CESifo WORKING PAPERS

6915
2018

February 2018

How Do Gender Quotas Affect Hierarchical Relationships? Complementary Evidence from a Representative Survey and Labor Market Experiments

Edwin Ip, Andreas Leibbrandt, Joseph Vecci

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6915*
Category 13: Behavioural Economics

How Do Gender Quotas Affect Hierarchical Relationships? Complementary Evidence from a Representative Survey and Labor Market Experiments

Abstract

Gender quotas are frequently proposed to address persistent gender imbalances in managerial roles. However, it is unclear how quotas for female managers affect organizations and whether quotas improve or damage relationships between managers and their subordinates. We conduct a representative survey to study opinions on quotas for female managers and based upon design a novel set of experiments to investigate how quotas influence wage setting and effort provision. Our findings reveal that both opinions about gender quotas and workplace behavior crucially depend on the workplace environment. In our survey, we observe that approval for gender quotas is low if women are not disadvantaged in the manager selection process, regardless of whether there are gender differences in performance. Complementing this evidence, we observe in our experiments that quotas lead to lower effort levels and lower wages in such environments. By contrast, in environments in which women are disadvantaged in the selection process, we observe a higher approval of quotas as well as higher effort levels and higher wages. These findings are consistent with the concept of meritocracy and suggest that it is important to evaluate the existence of gender disadvantages in the workplace environment before implementing quotas.

JEL-Codes: C920, J710, J300.

Keywords: gender quota, hierarchical relationships, fairness, meritocracy.

Edwin Ip
Monash University
Australia – Clayton VIC 3800
edwin.ip@monash.edu

Andreas Leibbrandt
Monash University
Australia – Clayton VIC 3800
andreas.leibbrandt@monash.edu

Joseph Vecci
Gothenburg University / Sweden
joseph.vecci@gu.se

February 13, 2018

1. Introduction

Proponents of gender quotas regard them as a last resort and a necessary evil to address gender gaps that have not closed organically.[1] However, during recent decades, resistance to gender quotas has weakened considerably and they are now implemented in many different environments. The number of countries to adopt gender quotas has risen from 5 to more than 120 in less than 50 years (Quota Project, 2016). In the private sector, gender quotas at the upper echelon have affected thousands of organizations and are likely to be implemented even more widely.[2] Such quotas not only affect the gender composition of superiors, but also could affect the relationship between superiors and their subordinates, and ultimately, organizational performance. For example, if subordinates believe that their female superior is in a position because of a quota rather than merit, they might be less motivated and less cooperative. By contrast, if subordinates believe that a quota helps to overcome gender disadvantages, then they might be more motivated and cooperative compared to when there is no quota.

Our study constitutes a first attempt to experimentally investigate the impact of gender quotas on the functioning of hierarchical relationships. A key consideration in our design is the link between gender quotas and meritocracy. Our main hypothesis is that the impact of gender quotas on the functioning of hierarchical relationships depends on how they affect the expected suitability of the selected superior: quotas improve their functioning if they increase suitability but harm their functioning if they decrease suitability. To test this hypothesis, and rule out alternative explanations, we design and conduct a representative survey as well as a set of laboratory experiments. Both survey and experiments capture reactions to quotas for female

[1] For example, Australian Institute of Company Directors chair Elizabeth Proust concedes that quotas might be the last resort in the battle to get even with the "blokes' club" in corporate Australia (http://www.abc.net.au/news/2017-09-26/companies-warned-they-could-face-gender-quotas/8987862).
Furthermore, according to a panel of female technology and business leaders, "Female quotas are a necessary evil to guarantee enough women rise to the top echelon of business and public life in Australia" (http://www.smh.com.au/it-pro/expertise/quotas-still-needed-to-guarantee-women-rise-in-business-leaders-20140313-hvidi.html).

[2] Norway became the first country to introduce a quota for female directors of listed companies in 2008. Since then, both mandatory and voluntary gender quotas for corporate boards have been imposed in Austria, Belgium, Brazil, France, Germany, Greece, Iceland, Italy, Malaysia, the Netherlands, and Spain, while the European Commission, Australia, Britain, and Sweden are considering quotas for female directors (The Economist, 2014; European Commission, 2016).

managers in environments in which gender differences in performance and treatment vary (i.e., there is disadvantage in the selection procedure).

The survey provides first suggestive evidence in favor of our main hypothesis and emphasizes the importance of the workplace environment. We observe that *general* opinions toward gender quotas for female managerial positions are divided and depend on the respondents' own workplace environment. Importantly, as hypothesized, we find that opinions shift when we present different workplace environments under which quotas are implemented. Quotas clearly lack majority support in environments in which the quotas do not increase the suitability of female managers, because women are not disadvantaged in the manager selection process. Alternatively, if the quotas increase this suitability because women are disadvantaged, female quotas are supported by a large majority.

Next, we test whether these opinions translate into workplace behavior. We use a set of economic experiments to study the causal impact of gender quotas on hierarchical workplace behavior in systematically different workplace environments. Taking this topic to the laboratory has four key advantages: it allows us to design experiments in which managerial selection can be exogenously manipulated such that gender quotas can be randomly implemented, the workplace environment can be exogenously manipulated, the emergence of hierarchical relations is completely transparent, and we can provide monetary incentives to counter potential social desirability bias. Taken together, our survey and experiment provide a wider glimpse into the relationship between gender quotas and organizational performance than most other studies can.

Our experiments consist of two connected parts. First, subjects participate in a performance task. Second, subjects are assigned to the role of manager or worker in a gift-exchange game, depending on their earlier task performance and treatment. We follow a 3-by-2 design. In the *no-quota* treatments, the best performers are promoted to the role of manager. By contrast, in the *quota* treatments, the procedure guarantees a proportion of managerial positions to the best-performing women. We test the impact of *quota* relative to *no quota* on behavior between managers and workers in three different stylized workplace environments akin to our survey, which are characterized by (*i*) the absence of gender differences in performance, (*ii*) the presence of gender differences in performance, or (*iii*) the presence of disadvantage against women in the selection procedure.

Our findings reveal workplace reactions to gender quotas. The experimental findings are consistent with the survey evidence and corroborate our main hypothesis. We find that gender quotas affect hierarchical relationships and that the impact crucially depends on the environment in which the quota is implemented. If the gender quota is implemented in an environment in which women's performance is believed to be on average lower than men's performance, it causes a decline in gift exchange, with lower wages and effort levels than when there is no quota. If the gender quota is implemented in the absence of any gender performance differences, it causes lower wages and effort levels than when there is no quota. However, if the quota is implemented in an environment in which there is discrimination against women, it leads to higher wage and effort levels than when there is no quota.

This study bridges and extends the scope of two growing research fields. It connects research on gender quotas (e.g., Chattopadhyay and Duflo, 2004; Adams and Ferreira ., 2009; Ahern and Dittmar, 2012; Balafoutas and Sutter, 2012; Matsa and Miller, 2013; Niederle et al., 2013, Balafoutas et al., 2016) to research on fairness in hierarchical relationships (e.g., Fehr et al., 1993; Charness and Haruvy,., 2002; Charness and Kuhn,., 2007; Charness and Kuhn, 2011; Brown et al., 2012; Herz et al., 2018).

The research on gender quotas is largely restricted to experimental studies investigating whether quotas assist women to compete in workplace settings (Schotter and Weigelt, 1992; Balafoutas and Sutter, 2012; Niederle et al., 2013; Leibbrandt et al., 2017) and field studies investigating mandated quotas for corporate boards in Norway (Ahern and Dittmar, 2013; Matsa and Miller, 2013; Bertrand et al., 2014) and village councils in India (for example, Chattopadhyay and Duflo,, 2004; Beaman et al., 2009).[3] By contrast, our study provides first insights on how quotas are perceived in different workplace environments, how quotas impact workplace relationships, and consequently, how they affect organizational performance. In

[3] The findings in these studies suggest that gender quotas are complex and their impact crucially depends on context and measurement. For example, while most experimental studies find that gender quotas encourage females to compete in tournaments, Leibbrandt et al. (2017) suggest that this encouragement effect is absent in an environment in which sabotage is possible. Furthermore, while Ahern and Dittmar (2013) find that the Norwegian quota led to less experienced boards and a decline in operating performance, Matsa and Miller (2013) find that the quota did not affect corporate decisions, except for employment policies. Firms affected by the female board quota undertook fewer workforce reductions, leading to lower short-term profits. Finally, while Beaman et al. (2009) find that quotas in India weaken gender stereotypes, the authors also find an increase in distaste for female leaders.

addition, our study advances the literature by investigating environments that capture key workplace features, allowing for a rigorous test of different behavioral models, and corroborates the importance of meritocracy in the context of gender quotas.

There is manifold evidence for the key role of fairness in hierarchical relationships (Kahneman et al., 1986; Konow, 1996, 2000; Charness and Haruvy, 2002; Fehr et al., 2008; Kube et al., 2012). In particular, there is experimental evidence that fairness is a crucial ingredient for efficiency wages and firm profit (Akerlof, 1982; Akerlof and Yellen, 1990; Fehr et al., 1993; Charness and Kuhn, 2007; Charness et al., 2012; Gächter et al., 2013; Englmaier and Strasser, 2014; Gilchrist et al., 2016; see also Fehr et al., 2009 for a survey). In the context of organizations and hierarchical relationships, fairness is often associated with meritocracy. Meritocracy refers to an environment in which advancement is based on performance and qualifications, and not demographic variables, background, or connections. Meritocracy implies that a typical workplace setting is regarded as fair when more productive workers are paid more than less productive workers and are more likely to be promoted. In turn, income inequalities due to differences in individual achievement are considered fair and as reward for hard work, while inequalities owing to luck are considered unfair (Alesina and Angeletos, 2005; Benabou and Tirole, 2006). In contrast to existing research, which corroborates the view that unequal outcomes do not harm cooperativeness if they are the result of meritocracy (Almås et al., 2010; Cappelen et al., 2013; Almås et al., 2016), we study the extent to which unequal treatment by quotas affects cooperativeness and depends on the meritocratic nature of the quota.

From a methodological standpoint, our study takes a new direction by combining both survey and economic experiment methods to increase the possible insights from either method in isolation. The survey on its own provides insights on opinions from a representative sample and exploits the variation in the participants' work backgrounds to uncover how quotas are perceived in different workplace environments. The experiment on its own provides insights on behavioral reactions in simulated and exogenously manipulated workplace environments. Together, and by building on each other, our survey and experiment improve the generalizability of our quota findings.

2. Representative Survey Evidence

Several large surveys measure general opinions toward affirmative action (AA). In particular, Gallup conducts an annual social science poll, which includes questions to elicit opinions about AA. On the topic of gender, one of their recent polls shows that 67% of Americans support AA for women (Gallup, 2015). By contrast, there are relatively few surveys about gender quotas. In a 2014 global survey of 12,500 senior executives, 45% of respondents expressed support for gender quotas (Grant Thornton, 2014). A large European survey found that more than 40% of Europeans believe that a 50% gender quota for females on listed companies is a realistic target (European Commission, 2012). Meanwhile, 22% of men and 49% of women support quotas for women in parliament in a recent Australian survey (Australian Broadcasting Corporation, 2016).

Our large survey complements the existing evidence in at least three ways. First, our survey was designed to capture opinions about gender quotas for leadership positions more generally (instead of only for parliament or corporate boards). Thus, we investigate opinions on gender quotas in a much broader space. Second, we analyze opinions based on the respondents' own workplaces. Thus, we start looking beyond the surface and take into consideration that opinions systematically differ across workplaces. Finally, we present respondents with systematically different workplace environments to further investigate the drivers behind opinions about gender quotas.

2.1. Survey Methodology

We commissioned a national survey of 1,011 US residents eliciting opinions toward gender quotas for leadership positions.[4] The sample was representative of the US population according to the US census regarding gender, age, ethnicity, and region. To ensure the validity of the results, (*i*) subjects were initially screened to ensure they resided in the US and were above the age of 18 years, (*ii*) three attention checks at different points in the survey were undertaken to discourage respondents from randomly selecting responses,[5] and (*iii*) respondents could not

[4] The survey was undertaken by Qualtrics using its survey platform.

[5] For instance, directly after eliciting opinions about quotas, respondents were told "Thank you for answering the prior questions. You will now be asked a number of attitude and demographic questions. To proceed, please select 'strongly disagree'." Subjects that failed to answer with strongly disagree where ejected from the survey without payment.

quickly click through the survey without reading questions, as they could move to the next question only after 12 seconds. The survey started by eliciting respondents' general opinions toward gender quotas for female representation in managerial leadership positions, by using a 5-point Likert scale asking to the extent to which they approved whether *"Gender quotas should be used to increase the number of women in leadership positions*."[6]

To study the role of the workplace environment and the presence/absence of gender skill gaps as well as disadvantages against women in the leadership selection procedure, we asked our respondents to what extent gender quotas should be used to increase the number of women in leadership positions in the following three workplace environments.

- Workplace environment with skill gap: *"Suppose that female candidates are on average less qualified for a certain leadership position than male candidates and there is no bias against female candidates in the selection process relative to male candidates."*
- Workplace environment with no difference: *"Suppose that female candidates are on average equally qualified for a certain leadership position relative to male candidates and there is no bias against female candidates in the selection process relative to male candidates."*
- Workplace environment with disadvantage: *"Suppose that female candidates are on average equally qualified for a certain leadership position relative to male candidates but there is bias against female candidates in the selection process relative to male candidates."*

The ordering of these three specific labor market environment questions was random.[7] Finally, the survey included standard demographic questions as well as the participants' perceptions of skill gap and disadvantage in their own profession. Precisely, we asked respondents the extent to which they agreed with the following two statements.

- "*In your profession, female candidates are on average less capable of taking leadership roles than male candidates.*"

[6] Possible responses are "strongly agree," "agree," "neither agree nor disagree," "disagree," and "strongly disagree."
[7] Preliminary analysis suggests that there are no significant order effects. A detailed analysis is available from the authors upon request.

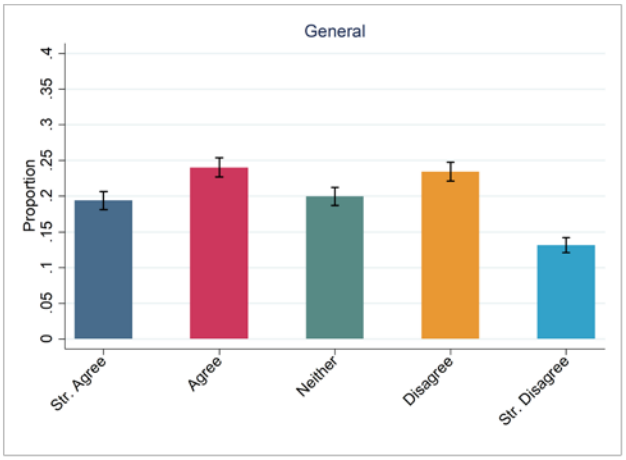
- *"In your profession, there is bias against female candidates in the selection process of leadership positions relative to male candidates."*[8]

The questions correspond to our work place environments. The first question is aimed at understanding whether respondents believe that a skill gap exists between male and females within their professions, while the second question is aimed at eliciting opinions about disadvantage within their professions.

2.2. Survey Findings

Figure 1 provides an overview of *general* opinions about quotas for female managerial positions. The figure shows that opinions are strongly divided if the workplace environment is not specified. We observe that a similar percentage of respondents express either positive (44%) or negative opinions (37%) toward a female quota for leadership positions.

Figure 1: General opinions about gender quotas in the United States



Notes: The figure represents average general opinions toward gender quotas. Confidence intervals on each bar illustrate significance at the 10% level. Data are from 1,011 representative US residents.

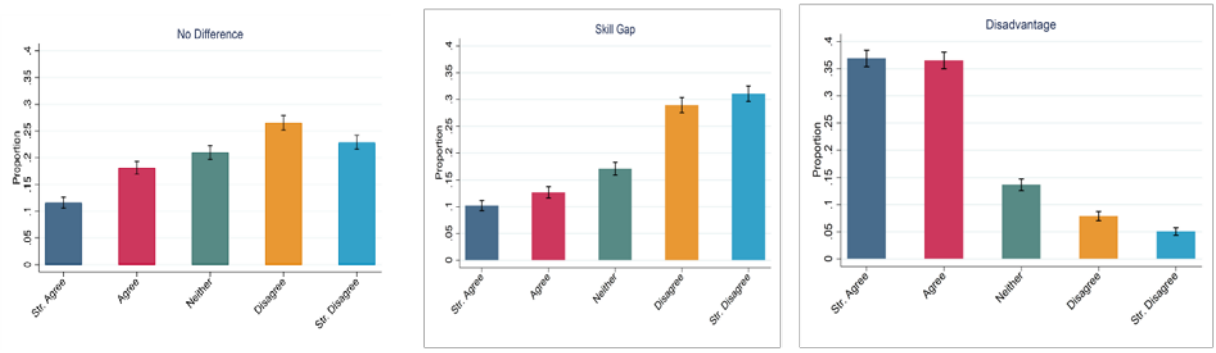
Figures 2a–c reveal that there is significantly more consensus about quotas once we specify the workplace environment. Figure 2a shows that 51% of respondents oppose gender quotas when there is no discrimination against women and when there are no gender differences in performance (*No Diff.*). The opposition to gender quotas is even more pronounced in an

[8] Possible responses are the same as those in our non-occupation-related opinion questions.

environment in which there is no discrimination against women and they underperform relative to men (Figure 2b). In this case, around 60% of respondents oppose gender quotas (*Skill Gap*). However, in Figure 2c, we clearly observe majority support for gender quotas in an environment in which women are disadvantaged in the leadership selection process. More than 70% of the respondents agree that gender quotas should be used and only 13% disapprove. Table 1 (panel A) shows that the differences in opinions across workplace environments are highly significant.[9]

{INSERT TABLE 1 ABOUT HERE}

Figures 2a–c: Opinions toward quotas in the United States



Notes: Figure (a) left: average opinions toward quotas for women in an environment in which there are no gender differences in skill and no disadvantage against women; Figure (b) center: opinions toward quotas for women in an environment in which women on average have lower skill than men and there is no disadvantage against women; Figure (c) right: opinions toward gender quotas in an environment in which there is disadvantage against women and there are no gender differences in skill. Each bar contains confidence intervals, which illustrate significance at the 10% level. Data are from 1,011 representative US residents.

Our representative survey provides first suggestive evidence that respondents' own profession is related to the expectation of gender performance difference and gender discrimination.[10] Specifically, when we examine responses to the general question by occupation group, we observe that 57.7% of respondents in office and admin occupations express positive

[9] Specifically, we take the mean of responses that range from 1= strongly agree to 5= strongly disagree. As expected, the mean response to the base question is 2.87, which is in-between the two extreme categories. By contrast, the survey responses to opinions about the skill gap and no difference are 3.58 and 3.31 respectively, indicating a strong dislike for quotas within these environments. The mean response for the disadvantage question is 2.08. At the mean, all questions are statistically different from each other at $p<.01$. Results are consistent when split by gender, the key exception being that females are more likely to agree with quotas.

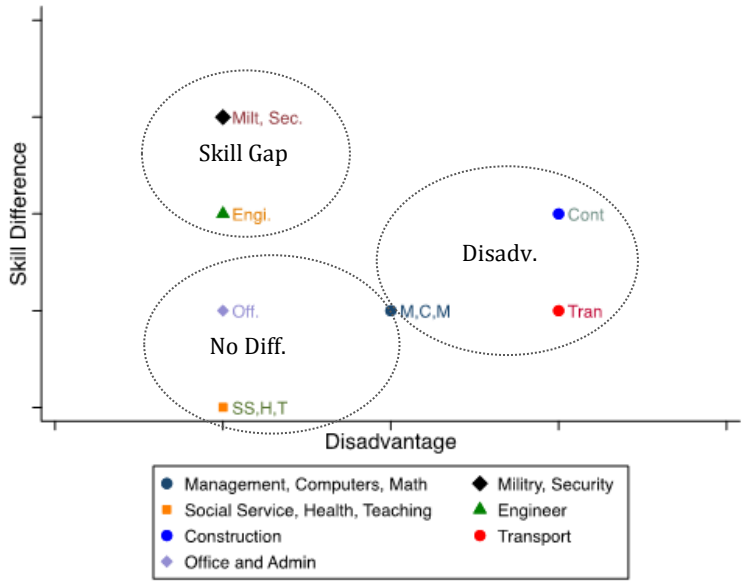
[10] To elicit occupation, we use a variant of the US census aggregate classification of occupation.

opinions about quotas, 50% do so in math and computers, but only 45.1% do so in health, 37.5% in teaching occupations, and 35.1% in construction.

Importantly, we observe that the respondents perceive their own profession differently in terms of gender skill gaps and gender disadvantages. Figure 3 illustrates the perceptions of employees in several key occupations.[11] The vertical axis plots median responses to the occupation-based skill-gap question and the horizontal axis plots median responses to the occupation-based disadvantage question. Employees in social service, health, and teaching occupations are much more likely than those in engineering, military, and protection occupations to believe that female candidates within their profession are similarly capable to male candidates. Meanwhile, employees in transportation and construction occupations are more likely to believe that female leaders suffer from bias in the leadership selection process. The key take-away message from this graph is that occupations are perceived to differ significantly on the performance and disadvantage dimensions. Many occupations show consistent characteristics and fall within the three distinct environments. Service occupations, such as teaching, social services, health, and office admin, are characterized by low relative skill gap and disadvantage, and thus, these occupations are consistent with our no-difference environment.[12] Military, protection, and engineering occupations are characterized by employees' belief that a skill gap exists, and as such, these occupations are more consistent with the skill-gap environment. By contrast, respondents in construction and transportation occupations are more likely to believe that a large relative disadvantage exists, which is consistent with our disadvantage environment.

[11] Responses from all occupation groups are in Appendix Table A1.
[12] It is plausible that such occupations as health and teaching have a skill gap favoring females or that males are disadvantaged.

Figure 3: Perceptions of skill gap and disadvantage within occupations



Notes: This figure illustrates perceptions of labor market conditions in respondents' own profession. For simplicity, a subset of occupations is shown; for a full list of occupations see Appendix 1, Table A1. Markers with more than one occupation, such as social services, health, and teaching, have (independently) the same median skill gap and disadvantage. We do not collapse occupation groups into separate categories. Data are from 1,011 representative US residents who indicate they participate in the labor force.

These survey findings provide suggestive evidence that gender quotas might affect organizational performance, assuming that opinions translate into behavior in the workplace. To investigate whether hierarchical relationships are affected by gender quotas and to test potential underlying channels, we design and conduct a set of laboratory experiments. A key difference between our survey and the experiment is the sample population. Our participants in the experiment are not nationally representative; they are all university students from an Australian university (Monash), young (89% are between 18 and 24 years), and have little workplace experience (70% are not in paid employment).

To identify the role of sampling on opinions on gender quotas, we conduct the same US representative survey with a sample of 171 students from the Monash laboratory subject pool.

Like the US survey, the Monash survey was conducted online, using the Qualtrics platform. Each subject received $10 for successfully completing the survey. Average opinions are reported in Table 1 (panel B). In general, we find very little difference in perceptions about quotas across our two samples, suggesting that our survey findings are robust.

3. Experimental Design in the Laboratory

In this section, we present our experimental design. We randomized the implementation of gender quotas and manipulated the perception of gender performance differences and discrimination against women not only to test the extent to which quotas affect hierarchical relationships but also to test different economic models.

There are two main parts in our experiments. The first consists of an arithmetic task and the second a gift-exchange game. The two parts are connected in two important ways. First, the performance in the first part co-determines who is promoted to the role of manager in the second part. Second, the task performances of the managers reflect managerial ability in the sense that they affect the payoffs of managers and their workers in the second part.

We used a 3-by-2 experimental design. Specifically, there are three experiments, each consisting of two treatments. The experiments and treatments were determined at the start of the second part. Subjects knew that there were two parts but they did not know the details of the second part until they finished the first part. The experiments varied the information that was given to the subjects and the treatments varied in the manager selection procedure. After each experiment, subjects took part in an incentivized belief elicitation task and a short post-experimental survey.

The experimental sessions were conducted at the Monash University Laboratory for Experimental Economics and lasted for approximately 70 minutes. Subjects earned an average A$33.40. The experiment was computerized using Z-Tree (Fischbacher, 2007) and subjects were recruited using ORSEE (Greiner, 2015). Subjects were required to answer a set of control questions correctly before the start of the experiment.

3.1. Part 1: The Arithmetic Task

In the first part, each subject was given 5 minutes to complete as many sums of five randomly chosen two-digit numbers as possible. We informed subjects that the performance in

this task would increase their chance of higher earnings in the second part. We did not inform subjects exactly how their performance would affect earnings because the impact is treatment-dependent and announcing it might have influenced performance differently across treatments. Moreover, we did not inform subjects about their actual performance in this task (i.e., how many sums they correctly answered).

We chose the arithmetic task for three main reasons. It provides a simple and objective performance measure, it has been used in related experiments (Niederle and Vesterlund, 2007), and there is variance in the extent of gender differences in performance in this task in the literature, which we made use of in our different treatments.[13]

In order to reduce gender saliency in the arithmetic task, subjects were never asked about their gender. Rather, before entering the laboratory, each subject was given a paper slip with a random number on it. Random odd (even) numbers were given to female (male) subjects. Subjects then entered their random numbers into the computer such that the computer could assign gender without knowing the subject's identity.

3.2. Part 2: The Gift-exchange Game

After the subjects finished the arithmetic task, we randomly assigned them into groups. Each group consists of two males and two females. There are two managers and two workers in each group. We paired each manager with one worker of the opposite sex, and thus, there are two manager–worker (hierarchical) pairs in each group. We employed a *partner design*, implying that each manager always interacted with the same worker throughout the second part.

The second part of the experiment consists of a gift-exchange game, which lasted for 15 periods (Fehr et al., 1993; Charness and Kuhn, 2011). In each period, managers first chose a wage level and then workers responded with an effort level. Specifically, at the beginning of each period, each manager received an endowment of 240 points and then chose a wage between 20 and 120 points for the worker s/he was paired with. Thereafter, each worker was informed about the wage and decided how much effort to provide for her/his manager. Higher effort

[13] Some studies report no statistically significant gender differences (Niederle and Vesterlund, 2007), while others identify significant differences (Cameron et al., 2015).

results in higher payoff for the manager but is at the same time costly for the worker. Table 2 shows the costs in points for each effort level.

{INSERT TABLE 2 ABOUT HERE}

After the worker chose her/his effort level, the manager was informed of it before the next period began. The base payoffs are:

- *Manager's payoff per period = (240 – wage) × effort*
- *Worker's payoff per period = (wage – effort cost – 20).*

A novel feature of our gift-exchange game is that final payoffs not only depend on the endowment, wage, and effort, but also on the arithmetic task performance. In particular, the payoffs of the manager and worker from the gift-exchange game were multiplied by the manager's arithmetic task (number of correct sums) at the end of the experiment to obtain the final payoff. This feature captures the reality that candidates are usually chosen as managers based on their performance in tasks that predict their performance as managers, and, in turn, the organization's performance and the managers' and workers'' wages. Making both a manager's and a worker's payoffs a function of the manager's performance in the first part accounts for these relationships in the simplest possible manner. In addition, it makes the management selection procedure transparent and objective for the subjects. At the end of the experiment, the points that subjects earned during the second task were converted into Australian dollars at an exchange rate of 100 points = A$1. We randomly selected 3 out of the 15 periods for payment.

We chose the framework of the gift-exchange experiment because it provides the most basic set-up of a hierarchical workplace relationship and has been widely studied (Fehr et al., 1993; Charness and Haruvy., 2002 Charness and Kuhn, 2007, 2011), facilitating the classification of our findings. The game also represents an analogy of hierarchical relationships in the field beyond wage-setting and effort provision. The wage chosen by the manager could be interpreted as how much time, effort, and attention a manager is willing to give to a worker and how much effort the worker puts in could be regarded as how cooperative and compliant workers are with their managers.

3.3 Implementation of Gender Quotas: Two Manager Selection Procedures

Subjects in each of our three experiments were randomly assigned into one of two manager selection procedures: *no quota* and *quota*. Under the *no-quota* procedure, the two subjects (in the group of four) with the highest task score (the number of sums answered correctly) were selected as managers. If two or more subjects had the same score, the computer randomly selected one of them to be the manager. Under the *quota* procedure, one of the manager roles was reserved for the highest-scoring female. The subject with the highest score out of the remaining three in the group became the other manager, regardless of gender.

While the implemented gender quota guarantees female representation at the manager level, its impact is relatively minor in our experiments. The quota only re-allocates 9.6% of the manager positions in our experiments. To capture typical promotions in the field, note that we never revealed to the subjects whether a woman had become a manager because of the quota, nor did we inform the promoted women whether they were promoted because of the quota. Workers and managers knew only that there was a positive probability that a female manager was promoted because of the quota.

3.4 Different Backgrounds for Gender Quotas: Three Experiments

We randomized subjects into three experiments (*skill gap, no difference,* and *disadvantage*) to capture the three environments that allow us to assess how gender performance and gender disadvantage affect the impact of quotas on hierarchical relationships. While the treatments and basic structure of parts 1 and 2 are the same, the experiments vary in the information given to the subjects on gender differences in task performance and whether there is bias against women's task performance in the selection process. The three experiments enabled us to disentangle reactions to a gender quota: if (*1 – "skill gap"*), there is an average gender difference in the task that determines the selection of manager positions; if (*2 – "no difference"*), there are no gender differences of any kind; and if (*3 – "disadvantage"*), there is a disadvantage against women in the task during the selection process. In all three experiments, we told our subjects that we would count the number of questions correctly answered in the arithmetic task. The key difference between the three experiments is the additional information we provided

before the start of task 2. The second quadrant in Appendix Figure A1 summarizes our 3-by-2 experimental design.[14]

In the *skill-gap* experiment, we informed the subjects that we observed gender differences in arithmetic task performance in a previous experimental session and that women solved on average 20% fewer arithmetic problems than did men. By contrast, in the *no-difference* experiment, we informed the subjects that in the previous experimental session we observed no significant gender differences in arithmetic task performance and that women and men solved approximately the same number of arithmetic problems. Because we held the selection procedure constant and used the same performance task, we could cleanly identify the role of information on average gender differences in performance. The information provided on task performance is based on actual pilot experiments.[15]

In the *disadvantage* experiment, women were disadvantaged and this was common knowledge among the subjects. Specifically, we informed subjects that, for the purpose of manager selection, performance in only the first 4 of the 5 minutes in the arithmetic task were counted for women but all 5 minutes were counted for men. In other words, the score for women was based on the number of sums answered correctly in 4 minutes whereas the score for men was based on the number of sums answered correctly in 5 minutes. We chose the level of disadvantage in this treatment (1 out of 5 minutes) such that it corresponds to the performance gap (average of 20%) in the performance treatment. In addition, we informed subjects that we did not find gender differences in the arithmetic task performance in the previous experimental session. For the purpose of payoff, all 5 minutes of the task were counted regardless of gender, and this fact was known to the subjects. Therefore, this environment represents a scenario in which women are not worse managers but are discriminated against in the selection process.

The goal of the three experiments is to investigate reactions to gender quotas when there is no uncertainty about the background of their implementation. Outside the laboratory

[14] It is entirely plausible that society or professions are located in another quadrant, but for the purpose of this study, we focused specifically on the second quadrant, as this is the most relevant for female quotas.

[15] We conducted four pilot experimental sessions before our three experiments, which included the arithmetic task. There were 24 subjects in each session. In two of the four sessions, we did not observe significant gender differences in task performance (females scored 9.6 and 9.9 compared to males, who scored 12.3 and 12.6, respectively). By contrast, in the other two sessions, we observed that women performed on average 20% less (females scored 12.8 and 10.3 on average compared to males, who scored 12.6 and 11.2, respectively).

environment, there are typically multiple reasons to implement gender quotas. In particular, the reason could be that gender quotas are the result of observed gender differences in ability and these gender differences may or may not be caused by disadvantage.[16] Importantly, the perception of the labor market environment and whether there are perceived gender differences in performance and treatment may differ amongst individuals and be linked to gender and position in an organization.

In our experiments, we not only made the labor market environment transparent and thereby aligned perceptions, we also made the determinants of the selection criteria transparent. We simplified the selection procedure as much as possible by making promotion dependent on only one performance measure—the score in the arithmetic task. This simplification has the advantage that the best applicants are always those with the highest task scores and whether quotas decrease or increase the quality of the successful applicant becomes objective.

3.5 Belief Elicitation and Survey

After the final period of part 2, we asked subjects to guess the average female and male arithmetic task scores in their session, as well as their own scores. We incentivized subjects to guess accurately by rewarding them with A$1 for each response that did not deviate more than 1 point from the actual score. The belief elicitation allowed us to check whether the treatments actually influenced subjects' beliefs about the arithmetic task performance of males and females. After the experiment, subjects completed a short survey that allowed us to obtain their standard demographic information for our regression analysis.

3.6. Theoretical Background: Fairness, Marginal Productivity of Labor, and Gender Quotas

Predictions of the standard neoclassical model are independent of experiment and treatment. In the first part, subjects attempted to solve as many sums as possible. In the second part, workers chose the lowest possible effort level (0.1) and managers chose the lowest possible wage (20). However, considerable evidence from gift-exchange experiments shows that managers offer more than the minimum wage and workers reciprocate by offering more than the

[16] According to a 2014 US Pew survey 33% of men and 52% of women believe that there are less women in top business positions because women are held to a higher standard than men (Pew, 2015).

minimum effort level. Importantly, standard reciprocity predictions do not predict differences across our various environments or treatments.

The experiments capture three stylized environments in which the different manager selection procedures can be linked to fairness considerations. We focused on meritocracy as our main fairness consideration.[17] In our framework, the meritocratic nature of the selection procedure is equivalent to the probability that the best two candidates are promoted as managers in each group of four. The selection process is expected to be more meritocratic if a higher-performing candidate is selected as manager.

Table 3 summarizes the predictions for the workers' effort choice. In the *skill-gap* experiment, the quota was implemented to increase female representation at the managerial level in an environment in which women perform on average worse than men. According to our definition of meritocracy, the *quota* treatment is unfair, because it allows the possibility that women with lower task scores are promoted instead of men with higher task scores, whereas the *no-quota* treatment always selects the best candidates in this setting. Similarly, in the *no-difference* experiment, the quota was implemented in an environment in which there are no gender differences in performance. In this experiment, it remains possible that women with lower task scores were promoted instead of men with higher task scores. The *quota* treatment reduces the meritocratic nature of the selection procedure compared to the *no-quota* treatment, which always promotes the best candidates.

{INSERT TABLE 3 ABOUT HERE}

In the *disadvantage* experiment, the *quota* treatment can improve the meritocratic nature of the selection procedure relative to the *no quota* treatment because women are discriminated against and there are no average gender differences in performance. In the presence of such discrimination, the *no-quota* procedure might naively promote men with lower actual task score (but higher observed score) instead of women with higher actual task score (but lower observed score), whereas this is much less likely under the *quota* treatment, which helps correct the discrimination.[18] In summary, if meritocracy is of concern to our subjects, then we should

[17] As discussed in Starmans et al. (2017), a meritocratic system is a key component of a fair society.
[18] Specifically, the quota helps to undo the disadvantage for females who would not become managers if only 4 minutes of their score were counted but would become managers if all 5 minutes were counted.

observe that implementing quotas lead to lower wage and effort levels in the *skill-gap* and *no-difference* environments, and higher wage and effort levels in the *disadvantage* environment.

Another important dimension of gender quotas is how they relate to the marginal productivity of labor. If gender quotas affect the likelihood with which the best applicants are promoted to managerial positions, they not only affect meritocracy but also can affect the marginal productivity of worker effort when workers' payoff is linked to their managers' performance. For example, if workers anticipate that gender quotas decrease the quality of the managers selected, then workers know that their marginal unit of effort translates into less expected surplus. Since quotas make it easier for females to become managers and more difficult for males to become managers in all environments, a male manager under the *quota* treatment is expected to be a better performer than the average male manager under the *no-quota* treatment. Similarly, a female manager under the *no-quota* treatment is expected to be a better performer than the average female manager under the *quota* treatment. Thus, if workers work harder when the expected marginal productivity of labor is higher, then we should observe that workers with male managers provide more effort in the *quota* treatment than in the *no-quota* treatment, regardless of the environment. As a result of lower expected marginal productivity of labor, we should observe that workers with female managers should provide less effort in the *quota* treatment, regardless of the environment.

4. Experimental Findings

4.1. Overview

In total, 516 subjects (258 females and 258 males) took part in our three experiments (N=152 in *skill gap*, 172 in *no difference*, and 192 in *disadvantage*). Overall in Task 1, males solved on average slightly more sums than did females (11.39 vs. 10.60, respectively, p=0.021), a difference that falls well within the range of previous literature, which often finds minor gender differences in performance. In Task 2, in accordance with previous evidence, wages and effort levels are higher than predicted by standard economic theory. Overall, managers on average offer a wage of 84.4 (minimal possible wage = 20, maximal possible wage = 120) and workers' average effort level is 0.56 (minimal possible effort level = 0.1, maximal possible effort level = 1). The minimal (maximal) possible wage was selected as 3.7% (10.5%) at the time. The minimal (maximal) effort level was chosen as 14.6% (10.9%) at the time. Both wages and effort

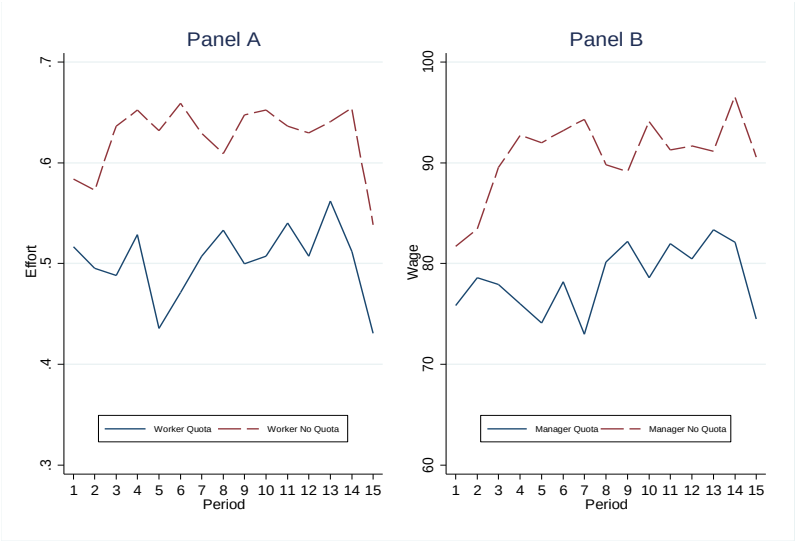
levels remain stable across periods, except for the last period. In period 15, both wages and effort levels significantly decline, which is consistent with the endgame effect.

4.2. Skill Gap: The Impact of Quotas with Gender Performance Difference

In the *skill-gap* experiment, subjects were informed that females solved on average 20% fewer arithmetic task questions than did males in the previous session. Our incentivized belief question after the experiment provides evidence that subjects believed this to be true. Subjects in this experiment believed that on average, males are 22.9% better than females in the arithmetic task.

Figure 4 illustrates the average effort level (panel A) and wage (panel B) in each period for both treatments. The red dashed line illustrates the patterns under *no quota* and the blue solid line the patterns under *quota*. In panel A, average effort levels are clearly lower if there is a quota and the difference between treatments is large in all periods. In panel B, average wages are clearly lower if there is a quota and the difference between treatments is large in all periods.

Figure 4: Effort and wages in the skill-gap experiment



Notes: Blue solid (red dashed) lines illustrate behavior under *quota* (*no quota*). Data are from the laboratory skill-gap experiment.

The treatment differences in wage and effort levels are substantial and statistically different. Managers offered a wage of 90.75 across all periods under *no quota* but only 78.47 under *quota*, which is a significant wage reduction of 13.5% (p=0.008, n=86; Mann–Whitney

two-tailed test).[19] Similarly, we found that workers have higher effort levels under *no quota* (0.625) than under *quota* (0.50), a difference of approximately 25% (p=0.012, n=86).

These differences are highly significant in a generalized least square (GLS) random-effects models reported in Table 4, which control in particular for *gender* (1 if the subject is male, 0 otherwise), *belief on partners' performance* (worker's belief about the average performance of the opposite sex in the arithmetic task), *belief on own* performance (belief about one's own arithmetic score), and a *period trend*. Standard errors were clustered at the group level. In the models, we found that workers in the *quota* treatment exerted 13.2-percentage points less effort than did their counterparts in the *no-quota* treatment (Table 4a, column 1, p<0.001). Furthermore, we found that wages are 13.4 points lower under *quota* (Table 4b, column 1, p<0.001). These findings are independent of the gender constellations (male manager–female worker, female manager–male worker) as we find that male and female managers offer lower wages under *quota* (11.9 and 13.4 points, respectively) and similarly, male and female workers provide lower effort (17.6% and 9.7%, respectively) (Tables 4a and b, columns 2 and 3).

{INSERT TABLES 4a & 4b ABOUT HERE}

In the preceding analysis, we focused on the impact of quotas on the functioning of hierarchical relationships and the interplay of managers and workers. However, it is also possible to provide an estimate of the impact of quotas on worker performance per se, that is, independent of wage. To do this, we controlled for the manager's wage offering in our regressions. We estimated a GLS random-effects model with effort as the dependent variable and quota as the key variable of interest. The results are reported in Table 4a, column 4. We found that after the inclusion of our standard set of controls and the wage offered by the leader, quotas marginally reduce worker performance relative to *no quota* by 4.6 percentage points (p=0.075). When the dependent variable was the wage offered by the manager and it was regressed on effort received in the last period (Table 4b, column 4), we found that wages are 7.3 points lower under *quota* (p<0.001). These findings are consistent with our meritocracy hypothesis and lead to our first main result.

[19] Unless otherwise reported, the behavior in all periods from one subject constitutes one independent observation.

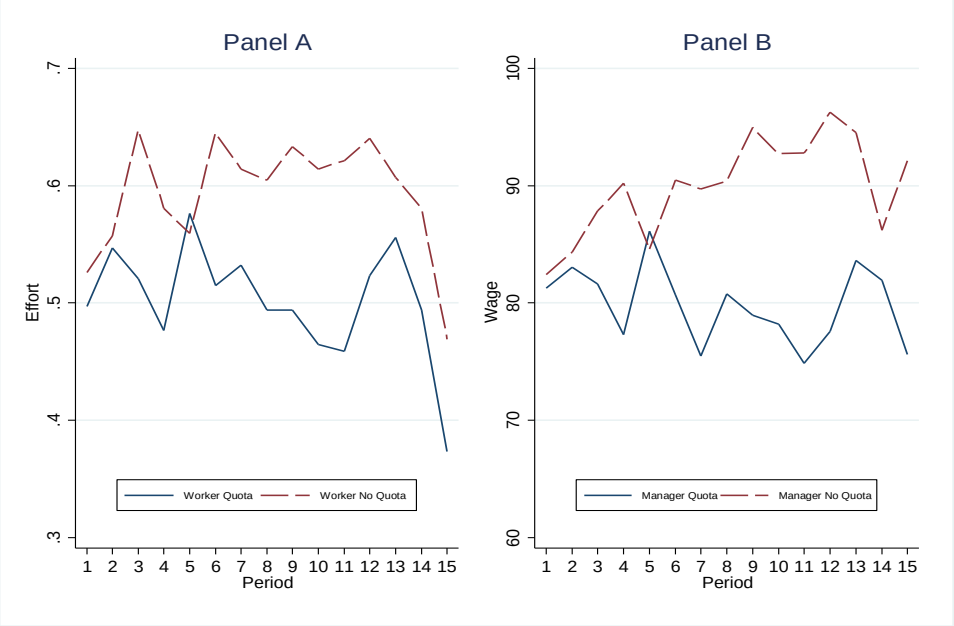
Result 1: *Female manager quotas cause lower effort and wages in an environment in which women are on average less suited to become managers and are not disadvantaged.*

4.3. No Difference: The Impact of Quotas without Discrimination and Gender Performance Differences

In the *no-difference* experiment, subjects were informed that there is no average gender performance difference in the arithmetic task. Our incentivized belief question provides suggestive evidence that subjects' actual beliefs about performance correspond to the information given. Subjects in this experiment believe that males are on average only 3.6% better than females in the arithmetic task, a significantly lower difference than in the *skill-gap* experiment ($p<0.001$).

Figure 5 illustrates the average effort level (panel A) and wage (panel B) in each period for both treatments. In panel A, average effort levels are generally lower if there is a quota and the difference between treatments remains substantial over the course of the experiment. In panel B, average wages are generally lower if there is a quota and the difference between treatments appears to increase over periods.

Figure 5: Effort and wages in the *no-difference* experiment



Notes: Blue solid (red dashed) lines illustrate behavior under *quota* (*no quota*). Data are from the laboratory no-difference experiment.

The differences between treatments over time are significant. Managers offered a wage of 89.97 across all periods when there is no quota but only 79.81 when there is a quota, a significant wage reduction of 11.3% (p=0.023). Similarly, workers have higher effort levels in the absence of a gender quota (0.59, n=42) than when there is a gender quota (0.50), a difference of almost 20% (p=0.045). These differences are robust in our random-effects models with controls. In column 1 of Table 5a, *quota* is significantly negative (p=0.009), showing that effort levels are 12 percentage points lower under *quota*. Similarly, introducing a *quota* reduces wages by 9.6 points (p=0.026) if we use wages as the dependent variable (Table 5b, column 1). These findings are independent of the gender constellations, as both male–female and female–male manager–worker pairs are responsible for the overall results (see columns 2 and 3 of Tables 5a and 5b).

{INSERT TABLES 5a & 5b ABOUT HERE}

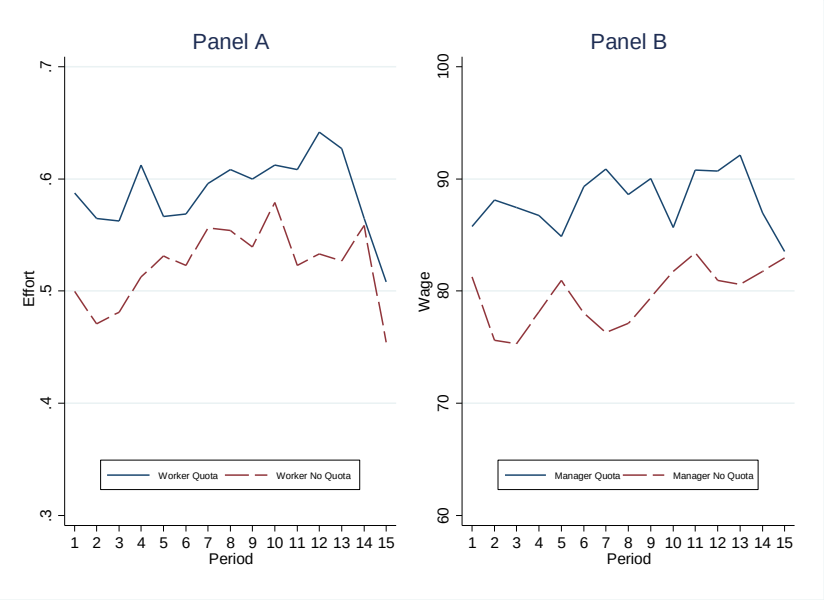
In Table 5a, column 4, in which we used our standard set of controls and the wage offered by the leader, *quota* significantly reduces worker performance relative to *no quota* by 5.0 percentage points (p=0.030). The results are similar when we replace effort with wage, with managers offering 5 points less for a given effort level (p=0.014, Table 5b, column 4). Thus, consistent with our meritocracy predictions, our second result is as follows.

Result 2: *Female manager quotas cause lower effort and wages in an environment in which women are equally suited to become managers and they are not disadvantaged.*

4.4. Disadvantage: The impact of Quotas when Women are Disadvantaged

Figure 6 illustrates average effort levels (panel A) and average wages (panel B) in each period for both *disadvantage* treatments. In stark contrast to the *skill-gap* and *no-difference* environments, we observe that average wages are *higher* under *quota* and this holds on average across all periods (panel A). Moreover, we observe that average effort levels are always higher under *quota* than under *no quota* (panel A).

Figure 6: Effort and wages in the *disadvantage* experiment



Notes: Blue solid (red dashed) lines illustrate behavior under *quota* (*no quota*). Data are from the laboratory disadvantage experiment.

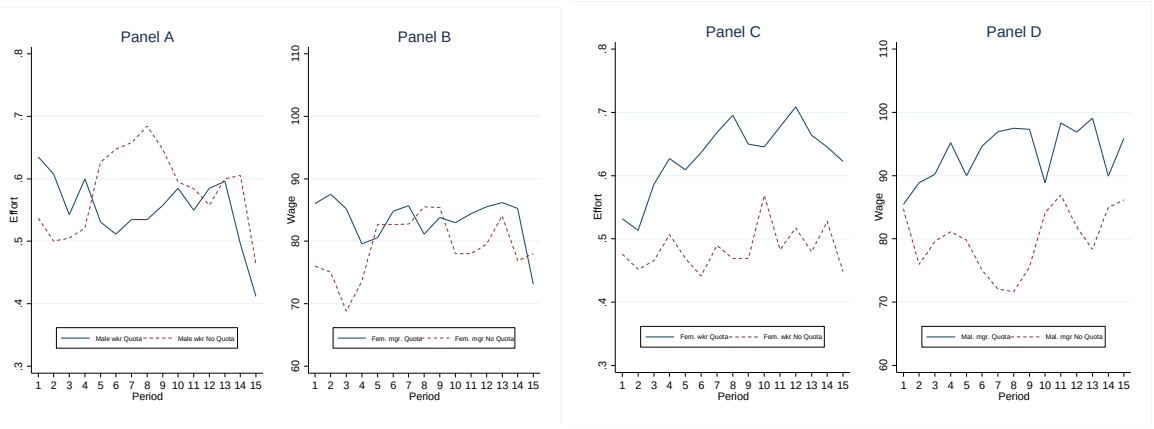
{INSERT TABLES 6a & 6b ABOUT HERE}

Managers offer an average wage of 79.79 across all periods under *no quota* but 88.14 under *quota*, a significant wage increase of more than 10% (p=0.033, n=96). Similarly, workers have higher effort levels under *quota* (0.59) than under *no quota* (0.521), a difference that is large in magnitude but not significant (p=0.213, n=96). These differences are more pronounced in random-effects models (Tables 6). Workers in the quota treatment exert 6.5% less effort than their counterparts do in the *no-quota* treatment, a difference that is significant at p=0.058 (Table 6a, column 1). Furthermore, wages are 9.5 points lower under *quota* (Table 6b, column 1, p=0.006). Controlling for wages, *quota* insignificantly increases effort by 1.4 percentage points (p=0.384, Table 6a, column 4). On the other hand, manager's wage increases by 6.5 points (p=0.001) under *quota* for a given lagged effort level (Table 6b, column 4). These results are largely consistent with the meritocracy hypothesis.

Result 3a: *Female manager quotas cause on average higher effort and wages in an environment in which women are disadvantaged and equally suited to become managers.*

Interestingly, this finding is not independent of the gender constellations. Figure 7 illustrates effort and wages depending on gender constellation (male manager–female worker, female manager–male worker) and the presence of a gender quota. Comparing panels A to C, effort levels in the quota environment only seem to be significantly higher in panel C (female workers) but not in panel A (male workers). Similarly, wages only seem to be higher in panel D (male managers) but not in panel B (female managers). Tables 6a and 6b corroborate these patterns and show that the average findings on higher effort and wages in the quota environment of the *disadvantage* experiment are entirely driven by the male manager–female worker pairs. The female manager–male worker pairs do not behave differently in the presence and absence of the quota (p=0.442 (female manager, see column 2 in Table 6a) and 0.445 (male worker, see column 3 in Table 6b) whereas the male manager–female worker pairs clearly change (p<.01, see column 3 in Table 6a and column 2 in Table 6b).

Figure 7: Behavior in the *disadvantage* experiment by gender



Notes: Panel A illustrates male workers' effort under *quota* and *no quota* while panel B illustrates female managers under *quota* and *no quota*. Panel C illustrates female workers' effort under *quota* and *no quota* while panel D illustrates male managers under *quota* and *no quota***.** Data are from the laboratory *disadvantage* experiment.

Result 3b: *Female manager quotas do not cause higher effort and wages for male workers in an environment in which women are disadvantaged and equally suited to become managers.*

5. Conclusion

Hierarchical relationships are key features of many organizations and determine organizational performance. There are many dangers to the functioning of hierarchical relationships. In particular, there is manifold evidence that fairness plays an important role when superiors interact with subordinates and conventional wisdom suggests that subordinates are less motivated if they question the selection mechanism of superiors. Persistent gender imbalances in managerial positions suggest that women are disadvantaged, providing scope for improving such selection mechanisms. Female gender quotas are a means to counter these disadvantages but they might also cause men to be disadvantaged.

This study navigated this challenging context to investigate the impact of gender quotas on the functioning of hierarchical relationships in different workplace environments. We identified significant reactions to gender quotas, which are crucially determined by the workplace environment. If the workplace disadvantages women, female quotas improve hierarchical relationships; however, if there is no such disadvantage, female quotas harm hierarchical relationships. These findings are particularly important, as we found in our representative survey that the perception of gender disadvantages differs across occupations. Thus, we believe that it is beneficial to evaluate the perception of the labor market environment and whether workers believe that gender quotas undermine the principle of meritocracy in the workplace *before* considering their implementation. Implementing quotas in workplaces in which employees do not perceive gender disadvantages could backfire and harm the organization, also hurting those who are meant to benefit from the quota.

Our study constitutes a first attempt to study the links between gender quotas, hierarchical relationships, and workplace environments. We believe there is merit in complementing our findings with field data on the dynamics of hierarchical relationships before and after the introduction of gender quotas. Furthermore, we believe that it is worthwhile to investigate the drivers behind perceptions of gender disadvantages, how they are determined by the workplace environment, and the extent to which the introduction of gender quotas affects perceptions. In particular, it seems important to verify whether employees perceive less gender disadvantages after the introduction of a gender quota and to relate this to the functioning of hierarchical relationships.

Acknowledgments

Leibbrandt acknowledges support from the Australian Research Council. Vecci acknowledges support from the Swedish Research Council (Project no. 348-2014-4030). For their insightful comments and suggestions, we thank audiences at the Australia and New Zealand Experimental Economics Workshop, Australian Conference of Economists, Autonomous University of Barcelona, University of Cologne, Economic Science Association Meetings, Economics and Social Research Institute (Ireland), University of Essex, Fordam University, GATE Lyon, University of Gothenburg, University of Granada, Lund University, University of Melbourne, Monash University, University of Sydney, Society of Advancement of Behavioural Economics Conference, RMIT University, and University of Vienna.

References

Adams, R.B. and Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics 94*(2), 291–309.

Ahern, K.R., and Dittmar, A. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *Quarterly Journal of Economics* 127(1), 137–197.

Akerlof, G. (1982). Labor contracts as partial gift exchange. *Quarterly Journal of Economics* 97, 543–569

Akerlof, G. and Yellen, J. (1990). The fair wage-effort hypothesis and unemployment. *Quarterly Journal of Economics* 105(2), 255–283.

Alesina, A. and Angeletos, G.M. (2005). Fairness and redistribution. *American Economic Review* 95(4), 960–980.

Almås, I., Cappelen, A.W., Sørensen, E.O., and Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science* 328(5982), 1176–1178.

Almås, I., Cappelen, A.W., and Tungodden, B. (2016). Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians? Discussion Paper No. 18/2016, NHH Dept. of Economics.

Australian Broadcasting Corporation (2016). Women and men split on idea of gender quotas in Parliament. Available at http://www.abc.net.au/news/2016-06-23/election-gender-quotas-parliament-women-split-vote-compass/7533006 (last accessed Feb. 5, 2018).

Balafoutas, L., Davis, B., and Sutter, M. (2016). Affirmative action or just discrimination? A study on the endogenous emergence of quotas. *Journal of Economic Behavior and Organization* 127, 87–98.

Balafoutas, L. and Sutter, M. (2012). Affirmative action policies promote women and do not harm efficiency in the lab. *Science* 335, 579–582.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009), Powerful women: Female leadership and gender bias. *Quarterly Journal of Economics* 124(4), 1497–1540.

Benabou, Roland and Tirole, Jean (2006). Belief in a just world and redistributive politics. *Quarterly Journal of Economics* 12(2), 699–746.

Bertrand, M., Sandra, B., Sissel, J., and Lleras-Muney. A. (2014). Breaking the glass ceiling? The effect of board quotas on female labor market outcomes in Norway. National Bureau of Economic Research, Working Paper, 20256.

Brown, M., Falk, A., and Fehr, E. (2012). Competition and relational contracts: The role of unemployment as a disciplinary device. *Journal of the European Economic Association* 10, 887–907.

Calsamiglia, C., Franke, J., and Rey-Biel, P. (2013). The incentive effects of affirmative action in a real-effort tournament. *Journal of Public Economics* 98, 15–31.

Cameron, L., Erkal, N., Gangadharan, L., and Zhang, M. (2015) Cultural integration: Experimental evidence of convergence in immigrants' preferences, *Journal of Economic Behavior & Organization*, 111, 38–58.

Cappelen, A.W., Konow, J., Sørensen, E.Ø., and Tungodden, B. (2013). Just luck: An experimental study of risk taking and fairness. *American Economic Review* 103, 1398–1413.

Charness, G., Cobo-Reyes, R., Jiménez, N., Lacomba, J., and Lagos. F. (2012). The hidden advantage of delegation: Pareto improvements in a gift exchange game. *American Economic Review* 102(5), 2358–2379.

Charness, G. and Haruvy, E. (2002). Altruism, equity, and reciprocity in a gift-exchange experiment: An encompassing approach. *Games and Economic Behavior* 40(2), 203–231.

Charness, G. and Kuhn, P. (2007). Does pay inequality affect worker effort? Experimental evidence. *Journal of Labor Economics* 25(4), 693–723.

Charness, G. and Kuhn, P. (2011). Lab labor: What can labor economists learn in the lab? *Handbook of Labor Economics*, Volume 4A, O. Ashenfelter and D. Card (eds.), 229–330.

Chattopadhyay, R. and Duflo, E. (2004).Women as policymakers: Evidence from a randomized policy experiment in India. *Econometrica* 72(5), 1409–1443.

Englmaier, F. and Strasser, J. (2014). Worker characteristics and wage differentials: Evidence from a gift-exchange experiment. *Journal of Economic Behaviour and Organisation* 97, 185–203.

European Commission (2012). Special Eurobarometer 376, Women in decision-making positions, Report, Available at http://ec.europa.eu/public_opinion/index_en.htm (last accessed Feb. 5, 2018).

European Commission (2016). Gender balance on corporate boards. Available at http://ec.europa.eu/justice/gender-equality/files/gender_balance_decision_making/1607_factsheet_final_wob_data_en.pdf (last accessed Feb. 5, 2018).

Fehr, E., Goette, L., and Zehnder, C. (2008). A behavioral account of the labor market: The role of fairness concerns. Institute for Empirical Research in Economics, University of Zurich, Working Paper, 394.

Fehr, E., Goette, L., and Zehnder, C. (2009). A behavioral account of the labor market: The role of fairness concerns. *Annual Review of Economics* 1, 355–384

Fehr, E., Kirchsteiger, G., and Riedl, A. (1993). Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics* 108(2): 437–459.

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2), 171–178.

Gallup (2015). Social science poll: Gallup, Inc., http://news.gallup.com/poll/184772/higher-support-gender-affirmative-action-race.aspx (last accessed Feb. 5, 2018).

Gächter, S., Nosenzo, D., and Sefton, M. (2013). Peer effecrs in pro-social behaviour: Social norms or social preferences? *Journal of the European Economic Association* 11, 548–573.

Gilchrist, D., Luca, M., and Malhotra, D. (2016). When 3 + 1 > 4: Gift structure and reciprocity in the field. *Management Science* 62(9), 2639–2650.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics* 118(3), 1049–1074.

Grant Thornton (2014). Women in business: From classroom to boardroom. Grant Thornton International Business Report 2014. Available at http://www.grantthornton.az/content/files/ibr2014-wib-report-final.pdf (last accessed Feb. 5, 2018).

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association* 1(1), 114–125.

Herz, H., and Taubinsk, D. (2018). What makes a price fair? An experimental analysis of market experience and endogenous fairness views. *Journal of the European Economic Association*, forthcoming.

Kahneman, D., Knetsch, J.L., and Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review* 76(4), 728–741.

Konow, J. (1996). A positive theory of economic fairness. *Journal of Economic Behavior & Organization* 31(1), 13–35.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review* 90(4), 1072–1091.

Kube, S., Maréchal, M.A., and Puppea, C. (2012). The currency of reciprocity: Gift exchange in the workplace. *The American Economic Review* 102(4), 1644–1662.

Leibbrandt, A., Wang, L.C., and Foo, C. (2017). Gender quotas, competitions, and peer review: Experimental evidence on the backlash against women. *Management Science*

Matsa, D.A. and Miller, A.R. (2013). A female style in corporate leadership? Evidence from quotas. *American Economic Journal: Applied Economics* 5(3), 136–169.

Niederle, M., Segal, C., and Vesterlund, L. (2013). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science* 59(1), 1–16.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics* 122(3), 1067–1101.

Pew (2015). Women and Leadership: Public Says Women are Equally Qualified, but Barriers Persist. Pew Research Center: Washington, D.CQuota Project (2016). Global database of quotas for women. International IDEA and Stockholm University. Available at http://www.quotaproject.org (accessed May 23, 2016).

Schotter, A. and Weigelt, K. (1992). Asymmetric tournaments, equal opportunity laws and affirmative action: Some experimental results. *Quarterly Journal of Economics* 107(2), 511–539.

Starmans, Christina, Sheskin, Mark, and Bloom, Paul (2017). Why people prefer unequal societies. *Nature Human Behaviour* 1(0082), 1–7.

The Economist (2014). The spread of gender quotas for company boards, March 25.

Thibaut, J.W. and Walker, L. (1975). *Procedural justice: A psychological analysis*. Hillsdale, N.J.: Lawrence Erlbaum Associates

Tables

Table 1: Field survey: Statistical test of differences.

	Base (1)	Skill Gap (2)	No Diff. (3)	Disadv. (4)	1-2	1-3	1-4	2-3	2-4	3-4	Obs
Panel A: US sample											
Overall	2.87	3.58	3.31	2.08	0.00	0.00	0.00	0.00	0.00	0.00	1011
Male	2.97	3.50	3.36	2.21	0.00	0.000	0.00	0.06	0.00	0.00	504
Female	2.77	3.66	3.26	1.94	0.00	0.000	0.00	0.00	0.00	0.00	507
p-value (male vs female)	0.012	0.041	0.233	0.00							
Panel B: Australian student sample											
Overall	2.772	3.579	3.251	1.825	0.00	0.00	0.00	0.00	0.00	0.00	171
Male	3.00	3.631	3.425	2.069	0.00	0.000	0.00	0.16	0.00	0.00	87
Female	2.536	3.524	3.071	1.571	0.00	0.000	0.00	0.01	0.00	0.00	84
p-value (male vs female)	0.01	0.57	0.06	0.00							
p-value (AUS vs US) overall sample	0.364	0.988	0.590	0.006							
p-value (AUS vs US) male sample	0.869	0.390	0.663	0.296							
p-value (AUS vs US) female sample	0.116	0.334	0.216	0.002							

Notes: This table shows the average response to the field survey questions. Male and Female rows restrict the sample by gender. Survey responses range from 1= strongly agree to 5= strongly disagree. Columns 5 to 6 show the difference in mean responses using a t-test. p-values from panel B are taken from t-tests. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Effort level and costs

Effort	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Effort costs	0	1	2	4	6	8	10	12	15	20

Table 3: Predictions for the impact of the gender quota on gift exchange levels

				Marginal Product of Labor	
		Standard Reciprocity	Meritocracy	Male Manager	Female Manager
Experiment	Skill gap	→	↓	↑	↓
	No difference	→	↓	↑	↓
	Disadvantage	→	↑	↑	↓

Notes: The arrows indicate the change in both wage and effort levels (constant, decrease, or increase) from treatment *no quota* to *quota* in each experiment. The predictions do not depend on the gender of the manager for standard reciprocity and meritocracy.

Table 4a: Predicting effort in the skill-gap experiment

	(1) All workers' effort	(2) Male workers' effort	(3) Female workers' effort	(4) All workers' effort	(5) Male workers' effort	(6) Female workers' effort
Quota	–0.132***	–0.176***	–0.097*	–0.046*	–0.063*	–0.024
	(0.038)	(0.044)	(0.054)	(0.026)	(0.035)	(0.039)
Male worker	0.013			0.048		
	(0.048)			(0.030)		
Wage				0.006***	0.006***	0.006***
				(0.000)	(0.001)	(0.001)
Constant	0.609***	0.709***	0.513***	0.020	0.170	–0.044
	(0.065)	(0.092)	(0.092)	(0.075)	(0.125)	(0.058)
Observations	1,290	660	630	1,290	660	630
Number of subjects	86	44	42	86	44	42

Table 4b: Predicting wage in the skill-gap experiment

	(1) All managers' wage	(2) Male managers' wage	(3) Female managers' wage	(4) All managers' wage	(5) Male managers' wage	(6) Female managers' wage
Quota	–13.372***	–11.864***	–13.446***	–7.308***	–5.765**	–8.347***
	(2.892)	(4.505)	(3.721)	(1.844)	(2.881)	(2.631)
Male	1.304			3.644		
	(5.161)			(3.334)		
Effort (t-1)				53.313***	62.982***	44.341***
				(3.894)	(3.190)	(6.015)
Constant	82.221***	66.503***	86.037***	50.998***	38.378***	59.737***
	(8.129)	(14.079)	(9.332)	(5.685)	(9.745)	(6.044)
Observations	1,290	630	660	1,204	588	616
Number of subjects	86	42	44	86	42	44

Notes: Random-effects models. All regressions include controls. Controls include gender of the worker, if they study psychology, average belief about partners' gender arithmetic score, and belief about own arithmetic score. Period control accounts for time trends in this model. In both components of Table 4, columns 2 (3) and 5 (6) restrict the sample to male (female) subjects. Data are from the laboratory skill-gap experiment. Significance levels are *** p<0.01, ** p<0.05, * p<0.1.

Table 5a: Predicting effort in the no-difference experiment

	(1) All workers' effort	(2) Male workers' effort	(3) Female workers' effort	(4) All workers' effort	(5) Male workers' effort	(6) Female workers' effort
Quota	–0.120***	–0.135**	–0.117*	–0.050**	–0.065**	–0.047
	(0.046)	(0.062)	(0.068)	(0.023)	(0.030)	(0.039)
Male	0.057*			0.048**		
	(0.030)			(0.023)		
Wage				0.006***	0.007***	0.005***
				(0.001)	(0.001)	(0.001)
Constant	0.419***	0.447***	0.448***	0.042	–0.003	0.009
	(0.111)	(0.155)	(0.112)	(0.073)	(0.097)	(0.053)
Observations	1,140	600	540	1,140	600	540
Number of subjects	76	40	36	76	40	36

Table 5b: Predicting wage in the no-difference experiment

	(1) All managers' wage	(2) Male managers' wage	(3) Female managers' wage	(4) All managers' wage	(5) Male managers' wage	(6) Female managers' wage
Quota	–9.625**	–11.075**	–8.090*	–5.039**	–7.219***	–3.010
	(4.321)	(5.139)	(4.629)	(2.061)	(2.156)	(3.151)
Male	1.002			1.825		
	(2.601)			(1.951)		
Lag Effort				55.854***	58.906***	53.522***
				(6.476)	(9.848)	(5.062)
Constant	79.861***	87.287***	70.324***	49.611***	56.004***	41.854***
	(10.594)	(12.520)	(9.444)	(6.586)	(8.677)	(6.315)
Observations	1,140	540	600	1,064	504	560
Number of subject	76	36	40	76	36	40

Notes: Random-effects models. All regressions include controls. Controls include gender of the worker, if they study psychology, average belief about partners' gender arithmetic score, and belief about own arithmetic score. Period control accounts for time trends in this model. In both components of Table 5, columns 2 (3) and 5 (6) restrict the sample to male (female) subjects. Data are from the laboratory no-difference experiment. Significance levels are *** p<0.01, ** p<0.05, * p<0.1.

Table 6a: Predicting effort in the disadvantage experiment

	(1) All workers' effort	(2) Male workers' effort	(3) Female workers' effort	(4) All workers' effort	(5) Male workers' effort	(6) Female workers' effort
Quota	0.065*	–0.046	0.141***	0.014	–0.076**	0.075***
	(0.034)	(0.060)	(0.051)	(0.016)	(0.030)	(0.027)
Male	0.010			0.041		
	(0.047)			(0.029)		
Wage				0.006***	0.006***	0.005***
				(0.000)	(0.000)	(0.000)
Constant	0.559***	0.547***	0.545***	0.050	0.012	0.105
	(0.103)	(0.130)	(0.100)	(0.067)	(0.097)	(0.073)
Observations	1,440	675	765	1,440	675	765
Number of subjects	96	45	51	96	45	51

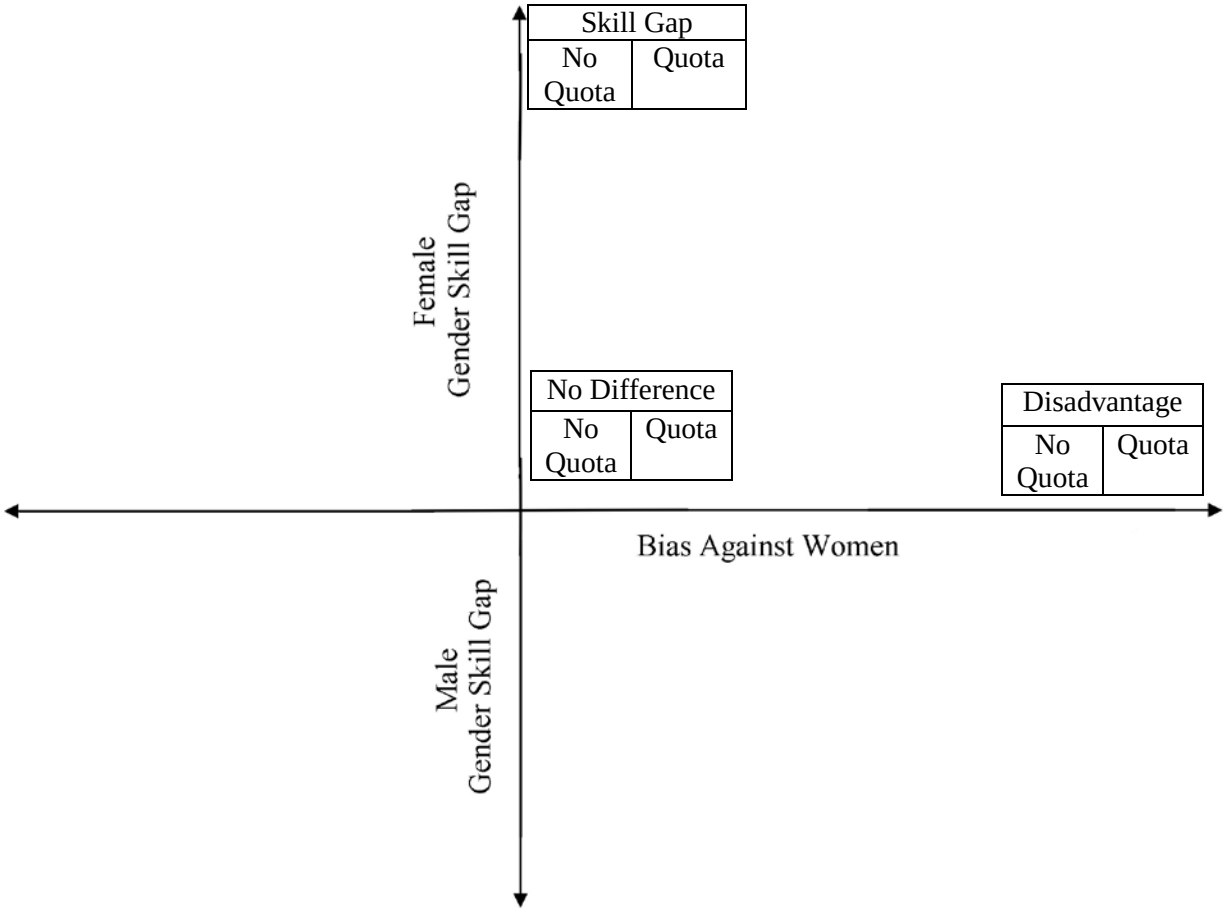
Table 6b: Predicting wage in the disadvantage experiment

	(1) All managers' wage	(2) Male managers' wage	(3) Female managers' wage	(4) All managers' wage	(5) Male managers' wage	(6) Female managers' wage
Quota	9.550***	15.134***	4.002	6.553***	8.032**	5.031
	(3.451)	(5.732)	(5.240)	(1.988)	(3.661)	(3.063)
Male	7.184			8.303**		
	(6.127)			(4.061)		
Lag effort				53.075***	54.016***	53.627***
				(4.103)	(6.518)	(3.462)
Constant	66.942***	64.688***	78.259***	43.349***	45.298***	48.842***
	(8.951)	(11.122)	(9.401)	(6.497)	(7.967)	(8.844)
Observations	1,440	765	675	1,344	714	630
Number of subjects	96	51	45	96	51	45

Notes: Random-effects models. All regressions include controls. Controls include gender of the worker, if they study psychology, average belief about partners' gender arithmetic score, and belief about own arithmetic score. Period control accounts for time trends in this model. In both parts of Table 6, columns 2 (3) and 5 (6) restrict the sample to male (female) subjects. Data are from the laboratory disadvantage experiment. Significance levels are *** p<0.01, ** p<0.05, * p<0.1.

Appendix

Figure A1: The three experiments and their two treatments



Notes: In each of the three experiments (*skill gap*, *no difference*, and *disadvantage*) there are two treatments (no quota and quota). Although this figure displays all four potential quadrants, for the purpose of this study, we focus specifically on the second quadrant, as it is most relevant for gender quotas.

Table A1: Opinion on Occupational Skill Gap and Disadvantage

Occupation	% of Sample (1)	Median Skill Gap (2)	Median Disadv. (3)
Farming, Fishing, and Forestry	1,03%	1	3
Management Occupations	5,51%	1	2
Business Operations Occupations	3,44%	0.5	2
Construction and Extraction Occupations	5,68%	2	3
Engineering Occupations	3,49%	2	1
Legal Occupations	1,03%	1.5	1
Primary/Secondary (K-12) Education Occupations	6,20%	0	1
Community and Social Services Occupations	2,07%	0	1
Healthcare Support Occupations	6,88%	0	1
Physician/Doctor Occupations	0,69%	1.5	2.5
Office and Administrative Support Occupations	8,95%	1	1
Finance and Insurance Occupations	2,58%	1	2.5
Arts, Design, Entertainment, and Sports,	4,30%	1	2
Sales Occupations	9,12%	1	2
Manufacturing and Production Occupations	6,20%	2	3
College or University Academic	2,75%	1	0
Food Preparation- and Serving-Related Occupations	3,79%	1	2
Building and Grounds Cleaning and Maintenance	1,20%	2	3
Personal Care and Service Occupations	3,10%	2	2
Computer and Mathematical Occupations	8,26%	1	2
Transportation and Material Moving Occupations	3,96%	1	3
Installation, Maintenance, and Repair	2,24%	0	2
Military and Protection-Specific Occupations	2,2%	3	1

Notes: Column 1 reports the percentage of the sample in the labor force within each occupation. Columns 2 and 3 report median responses from questions related to perception of skill gap and disadvantage within an individual's occupation group.

Table A2: Opinions on the appropriateness of gender quotas by occupation and environment

	Teaching, Health, and Social Services (1)	Engineering (2)	Management, Math, and Computers (3)	Transportation (4)	Construction (5)	Military and Protection (6)	Office and Admin (7)	K–W Test Chi Square (8)
General	2.803	2.750	2.772	2.565	3.027	3.143	2.500	10.80*
Skill Gap	3.672	3.167	3.435	3.565	3.514	4.00	3.615	9.71*
No Skill Gap	3.224	2.583	3.413	3.261	3.216	3.143	3.308	10.76*
Disadvantage	1.794	1.667	2.130	2.061	2.432	2.220	1.910	12.42**
% of Sample	18.41%	3.49%	15.83%	3.96%	6.37%	2.2%	8.95%	

Notes: Higher numbers indicate strongly disagree**.** Column 8 reports the Kruskal–Wallis test of overall significance. % of sample refers to % of sample in the labor force.