconditional talent distributions of male and female workers. Consider the possibility that the distributions of managerial potential are non-overlapping for male versus female loan officers, with women having strictly more potential. Facing these talent distributions, even if firms promoted every woman, the last (marginal) woman promoted would outperform all male managers (and especially the marginal male manager). This could lead to the marginal female manager outperforming her male counterpart (absent bias). While this extreme case could generate differences at the margin like we find, it is clearly an unlikely scenario, and it is strongly at odds with the gender gap in promotions we document in Table 2. Although we do not observe the underlying distributions of managerial potential, we can directly observe the distributions of male

and female managers' ex post performance. We plot these distributions in Figure IA1 in the Internet Appendix, both for the sample of all managers as well as those managers promoted during our 2014-2019 sample period. In both instances, we find the managerial effects by gender overlap heavily and are similar, ruling out this extreme case as an explanation for our results.

Finally, we consider a (slightly) less extreme case where the margin is set not by the unconditional talent distributions (as above), but by the talent distributions for the subset of workers willing to accept a promotion if offered one. Here we recognize the limitation that, similar to nearly all studies of promotions, we only observe accepted promotions, rather than all offers. However, an important advantage of our outcome tests is that so long as there are women (men) throughout the female (male) talent distribution that are willing to accept a promotion if offered one, then the marginal outcomes that we estimate will still be determined by the bar that firms set for making promotion offers. In order for the difference we document at the margin to be generated by worker preferences (rather than firms' biases) women with lower managerial potential would have to uniformly turn down essentially all promotion offers, until the lowest-potential woman willing to accept a promotion was above the firm's hypothetical gender neutral promotion bar.

We have not found any empirical or anecdotal evidence to suggest that women in finance are turning down promotions on a massive scale, or in the specific pattern outlined above. In fact, prior work suggests that women working in finance tend to be more career-oriented and less averse to competition than women in the broader population, greatly reducing gender differences along these dimensions (e.g., Fang and Huang, 2017). Therefore, a more plausible story is that a small percentage of women throughout the talent distribution may turn down a promotion (presumably due to family considerations) that a male counterpart would have taken. This pattern would leave the distribution of female talent for those willing to take promotions well populated throughout its support—leaving the margin to be set by firms' bar for making offers. Moreover, if firms face any pressure to maintain female representation in management, this pattern could push them to lower the bar for female promotions (e.g., to promote their second choice if their first turns it down). This would work directly against our results showing that women are held to a higher promotion standard.

In sum, the findings in this section show that marginally promoted female managers outperform their male counterparts by roughly 5%. Women appear to face similar levels of bias under both male and female decision makers, suggesting biased stereotypes about managerial

potential are the primary mechanism rather than gender animus. We also document that firms' bias toward overpromoting high performers relative to their managerial potential (the Peter Principle) is stronger for men, which contributes to overall levels of gender bias. Finally, we rule out salient alternative explanations for the differences we document at the margin, leaving gender bias as the most compelling explanation for our findings.

6. The economic consequences of firms' gender practices

In Becker's (1957) seminal theory, firms that forgo efficient employment practices due to biases incur costs that can destroy value.

If an individual has a 'taste for discrimination,' he must act as if he were willing to pay something either directly or in the form of a reduced income, to be associated with some persons instead of others. When actual discrimination occurs, he must, in fact, either pay or forfeit income for this privilege (Becker 1957:14).

However, it has been challenging for empirical researchers to document the costs resulting from gender bias at firms. In this section, we present four different tests, at both the worker and firm level, that investigate how underpromoting female workers to managers is costly to firms.

6.1 Manager gender and loan officers' performance

Our previous findings show that marginal female managers increase the productivity of loan officers under their supervision, whereas marginal male managers reduce productivity, suggesting that firms incur a significant opportunity cost when they overpromote male workers. In this section, we estimate these costs by investigating a potential channel through which female managers increase their subordinates' productivity. Prior research has argued that having more women in managerial roles creates an incentive for women to aspire to these positions (Blau and Kahn, 2017). If working under a female manager provides female loan officers additional incentives to become managers, we expect female loan officers to increase their productivity when working under a female supervisor. Therefore, we examine whether a loan officer's productivity depends on their gender and the gender of the manager who is supervising them. Our analysis is motivated by studies that examine such manager-subordinate gender effects on promotion (Cullen

and Perez-Truglia, 2018), termination (Egan, Matvos, and Seru, 2020), and performance evaluations (Benson, Li, and Shue, 2021).

We carry out a difference-in-differences test at the loan officer-year level. We focus on loan volume and quality in Panels A and B, respectively. Panel A of Table 7 reports OLS regression results examining the average effect of having a female manager on subordinate loan officers' performance. Columns 1 and 3 show that the average effect of female managers on all subordinates is positive but insignificant. Columns 2 and 4 show that this average effect masks significant heterogeneity based on subordinates' gender: male loan officers are slightly less productive under female managers (statistically insignificant), but female officers are considerably more productive. The regression coefficients on the interaction terms in columns 2 and 4 suggest that female loan officers working under female managers are 6% and 8% more productive in terms loan origination compared to their male counterparts who also work under female managers, respectively. Panel B shows that this does not come at the expense of decreased loan quality. Overall, the results in Table 7 document that, on average, female loan officers' productivity increases when they work under female managers compared to working under male managers. Taken together with our prior results that firms underpromote female workers to managers, these results suggest significant costs to firms that remain biased against promoting women.

[Insert Table 7 Here]

In additional tests, we examine whether female managers affect the composition of their subordinate loan officers' lending portfolios in terms of the fraction of loans made to women or low-income borrowers. Table IA4 in the Internet Appendix presents these results. The point estimates suggest that female managers increase lending to women, particularly from their male subordinates, but these effects are statistically insignificant. We also find a positive but statistically insignificant overall effect of female managers on lending to low-income borrowers, except among female loan officers, who increase their share of lending to low-income borrowers by around 2% under female managers. Overall, these results suggest that greater gender equity in promotions at financial institutions would likely benefit their female and low-income customers the most, but that these distributional effects would be modest, at least in the context of mortgage lending.

6.2 Manager gender and firm performance

Our next set of firm-level tests investigate whether firms that promote relatively more women to branch managers have higher performance. Our firm-level proxy for the representation of women managers in the firm, *% female manager/% female loan officer*, is the ratio of the percentage of managers who are female over the percentage of loan officers who are female at the mortgage company. We exploit two firm-level performance measures from our data, the aggregate number of loans and the aggregate dollar amount of loans per firm. While we estimate specifications that include year fixed effects as well as year and firm fixed effects, it is important to note that these firm-level tests are not designed to distinguish between treatment and selection effects.

Panel A of Table 8 shows that across specifications and performance measures, firms that have a higher percentage of female representation also have a higher amount of loan originations. Further, comparing models that include firm fixed effects to those without, this pattern holds both across firms, and within firms over time. The results are also highly economically significant, taking column 2 as an example, we find that a one standard deviation increase in female representation is associated with a 14% increase in the number of loans. Finally, Panel B of Table 8 shows that this increase in performance does not come at the expense of loan quality. Overall, these results are consistent with Becker's (1957) prediction that firms exhibiting bias (i.e., promoting relatively few women) incur significant costs in terms of fewer loan originations, the primary way mortgage companies generate cash flows.

[Insert Table 8 Here]

6.3 Manager gender, firm growth, and survival

Our previous results suggest that firms incur significant costs in terms of loan originations when they are biased against promoting women. Given loan originations are the main revenue source in mortgage companies, we next examine if firms with relatively fewer female branch managers have a lower growth rate and lower probability of survival. We are motivated by prior studies such as Faccio, Marchica, and Mura (2016) who show that firms with female CEOs are more likely to survive than those with male CEOs. We proxy for firms' growth potential using the *log (number of employees)*, and proxy for survival by using an indicator variable if the firm exits

the data the following year, *Firm exit next year*. Similar to our previous firm-level tests, we proxy for the representation of women managers in the firm using *% female manager/% female loan officer*. Table 8 presents the results.

Results for firm employment are shown in columns 1 and 2. Column 1 includes year fixed-effects and column 2 includes both firm and year fixed effects. Both tests show a positive relationship between the representation of women managers and the number of employees. For example, column 2 shows that positive changes in the relative number of female managers is associated with a positive change in the number of employees. The result is also economically significant – a one standard deviation increase in female representation translates to a 10% increase in employment. In columns 3 and 4, the dependent variable, *Firm exit next year*, is an indicator variable that equals one if the firm exits from the data in the next year and zero otherwise. In column 3, we find that firms with higher female representation are more likely to survive. In terms of economic significance, a one standard deviation increase in female representation is associated with a 10% decrease in the likelihood of exit. In column 4, which includes firm fixed effects, the results are insignificant, which is unsurprising given the limited within-firm time-series variation in the outcome variable. Overall, Panel C of Table 8 shows that firms with a higher representation of women managers not only grow faster, but are more likely to survive.

7. Conclusion

Because female leaders in financial firms are relatively rare, a strong presumption exists that women face discrimination that limits their climb up the corporate ladder. However, it is well-known that data limitations coupled with omitted variables and infra-marginality problems make it difficult to find convincing evidence of gender bias in promotion policies. We build a nationwide panel of mortgage loan officers and branch managers covering 72,000 workers from over 1,000 firms from 2014 to 2019 to document a significant gender gap in promotions: female loan officers are 15% less likely to be promoted than their male counterparts with similar experience and performance. We then use the two central predictions of Becker's (1957, 1993) model to test for gender bias at financial institutions: biased firms will (1) set a lower promotion bar for men, and (2) incur costs from forgoing efficient employment practices.

We find strong evidence that gender bias is an important driver of the gender promotion gap. First, outcome tests show that financial institutions demand better managerial performance

from marginally promoted women, showing that these firms set the promotion bar higher for women than men. Second, "in-group" tests show that female workers face bias from both male and female managers, consistent with inaccurate gender stereotypes (rather than animus). Third, firms' bias toward promoting the best men at the margin results in lower managerial performance (the Peter Principle), while in contrast, promoting more high-performing women at the margin does not sacrifice managerial performance (a finding we denote as the Penelope Principle). Finally, we find evidence consistent with Becker's (1957) prediction that biases are costly to firms: at the worker level, female loan officers' productivity decreases significantly when they work under a male supervisor. At the firm level, firms that promote relatively fewer women see a decrease in loan volume, lose employees, and are less likely to survive.

Our findings provide several new insights into gender bias in promotion policies. We provide direct evidence of bias in an industry that has received little attention in the academic literature, even though the finance sector employs millions of workers in highly compensated jobs that influence capital allocation in the economy. Our findings also speak directly to the biases women encounter at the very beginning of their career, a stage that has largely eluded researchers' study. Overall, we show that gender bias leads to a "Broken Rung" for women at the first step of the corporate ladder, long before women confront the better-known "Glass-Ceiling."

References

Abowd, John M., Francis Kramarz, and David N. Margolis, 1999, High wage workers and high wage firms, *Econometrica* 67, 251-333.

Adams, Renée B., Brad M. Barber, and Terrance Odean, 2019, The Math Gender Gap and Women's Career Outcomes, working paper.

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, Renée B., Roman Kräussl, Marco Navone, and Patrick Verwijmeren, 2021, Gendered prices, *The Review of Financial Studies* 34, 3789-3839.

Adhikari, Binay K., Anup Agrawal, and James Malm, 2019, Do women managers keep firms out of trouble? Evidence from corporate litigation and policies, *Journal of Accounting and Economics* 67, 202-225.

Aggarwal, Rajesh, and Nicole M. Boyson, 2016, The performance of female hedge fund managers, *Review of Financial Economics* 29, 23-36.

Alesina, Alberto F., Francesca Lotti, and Paolo E. Mistrulli, 2013, Do women pay more for credit? Evidence from Italy, *Journal of the European Economic Association* 11, 45-66.

Ambrose, Brent W., James N. Conklin, and Luis A. Lopez, 2021, Does borrower and broker race affect the cost of mortgage credit? *The Review of Financial Studies* 34, 790-826.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin, 1996, Identification of causal effects using instrumental variables, *Journal of the American Statistical Association* 91, 444-455.

Anwar, Shamena, and Hanming Fang, 2006, An alternative test of racial prejudice in motor vehicle searches: Theory and evidence, *American Economic Review* 96, 127-151.

Arnold, David, Will Dobbie, and Crystal S. Yang, 2018, Racial bias in bail decisions, *The Quarterly Journal of Economics* 133, 1885-1932.

Arrow, Kenneth, 1973, The Theory of Discrimination. Ed. Orley Ashenfelter and Albert Rees *Discrimination in Labor Markets* (Princeton University Press).

Ayres, Ian, 2002, Outcome tests of racial disparities in police practices, *Justice Research and Policy* 4, 131-142.

Barber, Brad M., Wei Jiang, Adair Morse, Manju Puri, Heather Tookes, and Ingrid Werner, 2021, What explains differences in finance research productivity during the pandemic? *The Journal of Finance* 76, 1655-1697.

Barber, Brad M., and Terrance Odean, 2001, Boys will be boys: Gender, overconfidence, and common stock investment, *The Quarterly Journal of Economics* 116, 261-292.

Barber, Brad M., Anna Scherbina, and Bernd Schlusche, 2017, Performance isn't everything: personal characteristics and career outcomes of mutual fund managers, working paper.

Becker, Gary S., 1957, *The Economics of Discrimination* (University of Chicago press).

---. 1993, Nobel lecture: The economic way of looking at behavior, *Journal of Political Economy* 101, 385-409.

Benson, Alan, Danielle Li, and Kelly Shue, 2019, Promotions and the Peter principle, *The Quarterly Journal of Economics* 134, 2085-2134.

---. 2021, "Potential" and the gender promotion gap, working paper.

Bertrand, Marianne, and Esther Duflo, 2017, Field experiments on discrimination, *Handbook of Economic Field Experiments* 1, 309-393.

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, 228-55.

Blau, Francine D., and Lawrence M. Kahn, 2017, The gender wage gap: Extent, trends, and explanations, *Journal of Economic Literature* 55, 789-865.

Bohren, J. A., Kareem Haggag, Alex Imas, and Devin G. Pope, 2019, Inaccurate statistical discrimination: An identification problem, working paper.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, 2016, Stereotypes, *The Quarterly Journal of Economics* 131, 1753-1794.

---. 2019, Beliefs about gender, *American Economic Review* 109, 739-773.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2018, Fintech, regulatory arbitrage, and the rise of shadow banks, *Journal of Financial Economics* 130, 453-483.

Chan, Eric W., 2018, Promotion, relative performance information, and the Peter principle, *The Accounting Review* 93, 83-103.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff, 2014a, Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates, *American Economic Review* 104, 2593-2632.

---. 2014b, Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood, *American Economic Review* 104, 2633-79.

Coffman, Katherine B., Christine L. Exley, and Muriel Niederle, 2020, The role of beliefs in driving gender discrimination, working paper.

Cullen, Zoë, and Ricardo Perez-Truglia, 2018, How much does your boss make? the effects of salary comparisons, NBER working paper.

Dobbie, Will, Andres Liberman, Daniel Paravisini, and Vikram Pathania, 2021, Measuring Bias in Consumer Lending, *Review of Economic Studies*, forthcoming.

Egan, Mark L., Gregor Matvos, and Amit Seru, 2020, When Harry fired Sally: The double standard in punishing misconduct, *Journal of Political Economy*, forthcoming.

Faccio, Mara, Maria-Teresa Marchica, and Roberto Mura, 2016, CEO gender, corporate risk-taking, and the efficiency of capital allocation, *Journal of Corporate Finance* 39, 193-209.

Fang, Lily H., and Sterling Huang, 2017, Gender and connections among Wall Street analysts, *The Review of Financial Studies* 30, 3305-3335.

Farrell, Kathleen A., and Philip L. Hersch, 2005, Additions to corporate boards: The effect of gender, *Journal of Corporate Finance* 11, 85-106.

Glover, Dylan, Amanda Pallais, and William Pariente, 2017, Discrimination as a self-fulfilling prophecy: Evidence from French grocery stores, *The Quarterly Journal of Economics* 132, 1219-1260.

Goldin, Claudia, and Cecilia Rouse, 2000, Orchestrating impartiality: The impact of" blind" auditions on female musicians, *American Economic Review* 90, 715-741.

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

Hjort, Jonas, 2014, Ethnic divisions and production in firms, *The Quarterly Journal of Economics* 129, 1899-1946.

Hospido, Laura, Luc Laeven, and Ana Lamo, 2019, The gender promotion gap: evidence from central banking, *The Review of Economics and Statistics* , 1-45.

Huang, Jiekun, and Darren J. Kisgen, 2013, Gender and corporate finance: Are male executives overconfident relative to female executives? *Journal of Financial Economics* 108, 822-839.

Huber, Kilian, Volker Lindenthal, and Fabian Waldinger, 2021, Discrimination, managers, and firm performance: Evidence from "aryanizations" in nazi germany, *Journal of Political Economy* 129, 2455–2503.

Kumar, Alok, 2010, Self-selection and the forecasting abilities of female equity analysts, *Journal of Accounting Research* 48, 393-435.

Niederle, Muriel, and Lise Vesterlund, 2007, Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics* 122, 1067-1101.

Niessen-Ruenzi, Alexandra, and Stefan Ruenzi, 2019, Sex matters: Gender bias in the mutual fund industry, *Management Science* 65, 3001-3025.

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

Post, Corinne, and Kris Byron, 2015, Women on boards and firm financial performance: A meta-analysis, *Academy of Management Journal* 58, 1546-1571.

Sherman, Mila G., and Heather Tookes, 2021, Female representation in the academic finance profession, *Journal of Finance*, forthcoming.

Staiger, Douglas O., and James H. Stock, 1997, Instrumental variables regression with weak instruments, *Econometrica* 65, 557-586.

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

Zillow, 2021, Zillow Transaction and Assessment Dataset (ZTRAX).

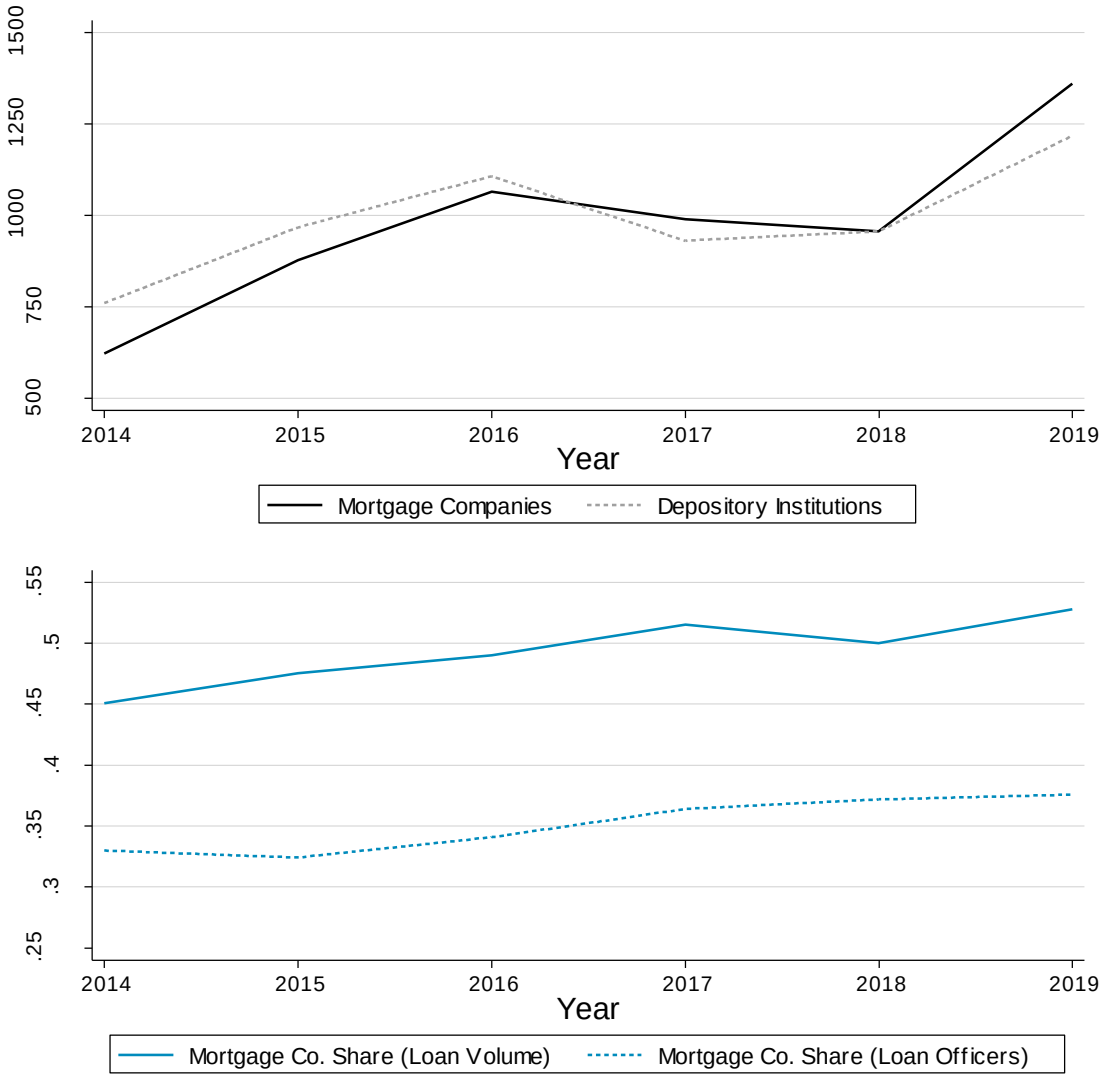

Figure 1: The importance of mortgage companies in the U.S. mortgage market
This figure shows the importance of mortgage companies in the U.S. mortgage market during our sample period (2014 to 2019). The top plot presents the total mortgage origination volume across all loan types by mortgage companies (black line) and depository institutions (dashed line). The bottom plot presents mortgage companies' share of the market based on loan volume (blue line) and number of loan officers (dashed line).

Table 1: Descriptive statistics

This table presents descriptive statistics for our sample. Panel A reports female representation among loan officers and managers. Columns 1 shows the total number of loan officers in our sample each year, Column 2 shows the percentage with gender identified, and Column 3 shows the percentage of female officers. Columns 4-6 present the same statistics for managers in our sample. Panel B reports summary statistics from our loan officer-year panel dataset for the variables used in the analysis. Columns 1-6 present the sample size (N), mean, standard deviation (SD), 25th percentile (P25), 50th percentile (P50), and 75th percentile (P75), respectively. All the variables are defined in Appendix A.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Female representation						
	Total number of loan officers	% with gender identified	% Female	Total number of managers	% with gender identified	% Female
2014	20,634	99.32%	34.40%	4,369	99.89%	26.31%
2015	25,995	99.40%	34.17%	5,185	99.90%	26.37%
2016	29,833	99.48%	34.04%	5,814	99.90%	27.03%
2017	32,587	99.45%	34.44%	6,374	99.91%	27.50%
2018	32,615	99.45%	35.60%	6,590	99.89%	28.35%
2019	33,669	99.25%	36.38%	6,868	99.93%	28.65%

Panel B: Summary statistics						
	N	Mean	SD	P25	P50	P75
Promoted	174,272	0.0115	0.1066	0	0	0
Number of loans	174,272	36.5262	47.9252	6	18	47
Log(number of loans)	174,272	2.8910	1.2801	1.9459	2.9444	3.8712
Loan amount (million $)	174,272	8.5121	11.4459	1.2291	4.1125	10.8284
Log(loan amount)	174,272	15.0676	1.4974	14.0218	15.2296	16.1977
Foreclosure ratio	174,272	0.0079	0.0293	0	0	0
Foreclosure amount	174,272	0.0009	0.0053	0	0	0
Female	174,272	0.3491	0.4767	0	0	1
White	174,272	0.7422	0.4374	0	1	1
Tenure	174,272	2.3245	1.8195	1	2	3
Top LO at branch	174,272	0.2010	0.4007	0	0	0

Table 2: The gender gap in promotions

This table reports OLS regressions testing whether loan officers' gender affects the likelihood of being promoted to manager. The unit of observation is at the loan officer-year level. The dependent variable is an indicator for the loan officer being promoted in the next year, and the key independent variable is the *Female* indicator. All variables are defined in Appendix A. The standard errors are clustered at the loan officer level, and statistical significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	(1)	(2)	(3)
Dependent variable:		Promoted	
Female	-0.0014**	-0.0010*	-0.0017***
	(0.0006)	(0.0006)	(0.0006)
Tenure		-0.0007***	0.0002
		(0.0002)	(0.0002)
Top LO at branch		0.0035***	-0.0036***
		(0.0008)	(0.0009)
Log(number of loans)		0.0040***	0.0044***
		(0.0003)	(0.0003)
White		0.0022***	0.0014**
		(0.0006)	(0.0006)
Firm x year FE	Y	Y	Y
Branch FE	N	N	Y
Observations	174,015	174,015	171,182
R-squared	0.0283	0.0299	0.1422

Table 3: Gender bias in promotions

This table presents IV regressions testing for differences in managerial effectiveness between marginally promoted male and female managers. The unit of observation is at the loan officer-year level. The dependent variables in Columns 1 and 2 and Columns 3 and 4 are the managerial effect on subordinates' number of loans, and on total loan amount, respectively. We extract these managerial effects from the manager fixed effects estimation for each manager by regressing each subordinate's loan volume on loan officer fixed effects, manager fixed effects, firm-year fixed effects, and a control for loan officer tenure (see equation (2)). The key independent variable, *Promoted*, is an indicator that equals one if the loan officer is promoted in the next year and zero otherwise. We instrument for *Promoted* with a leave-out measure of the average promotion rate of loan officers at the same firm in the same year, excluding the focal loan officer and other officers working at the same branch. The sample is split on gender in order to identify differences between the marginally promoted male manager (Columns 1 and 3) and the marginally promoted female manager (Columns 2 and 4). The final row reports p-values from a Wald test of equality at the margin, i.e., between the *Promoted* coefficients in the male/female tests. The controls include *Tenure*, *Top LO at branch*, *Log(number of loans)*, and *White*. All variables are defined in Appendix A. The standard errors are clustered at the loan officer level, and statistical significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	(1)	(2)	(3)	(4)
Dependent variable:	Managerial effect on number of loans		Managerial effect on loan amount	
Subsample:	Male	Female	Male	Female
Promoted	-0.0308**	0.0202**	-0.0400***	0.0206*
	(0.0126)	(0.0092)	(0.0153)	(0.0113)
Controls	Y	Y	Y	Y
Observations	113,425	60,847	113,425	60,847
First stage F-stat	95.7811	45.5624	95.7811	45.5624
P-value	0.0011***		0.0014***	

Table 4: In-group tests—the previous manager's gender and promotion equity

This table presents IV regressions testing for differences in managerial effectiveness between marginally promoted male and female managers, split based on whether the previous manager (a likely participant in promotion decisions) was male versus female. The unit of observation is at the loan officer-year level. The dependent variables in Panels A and B are the managerial effect on subordinates' number of loans, and on total loan amount, respectively. We extract these managerial effects from the manager fixed effects estimation for each manager by regressing each subordinate's loan volume on loan officer fixed effects, manager fixed effects, firm-year fixed effects, and a control for loan officer tenure (see equation (2)). The key independent variable, *Promoted*, is an indicator that equals one if the loan officer is promoted in the next year and zero otherwise. We instrument for *Promoted* with a leave-out measure of the average promotion rate of loan officers at the same firm in the same year, excluding the focal loan officer and other officers working at the same branch. The sample is split first based on whether the previous manager was male (Columns 1 and 2) or female (Columns 3 and 4), and then based on loan officers' own gender to test for gender differences at the margin. The final row reports p-values from a Wald test of equality at the margin, i.e., between the *Promoted* coefficients in the male/female loan officer tests. The controls include *Tenure*, *Top LO at branch*, *Log(number of loans)*, and *White*. All variables are defined in Appendix A. The standard errors are clustered at the loan officer level, and statistical significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	(1)	(2)	(3)	(4)
Panel A: Managerial effect on number of loans				
Subsample:	Male previous manager		Female previous manager	
	Male	Female	Male	Female
Promoted	-0.0372**	0.0004	-0.0096	0.0515**
	(0.0159)	(0.0004)	(0.0098)	(0.0248)
Controls	Y	Y	Y	Y
Observations	95,725	45,240	17,630	15,536
First stage F-stat	99.7674	23.6490	8.7005	22.2344
P-value	0.0184**		0.0218**	
Panel B: Managerial effect on loan amount				
Subsample:	Male previous manager		Female previous manager	
	Male	Female	Male	Female
Promoted	-0.0505***	-0.0003	-0.0050	0.0566*
	(0.0195)	(0.0003)	(0.0051)	(0.0309)
Controls	Y	Y	Y	Y
Observations	95,725	45,240	17,630	15,536
First stage F-stat	99.7674	23.6490	8.7005	22.2344
P-value	0.0101**		0.0488**	

Table 5: Gender bias and the Peter Principle

This table presents IV regressions testing for differences in managerial effectiveness between marginally promoted managers with low versus high pre-promotion sales volume, split based on whether the person is male versus female. The unit of observation is at the loan officer-year level. The dependent variables in Panels A and B are the managerial effect on subordinates' number of loans, and on total loan amount, respectively. We extract these managerial effects from the manager fixed effects estimation for each manager by regressing each subordinate's loan volume on loan officer fixed effects, manager fixed effects, firm-year fixed effects, and a control for loan officer tenure (see equation (2)). The key independent variable, *Promoted*, is an indicator that equals one if the loan officer is promoted in the next year and zero otherwise. We instrument for *Promoted* with a leave-out measure of the average promotion rate of loan officers at the same firm in the same year, excluding the focal loan officer and other officers working at the same branch. The sample is split first based on whether the loan officer is male (Columns 1 and 2) or female (Columns 3 and 4), and then based on whether they have low versus high prepromotion sales volume. The final row reports p-values from a Wald test of equality at the margin, i.e., between the *Promoted* coefficients in the low versus high sales tests. The controls include *Tenure*, *Top LO at branch*, *Log(number of loans)*, and *White*. All variables are defined in Appendix A. The standard errors are clustered at the loan officer level, and statistical significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	(1)	(2)	(3)	(4)
Panel A: Managerial effect on number of loans				
Subsample:	Male		Female	
	Low sales	High sales	Low sales	High sales
Promoted	-0.0048	-0.0437**	0.0068	0.0274**
	(0.0068)	(0.0188)	(0.0098)	(0.0134)
Controls	Y	Y	Y	Y
Observations	56,952	56,473	32,704	28,143
First stage F-stat	34.7210	73.1238	19.0632	28.8031
P-value	0.0517*		0.2148	
Panel B: Managerial effect on loan amount				
Subsample:	Male		Female	
	Low sales	High sales	Low sales	High sales
Promoted	-0.0075	-0.0558**	0.0017	0.0310**
	(0.0073)	(0.0227)	(0.0147)	(0.0156)
Controls	Y	Y	Y	Y
Observations	56,952	56,473	32,704	28,143
First stage F-stat	34.7210	73.1238	19.0632	28.8031
P-value	0.0430**		0.1723	

Table 6: Robustness test ruling out differences in loan quality under female managers

This table reports robustness tests based on loan quality measures. We use IV regressions to test for differences between the marginally promoted male versus female manager's impact on the loan quality of their subordinates. The unit of observation is at the loan officer-year level. The dependent variables in Columns 1 and 2 and Columns 3 and 4 are the managerial effect on subordinates' foreclosure ratio, and on their foreclosure amount, respectively. We extract these managerial effects from the manager fixed effects estimation for each manager by regressing each subordinate's foreclosure statistics on loan officer fixed effects, manager fixed effects, firm-year fixed effects, and a control for loan officer tenure (see equation (2)). The key independent variable, *Promoted*, is an indicator that equals one if the loan officer is promoted in the next year and zero otherwise. We instrument for *Promoted* with a leave-out measure of the average promotion rate of loan officers at the same firm in the same year, excluding the focal loan officer and other officers working at the same branch. The sample is split on gender in order to identify differences between the marginally promoted male manager (Columns 1 and 3) and the marginally promoted female manager (Columns 2 and 4). The final row reports p-values from a Wald test of equality at the margin, i.e., between the *Promoted* coefficients in the male/female tests. The controls include *Tenure, Top LO at branch, Log(number of loans)*, and *White*. All variables are defined in Appendix A. The standard errors are clustered at the loan officer level, and statistical significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	(1)	(2)	(3)	(4)
Dependent variable:	Managerial effect on foreclosure ratio		Managerial effect on foreclosure amount	
Subsample:	Male	Female	Male	Female
Promoted	-0.0090	0.0014	-0.0282	0.0025
	(0.0063)	(0.0026)	(0.0265)	(0.0377)
Controls	Y	Y	Y	Y
Observations	113,425	60,847	113,425	60,847
First stage F-stat	95.7811	45.5624	95.7811	45.5624
P-value	0.1268		0.5055	

Table 7: Manager gender and loan officers' performance

This table reports OLS differences-in-differences regressions examining the average effect of having a female manager on subordinate loan officers' performance. The unit of observation is at the loan officer-year level. Panel A focuses on loan volume measures. The dependent variables in Columns 1 and 2 and Columns 3 and 4 are *Log(number of loans)* and *Log(loan amount)*, respectively. Panel B focuses on loan quality measures. The dependent variables in Columns 1 and 2 and Columns 3 and 4 are *Foreclosure ratio* and *Foreclosure amount*, respectively. The key independent variables are *Female manager*, which indicates the loan officer's current manager is female, and its interaction with *Female LO*, which indicates the loan officer herself is female. All variables are defined in Appendix A. The standard errors are clustered at the loan officer level, and statistical significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	(1)	(2)	(3)	(4)
Panel A: Loan volume				
Dependent variable:	Log(number of loans)		Log(loan amount)	
Female manager	0.0039	-0.0226	0.0144	-0.0190
	(0.0198)	(0.0233)	(0.0232)	(0.0274)
Female manager x female LO		0.0630**		0.0794**
		(0.0274)		(0.0325)
LO tenure control	Y	Y	Y	Y
LO FE	Y	Y	Y	Y
Branch FE	Y	Y	Y	Y
Firm x year FE	Y	Y	Y	Y
Observations	144,254	144,254	144,254	144,254
R-squared	0.6513	0.6513	0.6355	0.6356
Panel B: Loan quality				
Dependent variable:	Foreclosure ratio		Foreclosure amount	
Female manager	-0.0006	0.0002	0.0001	0.0002
	(0.0008)	(0.0009)	(0.0001)	(0.0001)
Female manager x female LO		-0.0018		-0.0002
		(0.0011)		(0.0002)
LO tenure control	Y	Y	Y	Y
LO FE	Y	Y	Y	Y
Branch FE	Y	Y	Y	Y
Firm x year FE	Y	Y	Y	Y
Observations	144,254	144,254	144,254	144,254
R-squared	0.4825	0.4825	0.4244	0.4244

Table 8: Firms' gender promotion practices and firm-level outcomes

This table reports OLS regressions of firm outcomes on a firm-level measure of gender equity in promotions. The unit of observation is at the firm-year level. Panel A focuses on loan volume measures. The dependent variables in Columns 1 and 2 and Columns 3 and 4 are the firm's *Log(number of loans)* and *Log(loan amount),* respectively. Panel B focuses on loan quality measures. The dependent variables in Columns 1 and 2 and Columns 3 and 4 are the firm's *Foreclosure ratio* and *Foreclosure amount,* respectively. Panel C focuses on firm employment and exits. The dependent variables in Columns 1 and 2 and Columns 3 and 4 are the firm's *Log(number of employees)* and *Firm exit next year*, respectively. The key independent variable, *% female manager/% female LO*, is the ratio of the percentage of managers who are female over the percentage of loan officers who are female at the firm. All variables are defined in Appendix A. The standard errors are clustered at the firm level, and statistical significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	(1)	(2)	(3)	(4)
Panel A: Loan volume				
Dependent variable:	Log(number of loans)		Log (loan amount)	
% female manager/% female LO	0.7231***	0.2198***	0.7733***	0.2410***
	(0.0877)	(0.0405)	(0.0890)	(0.0434)
Firm FE	N	Y	N	Y
Year FE	Y	Y	Y	Y
Observations	2,807	2,675	2,807	2,675
R-squared	0.0656	0.8716	0.0714	0.8648
Panel B: Loan quality				
Dependent variable:	Foreclosure ratio		Foreclosure amount	
% female manager /% female LO	-0.0001	0.0004	-0.0000	0.0001
	(0.0004)	(0.0006)	(0.0001)	(0.0002)
Firm FE	N	Y	N	Y
Year FE	Y	Y	Y	Y
Observations	2,807	2,675	2,807	2,675
R-squared	0.0946	0.3108	0.0513	0.2611
Panel C: Firm employment and exits				
Dependent variable:	Log(number of employees)		Firm exit next year	
% female manager /% female LO	0.5247***	0.1554***	-0.0137*	0.0010
	(0.0692)	(0.0283)	(0.0081)	(0.0102)
Firm FE	N	Y	N	Y
Year FE	Y	Y	Y	Y
Observations	2,807	2,675	2,300	2,195
R-squared	0.0457	0.8982	0.0084	0.1462

Appendix A—Variable Definitions

Variables	Definition
Promoted	Equals one if the loan officer is promoted to manager and zero otherwise
Log(number of loans)	Natural logarithm of the number of loans originated
Log(loan amount)	Natural logarithm of the total dollar amount of loans originated
Foreclosure ratio	The ratio of number of foreclosures over number of loans originated
Foreclosure amount	The ratio of total dollar amount of foreclosures over total dollar amount of loans originated
Female	Equals one if the loan officer is female and zero otherwise
White	Equals one if the loan officer is white, and zero otherwise
Tenure	Number of years working at the firm
Top LO at branch	Equals one if the loan officer is ranked number one in number of loans originated at the branch
Log(number of employees)	Natural logarithm of the total number of employees
Firm exit next year	Equals one if the firm exits the data in the next year and zero otherwise

Internet Appendix

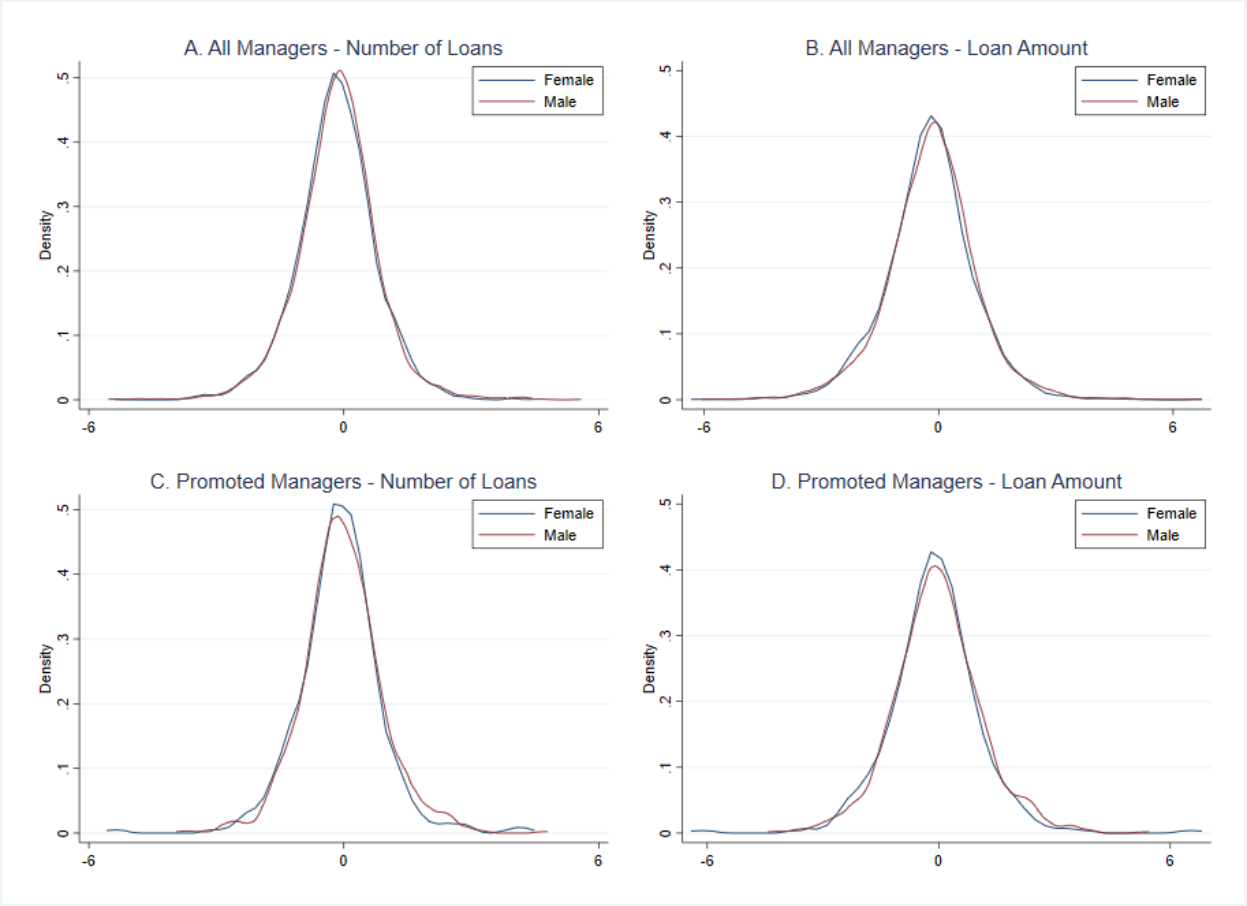

Figure IA1: Distribution of managerial effects by gender
These figures show the distribution of managerial effects by gender using both the effect on subordinates' number of loans and dollar volume. Plots A and B include all managers. Plots C and D include only managers who were promoted during our 2014-2019 sample period.

Table IA1: First-stage regressions using the leave-out measure as an instrument

This table reports the first-stage regressions for our IV analysis. The unit of observation is at the loan officer-year level. The dependent variable, *Promoted*, is an indicator for the loan officer being promoted in the next year. The key independent variable (the instrument) is the *Leave-out measure*, which is the fraction of loan officers promoted within the same firm in the same year, excluding the focal loan officer and other loan officers working at the same branch. Columns 1-2, 3-4, and 5-6 present the results for the full sample, the male loan officer sample, and the female loan officer sample, respectively. The controls include *Tenure*, *Top LO at branch*, *Log(number of loans)*, and *White*. All variables are defined in Appendix A. The standard errors are clustered at the loan officer level, and statistical significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Promoted					
Subsample:	All loan officers		Male loan officers		Female loan officers	
Leave-out measure	0.3417***	0.3322***	0.3131***	0.3035***	0.3924***	0.3790***
	(0.0303)	(0.0302)	(0.0329)	(0.0327)	(0.0589)	(0.0587)
Controls	N	Y	N	Y	N	Y
Observations	174,272	174,272	113,425	113,425	60,847	60,847
R-squared	0.0019	0.0019	0.0016	0.0033	0.0027	0.0051

Table IA2: Managerial effects and the leave-out measure

This table reports OLS regressions examining how firm characteristics affect manager's managerial effect at promotion. The sample is based on all newly promoted managers during our sample period (2014 to 2019). The dependent variables in columns 1 and 3 and columns 4 and 6 are the managerial effect on the number of loans, and on the total loan amount, respectively. These managerial effects are extracted from the manager fixed effects estimation for each manager by regressing each subordinate's loan volume on loan officer fixed effects, manager fixed effects, firm-year fixed effects, and a control for loan officer tenure (see equation (2)). All variables are defined in Appendix A. The standard errors are clustered at the manager level, and statistical significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Managerial effect on number of loans			Managerial effect on loan amount		
Leave-out measure	0.0115			0.0187		
	(0.1807)			(0.2132)		
Log(number of loans)		0.0013			0.0015	
		(0.0022)			(0.0026)	
Log(number of employees)			0.0018			0.0018
			(0.0026)			(0.0030)
Controls	Y	Y	Y	Y	Y	Y
Observations	2,005	2,005	2,005	2,005	2,005	2,005
R-squared	0.0015	0.0017	0.0017	0.0025	0.0026	0.0026

Table IA3: The relationship between gender and average managerial effect

This table reports OLS regressions testing for differences in managerial effectiveness between the average promoted male and female managers. The unit of observation is at the loan officer-year level. The dependent variables in Columns 1 and 2 and Columns 3 and 4 are the managerial effect on subordinates' number of loans, and on total loan amount, respectively. We extract these managerial effects from the manager fixed effects estimation for each manager by regressing each subordinate's loan volume on loan officer fixed effects, manager fixed effects, firm-year fixed effects, and a control for loan officer tenure (see equation (2)). The key independent variable, *Promoted*, is an indicator that equals one if the loan officer is promoted in the next year and zero otherwise. The sample is split on gender in order to identify differences between the marginally promoted male manager (Columns 1 and 3) and the marginally promoted female manager (Columns 2 and 4). The final row reports p-values from a Wald test of equality, i.e., between the *Promoted* coefficients in the male/female tests. The controls include *Tenure*, *Top LO at branch*, *Log(number of loans)*, and *White*. All variables are defined in Appendix A. The standard errors are clustered at the loan officer level, and statistical significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	(1)	(2)	(3)	(4)
Dependent variable:	Managerial effect on number of loans		Managerial effect on loan amount	
Subsample:	Male	Female	Male	Female
Promoted	-0.0115***	-0.0041	-0.0148***	-0.0089
	(0.0044)	(0.0060)	(0.0054)	(0.0066)
Controls	Y	Y	Y	Y
Observations	113,425	60,847	113,425	60,847
R squared	0.0051	0.0007	0.0057	0.0028
P-value	0.3157		0.4905	

Table IA4: Manager gender and subordinate loan officers' lending

This table reports OLS differences-in-differences regressions examining the average effect of having a female manager on subordinate loan officers' lending. The unit of observation is at the loan officer-year level. Panel A focuses on the percentage of loans made to female borrowers (defined as having a woman as the first-signer). The dependent variable in Columns 1 and 2 is the *Fraction of loans* made to women (we focus on only home purchase mortgages by non-fintech lenders where loan officers may have discretion). The dependent variable in Columns 3 and 4 is the *Fraction of loan amount*. Panel B focuses on the percentage of loans made to low-income borrowers (people living in ZIP codes in the bottom quartile of personal income per capita). The key independent variables are *Female manager*, which indicates the loan officer's current manager is female, and its interaction with *Female LO*, which indicates the loan officer herself is female. All variables are defined in Appendix A. The standard errors are clustered at the loan officer level, and statistical significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	(1)	(2)	(3)	(4)
Panel A: Loans to female borrowers				
Dependent variable:	Fraction of loans		Fraction of loan amount	
Female manager	0.0018	0.0061	0.0025	0.0065
	(0.0050)	(0.0061)	(0.0051)	(0.0061)
Female manager x female LO		-0.0100		-0.0095
		(0.0074)		(0.0076)
LO tenure control	Y	Y	Y	Y
LO FE	Y	Y	Y	Y
Branch FE	Y	Y	Y	Y
Firm x year FE	Y	Y	Y	Y
Observations	144,254	144,254	144,254	144,254
R-squared	0.3693	0.3693	0.3232	0.3232
Panel B: Loans to low-income borrowers				
Dependent variable:	Fraction of loans		Fraction of loan amount	
Female manager	0.0054	-0.0040	0.0047	-0.0050
	(0.0080)	(0.0090)	(0.0088)	(0.0099)
Female manager x female LO		0.0224*		0.0228*
		(0.0119)		(0.0131)
LO tenure control	Y	Y	Y	Y
LO FE	Y	Y	Y	Y
Branch FE	Y	Y	Y	Y
Firm x year FE	Y	Y	Y	Y
Observations	144,254	144,254	144,254	144,254
R-squared	0.5773	0.5773	0.4945	0.4945