Universitat Autònoma de Barcelona

Department of Business Economics

Doctorate in Economics, Management, and Organization

DOCTORAL DISSERTATION

The role of gender in accountability and competition

Author:
Orsola Garofalo

Dissertation Advisor:
Jordi Brandts Bernad, Ph.D.

Barcelona, May 2014

Universitat Autònoma de Barcelona

Department of Business Economics

Doctorate in Economics, Management, and Organization

DOCTORAL DISSERTATION

The role of gender in accountability and competition

Author:
Orsola Garofalo

Dissertation Advisor:
Jordi Brandts Bernad, Ph.D.

Barcelona, May 2014

ABSTRACT

A growing literature explores gender differences in economics behavior and their implications for organizations and labor markets. My dissertation contributes to this literature by investigating on the role of gender as well as gender interactions using lab experiments and empirical data.

In the first chapter, we conduct an experiment to investigate how the gender of a decision-maker interacts with the gender composition of an audience that calls for an explanation of the decision-maker's choices. In particular, we look at how this gender interaction shapes the decision-maker's level of cognitive effort in an accountable setting. Together with the measure of accountability based on decision theory, we employ two physiological measures, blood pressure and heart rate variability, which allow us to disentangle the effects of stress and accountability. Our results show that men are more sensitive to changes in the gender composition of the audience. By contrast, women provide a level of cognitive effort that does not change depending on the gender of the audience. Finally, we find that the variation in blood pressure has a significant but small effect only on men's behavior.

In the second chapter, we extend the previous experiment introducing two new elements, namely that participants make incentivized decisions and that share the payoffs with the audience members. Thus, we study how a decision-maker's cognitive effort is influenced by (1) the interaction of his own gender with the one of an audience; (2) the compensation scheme adopted. The main results show that the combination of incentives and payoff sharing has a positive effect on female's cognitive effort provision; in particular, women increase their effort level when paired with an audience of the opposite gender. By contrast, male's cognitive effort is influenced when the compensation scheme includes only incentives; men decrease their effort provision when paired with an audience of the opposite gender.

In the third chapter, using a data set of family-controlled firms in Italy we analyze whether gender interactions at the top of the corporate hierarchy affect corporate performance. Our findings show that female directors significantly improve the operating profitability of female-led companies. To mitigate endogeneity concerns, we assess executive transitions using a triple-difference approach complemented by propensity score matching and instrumental variables. Finally, we show that the positive effect of female interactions on profitability is reduced when the firm is located in geographic areas characterized by gender prejudices and when the firm is large.

In the last chapter, we test whether female and male executives react differently to an increase in competitive pressure. Our identification is based on the staggered change of barriers interstate branching, which varied the exposure to competitive pressure in the US banking sector during the mid 1990s-early 2000s. Our results indicate that when competition increases, female-led banks experience worse accounting and market-based performance. However, we find that despite lower profitability, female leadership mitigates the potential exacerbating effect of competition on risk-taking.

ACKNOWLEDGEMENTS

Do you think your name should be in this page?

If you think that during my stays in Barcelona and San Diego you ever supported (and tolerated) me somehow, or gave me good and terrible advices, or spent with me unforgettable moments, or laughed without reason, or spent hours and hours working together, or if we had a siesta in the classroom, or we shared food and fruit during summer holidays...

Then, here it is! Thanks a lot ☺

A special thanks to Jordi and my family!

TABLE OF CONTENTS

Chapter IV: Gender interactions within the family firm

Chapter V: How do female bankers react to competitive pressures?

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Motivated by the introduction of policies that aim to increase the presence of women in political and corporate positions, a growing literature explores gender differences in economic behavior and their implications for organization. My dissertation contributes to this literature by investigating on the role of gender as well as gender interactions in the lab and in the field.

In the first two chapters we use a laboratory experiments to study whether the gender composition of an audience interacts with the gender of a decision-maker thereby shaping effort provision in an accountable setting. In other words, we do not analyze single gender effects in accountability but rather we propose to investigate the effect induced by gender pairings between audience members and decision-makers. We focus on gender pairings because they are crucial to provide a comprehensive description of male and female behavior, as indicated by research in evolutionary psychology (Buss 1999) and economics (Sutter, Bosman, Kocher and van Winden 2009; Gneezy and Rustichini 2004). An important question is whether, in an accountable setting where individuals have to justify their choices to an audience, the provision of effort by a player depends on the demographic characteristics of the player herself and the audience members. Gender is, in our opinion, one of the most potentially interesting variables. For example, in relation to the presence of women both in CEO positions and on the board of directors of corporations, it is crucial to understand whether e.g., a monitoring committee spurs accountability differently depending on the genders of participants and committee members.

In particular, in Chapter II we look at the interaction between the gender of the decision-maker and the gender of the audience in a non-incentivized setting. Our results show that men's choices entail a different level of accountability, measured by the correct answers in the decision task, depending on the gender of the audience; men make more responsible decisions when paired with the same sex. By contrast, women behave equally, regardless of the audience's gender. In Chapter III, we create again an accountable setting and study how effort is influenced by both gender pairings between the decision-maker and the audience, and the compensation scheme. In this chapter we introduce a new element, the decisions are incentivized and the final payoffs are shared with the audience. The result of this experiment shows that women’s effort is influenced by the audience’s gender when their compensation

scheme includes both incentives and payoff sharing; and that men in this new context do not behave differently depending in the audience's gender.

In chapter IV, using a comprehensive data set of family-controlled firms in Italy, we analyze whether gender interactions in the top layers of a firm's hierarchy affect corporate profitability. The findings show that an increasing fraction of women on the board of directors significantly improves the profitability of firms led by female CEOs. Our results have several implications. Existing evidence suggests that women are more sensitive than men to context (Croson and Gneezy 2009). By showing that the interactions with a board in which women are largely represented mitigate the observed underperformance of lone female CEOs, our results augment recent discussions (e.g., Dezsö and Ross 2012) of the contexts under which female leadership can be more effective. Our findings also suggest that the criteria by which women gain seats on the board may shape the performance effect. That is, the female interactions that are most beneficial to firm profitability are those involving nonfamily female directors, who are more likely to have been selected via a meritocratic process. This notion is reinforced by our finding that female interactions are less effective in areas with strong traditional family values, where companies more likely appoint executives on the basis of family ties. Our results thus complement those of Ahern and Dittmar (2012), who document a decline in the market value of firms subsequent to the passage in Norway of gender quotas for boards (a finding that may reflect the tendency of those companies to appoint relatively inexperienced female directors).

In Chapter V, we focus on legislative changes in the US banking industry to test whether women and men executives respond differently to an exogenous increase in competitive pressures. The US banking sector in the late 20th century represents an ideal *laboratory* for our research. While historical regulations severely limited the geographic expansion of banks, US states gradually lifted these restrictions starting in the 1970s. Specifically, our identification is based on the legal roadblocks US states passed to limit the scope of deregulation of interstate branching activities introduced in 1994 with the Interstate Banking and Branching Efficiency Act (IBBEA). The staggered introduction and removal of these restrictions determined both temporal and geographic variations in the intensity of competition from the mid 1990s to the early 2000s, variations that are useful for mitigating endogeneity concerns (Cornaggia, Mao, Tian and Wolfe 2013; Rice and Strahan 2010; Johnson and Rice 2008). Our results show that when competitive pressure increases, female-led banks experience lower performance than their male counterparts. Focusing on risk-taking, we find that banks led by women were less subject to an increase in risk after

competition increased. Thus, our findings indicate that despite lower profitability, female leadership generates superior bank capitalization and more stable streams of profit.

Bibliography

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

Buss D.M. (1999). *Evolutionary psychology: The new science of the mind*: Allyn & Bacon.

Cornaggia J., Mao Y., Tian X. and Wolfe B. (2013) "Does banking competition affect innovation", *Journal of financial economics*.

Croson R. and Gneezy U. (2009) "Gender differences in preferences", *Journal of Economic Literature*, 448-474.

Dezsö C.L. and Ross D.G. (2012) "Does female representation in top management improve firm performance? A panel data investigation", *Strategic Management Journal* 33, 1072-1089.

Gneezy U. and Rustichini A. (2004) "Gender and competition at a young age", *American Economic Review*, 377-381.

Johnson C.A. and Rice T. (2008) "Assessing a decade of interstate bank branching", *Wash. & Lee L. Rev.* 65, 73.

Rice T. and Strahan P.E. (2010) "Does Credit Competition Affect Small- Firm Finance?", *The Journal of Finance* 65, 861-889.

Sutter M., Bosman R., Kocher M.G. and van Winden F. (2009) "Gender pairing and bargaining—Beware the same sex!", *Experimental Economics* 12, 318-331.

CHAPTER II

GENDER PAIRINGS AND ACCOUNTABILITY EFFECTS

1. Introduction

Individuals living in social groups do not make decisions in isolation; rather, they are often called on to defend their conclusions and their reasoning to peers, subordinates or superiors (Lerner and Tetlock 1999). Following this notion, accountability can be thought as the social pressure to justify one's views to others.

The role of accountability[1] in decision-making has been extensively investigated by psychology scholars. According to this strand of literature, accountable individuals exhibit greater coherence between gain and loss frames (Takemura 1994, 1993; Miller and Fagley 1991) and lower overconfidence (Arkes, Christensen, Lai and Blumer 1987). By contrast, economics scholars have only recently turned their attention to the concept of accountability, despite the numerous real-world applications to politics or corporate governance. Recent studies argue that when a decision-maker expects to justify ex post his choices in front of an audience, the likelihood that he will make the right decisions is significantly higher (Lerner and Tetlock 1999). Exploring this effect in an experimental setting, Vieider (2011) shows that indeed the presence of an audience that calls for the explanation of the choices made increases the individual effort in solving a decision task and thus the number of superior events chosen.

This paper offers a novel examination of the audience-based accountability by testing if the gender composition of the audience interacts with the gender of a decision-maker thereby shaping her/his degree of responsibility. In other words, we do not analyze single gender effects in accountability but rather we propose to investigate the effect induced by gender pairings between audience members and decision-makers.

We focus on gender pairings because they are crucial to provide a comprehensive description of male and female behavior, as indicated by the extant evidence in evolutionary psychology (Buss 1999) and recent economic works[2]. Failing to account for gender pairing effects may lead to severely biased estimates of gender differences in several outcomes.

Quantifying accountability has proven to be challenging. In the spirit of Lerner and Tetlock (1999), Vieider (2011) adopts measures based on decision theory, such as the time

[1] "Someone who is accountable is completely responsible for what they do and must be able to give a satisfactory reason for it" (Cambridge Advanced Learner's Dictionary2008)

[2] For example Gneezy and Rustichini (2004) find that in a single (non-competitive) race there is no difference between boys and girls at young age. However, a significant difference shows up when the children are paired with other children of the same or opposite sex: whereas boys competing against boys perform better than in the non-competitive setting, girls paired with girls perform worse than when they run alone. In the mixed race, the performance improvement of boys is much larger than the one of girls. Another paper that has explored gender pairings is Sutter, Bosman, Kocher and van Winden (2009).

employed in solving a task and the number of correct answers. We claim that introducing an audience might induce not only an effect on accountability (i.e. by anticipating that a player has to explain her/his choices, she/he should make more responsible choices), but also an effect on psychological stress. In fact, explaining or making a presentation has been considered as a behavioral stressor in the medical literature (Girdler, Turner, Sherwood and Light 1990). This confounding factor turns out to be particularly relevant in our context because women and men tend to react differently to behavioral stressors (Allen, Stoney, Owens and Matthews 1993; Girdler et al. 1990). In order to mitigate this concern, we adopt two physiological measures throughout the experiment: heart rate variability and blood pressure[3]. By combining scores in decision tasks (Vieider 2011) and physiological measures, we are able to examine a player's accountability behavior from an economic viewpoint (i.e. effort and choices of the superior simple event) while monitoring her/his physiological activities during the resolution of the tasks.

Our results show that men's choices entail a different level of accountability, measured by the correct answers in the decision task, depending on the gender of the audience; men make more responsible decisions when paired with the same sex. By contrast, women behave equally, regardless of the audience's gender. These results are consistent with the patterns highlighted by the physiological measures: we record a change in the blood pressure only among male subjects that are paired with a female audience; yet among female subjects we do not detect any significant pattern in neither blood pressure nor heart rate.

Our paper contributes to different strands of research. First, we provide new insights into the economic literature about gender differences. Existing works have focused on a wide range of economic and social behaviors[4] but none thus far has analyzed the effect of gender pairings on the responsibility in decision-making. Second, our experimental design allows us to disentangle the accountability effect induced by the audience from the stress that the

[3] As discussed by prior medical research, such indicators represent reliable measures of an individual's cardiovascular response to stress. Matthews and Stoney (1988) adopt similar measures to study gender differences in response to behavioral stressors. In their framework, the participants are asked to do serial subtractions and mirror image tracing, while the experimenter measures the heart activity and the blood pressure. Results show that gender is a significant determinant of the cardiovascular adjustments to stressors. In line with this result, Lawler, Wilcox and Anderson (1995) find that in three laboratory tasks (mental arithmetic, video-game and anger recall interview) men have higher systolic blood pressure at all intervals and higher cardiac output during math and video-game tasks.

[4] Gender differences have been investigated with respect to, among other issues, competition (Niederle and Vesterlund 2007; Larson 2005; Gneezy and Rustichini 2004; Gneezy, Niederle and Rustichini 2003), cooperation (Charness and Rustichini 2011; Simpson 2003), coordination (Croson, Marks and Snyder 2008; Cadsby, Hamaguchi, Kawagoe, Maynes and Song 2007), risk aversion (Fehr-Duda, De Gennaro and Schubert 2006; Eckel and Grossman 2002), altruism (Andreoni and Vesterlund 2001; Eckel and Grossman 2001) trust (Schwieren and Sutter 2008; Bohnet, Herrmann and Zeckhauser 2006), and investment behavior (Atkinson, Baird and Frye 2003; Dwyer, Gilkeson and List 2002). With respect to some more specific issues, Adams and Ferreira (2009) suggest that women on the board of directors tend to have a better attitude in monitoring, Matsa and Miller (2013) that they are more stakeholders oriented, and Rivas (2008) that they are less corrupt than men. See also Croson and Gneezy (2009) for a review of literature on gender differences.

participant is subject to during the task resolution, thereby adding to the research on how to measure accountability.

Although our experiment does not have a specific characterization, it is far from being a merely abstract scenario; by contrast, it displays numerous potential real-world applications. For example, our design can resemble a corporate situation in which a CEO has to make decisions and justify them in front of the board of directors. Managerial behavior may be influenced by the gender composition of the board, which leads a CEO to make decisions entailing a different degree of accountability[5]. Our setting can also be related to a political context in which a mayor is responsible for governing a municipality while the counselors conduct monitoring activities. Gagliarducci and Paserman (2012) document that in municipalities headed by female mayors the probability of an early termination of the legislature is higher. In line with the notion that gender pairings matter, they also find that "the likelihood that a female mayor survives until the end of her term is lowest when the council is entirely male" (Gagliarducci and Paserman 2012).

In Paragraph 2, we propose a background and advance possible relationships between gender pairings and accountability. In Paragraph 3, we illustrate the experimental design. In Paragraph 4, we introduce the physiological measures adopted throughout the experiment. In Paragraph 5, we illustrate and interpret our empirical findings. In Paragraph 6, we conclude.

2. Background

Following Scott and Lyman (1968), Semin and Manstead (1983), and Tetlock (1992) accountability can be considered as a multiple phenomenon grounded on the implicit or explicit expectation that one may be called on to justify one's beliefs, feelings and actions to others. Several context-based types of accountability can be discerned: (a) mere presence of another (individuals expect that another will observe their performance); (b) identifiability (individuals expect that what they say or do in a study will be linked to them personally); (c) evaluation (individuals expect that their performance will be assessed by others according to some normative ground rules and with some implied consequences); (d) reason-giving (individuals except that they must give reasons for what they say or do)

Although the results may vary depending on the dimension considered, accountability-enhancing contexts have been repeatedly found to influence an individual 's engagement in high-effort elaborations and coherence among choices (Takemura 1994, 1993; Miller and

[5] In this example we leave out the fact that in a natural settings the CEO can be punished or rewarded by the board of directors.

Fagley 1991). Janis and Mann (1977) argue that people who expect to justify their views will be more vigilant information processors – more likely to perform the difficult cognitive tasks widely regarded as sign of high quality decision-making. Examining the roots of such an impact, Tetlock (1983) and Tetlock and Boettger (1989) present evidence that accountability is ultimately shaped by what people think, i.e. preferences, cultural values and beliefs.

To motivate the importance of looking at gender pairings in accountability, we posit that decision-maker's attitudes and beliefs toward the opposite gender might either foster or diminish her/his accountability. Social values and gender prejudices are receiving a great deal of attention by scholars because, in addition to biological considerations, they contribute to explaining variations in gender outcomes. An example is Gneezy, Leonard and List (2009), which finds that in a patriarchal society men are more competitive than women, while in a matrilineal society the opposite is true. Thus Gneezy et al. (2009)'s results seem to indicate that the way in which people look at the opposite gender shapes distinct individual behaviors.

We draw on these arguments and extend them to our context: since accountability might embody values, prejudices and beliefs toward the opposite gender, gender pairings can determine important variations in accountability.

To test whether gender pairings influence accountability, we conduct an experiment adopting the audience context such as in Vieider (2011) based upon the conceptual framework in Lerner and Tetlock (1999). However, measuring gender differences in accountability using an audience while neglecting biological factors might be misleading. In fact, the extant evidence shows that women express a wide variety of emotions more intensively than do men. Certainly, biological factors, including genetic, hormonal, and neuropsychological variables, as well as social factors, including different gender roles, and status differences between the two genders, contribute to determine gender differences in emotions (Ablon, Brown, Khantzian and Mack 2013). Emotions can be expressed in four basic ways: verbally, behaviorally, through non-verbal facial expressions, and through physiological arousal (increased heart rate, galvanic skin response, respiration, temperature) (Brody and Hall 2000; Adelmann and Zajonc 1989). Buck, Miller and Caul (1974) suggest that men tend to internalize emotions; they manifest emotion in their levels of physiological arousal and not in their facial expressions. In contrast, women tend to externalize emotions; they manifest emotion in facial expressions and not in levels of physiological arousal. The medical literature on gender and cardiovascular activities reports that men exhibit larger blood pressure (in particular, systolic values) changes than women in response to psychological challenges such as mental arithmetic, mirror image tracing, speech-making (Stoney,

Matthews, McDonald and Johnson 1988). By contrast, women respond to stress by increasing heart rate activity (McAdoo, Weinberger, Miller, Fineberg and Grim 1990). To detect potential gender differences of this sort during the resolution of the task, we track blood pressure and heart rate for each participant involved in the experiment.

3. Experimental design

Participants are recruited either using flyers on campus or by email, from a database (ORSEE) of students at Universitat Autònoma de Barcelona who voluntarily registered in the database to participate in prior experiments. The total number of participants is 83, of whom 44 women (16 of them participate at the experiment with a female audience, 17 with a male audience, and 11 without the audience) and 39 men (15 of them with a female audience, 13 with a male audience, and 11 without the audience). Each subject can participate only in one session and she/he is randomly allocated to a male or female audience. As reported in Table 2.1, Panel A, the average age is around 24, without significant differences between women and men. The audience is composed by three persons (either three women or three men) recruited from PhD students at Universitat Autònoma de Barcelona. In the instructions, we specify that the audience is composed by three experts in that specific kind of task, so that each participant is aware of it before starting the resolution of the task. The average age of the audience is about 33 and it is statistically larger than the participants' average age[6]. The differences in participants' age by audience's gender (untabulated) are small and statistically not significant.

To mitigate the impact of confounding factors when using physiological measures, we collected, through a questionnaire, information on health characteristics that are typically associated with blood pressure or heart rate activities, such as weight, height, problems of high blood pressure, smoking status and so on. Table 2.1, Panel B, reports summary statistics. The average weight for male subjects is 73 Kg and the average height is 176 cm, while for females it is 60 Kg and 164 cm, respectively. Only one participant has hypertension problems and only 20% of them smoke (Table 2.1, Panel B).

On average, each session takes about 45 minutes. It involves only one participant and three members of the audience; each participant meets the audience members for the first time during the experiment. All subjects (both players and audience members) are paid a flat fee of

[6] For the composition of the audience we chose people older than the participants in order to make the audience credibly expert from the student's point of view. The audience was composed by Ph.D. students but they were not really experts in this kind of problems.

9€ for their participation. We pay a flat fee to avoid that the participant's effort in the resolution of the task is affected by monetary incentives and not only by the gender of the audience.

The experiment involves the analysis of a treatment represented by the gender of the audience paired with the opposite gender of the participant. Overall, we have four gender pairs: Female Participant-Male Audience; Female Participant-Female Audience; Male Participant-Female Audience; Male Participant-Male Audience.

In order to measure the individual ability of the decision-makers, with no audience involved, we also run some control sessions. The experimental setting of the control sessions is the same of the one in the treatment sessions, the only difference is that the participant does not justify her/his choices in front of an audience, she/he only has to solve the task. In the treatment and control groups the personal characteristics of the participants are similar.

Each session takes place in a semi-empty room containing a big table in which the participant sits on one side and the audience on the opposite side. The session is divided into several parts. In the first part, the experimenter explains a part of the rules and makes the participant ready for the blood pressure and heart rate measures. In particular, the first part of the instructions only specifies that the experiment consists in measuring the blood pressure and the heart rate, and in solving a decision task. The experimenter helps the participant in putting on the chest belt to measure the heartbeat throughout the experiment and then, in order to take time for the stabilization of the blood pressure, she gives her/him a document to fill in with personal information and a questionnaire[7] (Multidimensional Mood State Questionnaire Steyer, Schwenkmezger, Notz and Eid 1997) containing questions about personal emotional status at that moment.[8] Afterwards, the experimenter measures the blood pressure of the subject. Up to this point, the subject does not know the main structure of the experiment, thus we can use the first measure of blood pressure as a baseline.

In the second part, the experimenter hands to the participant the second part of the instructions and then reads them aloud to be sure that the participant understands the structure of the experiment. This part of the experiment consists in solving the task, and in justifying the options chosen in front of an audience. Before leaving the participant alone in the room for the resolution of the task, the experimenter tells the participant that the audience will enter the room to fill in a document, to deliver the task sheet to her/him and then to briefly explain

[7] We do not report the results obtained from the questionnaire because we are interested only in making time in order to stabilize the blood pressure.

[8] It is a list of expressions that characterize different moods. For each word, the participant has to mark a number (ranged from 1 to 5) that best represents the actual intensity of her/his mood status.

how the resolution of the task works.

The audience enters the room, sits down and fills in the document for a few seconds, whereas the experimenter goes out of the room. While the audience fills in the document, the participant can observe the audience members for the first time, and thereby she/he creates her/his own idea about them. Afterwards, one person of the audience explains to the participant the rules to solve the task and gives her/him the sheet containing the exercises. Then the audience leaves the room and the participant remains alone while solving the task. As soon as the participant has solved all six problems, she/he has to ring a small bell and then the experimenter enters in the room and takes the second blood pressure measure. Afterwards, the audience enters the room for the second time and the last part of the experiment starts: the participant justifies the decisions made.[9] At the end of the explanation, the audience leaves the room and the experimenter enters the room to take the third blood pressure measure. To conclude the experiment, the participant fills in a slightly different version of the same questionnaire she/he filled in at the beginning of the experiment and signs a receipt for the payment.

In the experiment the audience plays a neutral role, it has only to pay attention to the subject during the explanation. We asked the members of the audience to be serious during the first meeting and the explanation, but not to directly stress or pressure the participant. The main role of the audience is to increase the player's commitment in the resolution of the task.

The task consists of six problems in which the participant has to choose between simple and compound prospects (as in Bar-Hillel 1973). Bar-Hillel (1973) presents a case in which subjects were given the opportunity to choose between one of two events. Three types of events were used: (1) simple events, such as drawing a red ball from an urn containing 50 red balls and 50 black balls; (2) conjunctive events, such as drawing a red ball seven times in succession, with replacement, from a bag containing 90 red balls and 10 black balls; and (3) disjunctive events, such as drawing a red ball at least once in seven successive tries, with replacement, from a bag containing 10 red balls and 90 black balls. A significant majority of subjects chose the conjunctive event (probability of 0.48) rather than the simple event (probability of 0.50); they also preferred the simple event rather than the disjunctive event (probability of 0.52). These biases can be explained as effects of an anchoring and adjustment process (Holtgraves and Skeel 1992; Kruglanski and Freund 1983; Tversky and Kahneman 1974). The stated probability of the elementary event (successful at any stage) provides a

[9] We did not record the explanations because we are not interested in the reason given to justify the choices, yet we focus the attention only on the effect that the presence of the audience has on decisions made before the explanation.

natural starting point for the estimation of the probabilities of both conjunctive and disjunctive events. Since the adjustment from the starting point is typically insufficient, the final estimates remain too close to the probability of the elementary events in both cases (Tversky and Kahneman 1981).[10]

The problems are all of the same type but with different level of analytic difficulty. Lower number of extractions in the compound event makes the calculation easier and should thus increase the choice of the superior event. In the problems chosen, the compound probability to draw the winning balls is always lower than the simple prospect probability.

4. Data and measures

In this section, we describe in detail the physiological measures adopted in the experiment. The heart rate variability was measured non-invasively using the Polar Heart Rate Monitor RS400, consisting in a belt and a wear-link transmitter that the participant wears on the chest from the beginning until the end of the experiment. The heartbeats are recorded in a watch and then downloaded on a computer using an infrared USB.

The software records the heart rate expressed in number of heartbeats per minute and it provides several indicators that describe the heart rate variability (HRV), such as standard deviation, mean, maximum, RMSSD[11], pNN50[12], Low Frequency (LF), High Frequency (HF) and LF/HF ratio expressed in milliseconds. The HRV represents the variability of the cardiac frequency as response to respiratory rhythm, emotional stress, anger and relaxation. In a healthy heart, the cardiac frequency reacts quickly to these factors in order to better adapt the body to the external environment. The HRV is correlated with the interaction between the Sympathetic and Parasympathetic Nervous System[13]. Moreover, the beginning and the end of each session are marked in the sequential measure of the heart activity to isolate it during specific moments of the task resolution.

Figure 1 contains an example of heart activity together with three markers. The first marker is taken at the eighth minute of the experiment that corresponds with the entrance of the audience in the room for the first time; the second one is at the nineteenth minute of the

[10] Given these findings, making the simple lottery the correct choice allows for a better detection of the effort exerted in making the decision.
[11] Square root of the mean squared difference of successive normal-to-normal intervals (NNs).
[12] Proportion of NN50 divided by total number of NNs.
[13] The Sympathetic Nervous System is responsible for acceleration of the heart rate, bronchial dilatation, increase of the blood pressure, pupillary dilatation, etc.. By contrast, the Parasympathetic Nervous System produces a decrease of the cardiac rhythm, increase of the bronchial muscular tone, dilatation of blood vessels, decrease of the blood pressure, slowdown of the breath, increase of the muscular relaxation, etc… The ability of our body to tend to one or the other nervous system represents the possession of a dynamic equilibrium fro a physiological and psychological point of view.

experiment and it coincides with the end of the task and the beginning of the explanation; the third one is just after the explanation that was at the twenty-fourth minute.

To further quantify the level of stress during the task resolution, we also measure the blood pressure, measured non-invasively using an automated blood pressure meter. Such an instrument provides the two main indicators of blood pressure: systolic and diastolic levels, respectively the maximum and minimum values of blood pressure during each heartbeat.

5. Experimental results

Following Vieider (2011), we assume that being held accountable by an audience leads to a more thorough assessment of probabilities and thereby to more frequent choices of the superior simple event by focusing the attention on the probabilities involved. In this section, we test how the accountability measured in such a way is influenced by gender pairings between decision-makers and audience members.

The choices made by the participant are encoded as dummy variables, with 1 indicating the choice of the normatively superior event (simple event), and with 0 the compound event. These dummies are summed for all six choice pairs to obtain a general index ranging from 0 to 6. The number of correct answers is not influenced by the type of education of the participants (e.g. maximum and the minimum indexes were obtained by people from law, mathematics, sociology, or economics degrees). These results also indicate that the education level of the sample was adequate to solve the task and that the choices were influenced by some other factors.

5.1. Descriptive statistics

Figure 2.2 shows the average index values in the unpaired cases, first by male and female audience, and second by male and female participants. Panel A shows that the mean number of superior simple event choices in the whole sample is 2.7 with male audience and 1.9 with female audience. Panel B reports that, independently of the audience's gender, men choose on average 2.8 simple event choices, and women about 1.9 out of six. In all cases, we do not find any statistically significant difference.

Figure 2.3 shows the average index values once we take into account gender pairings. The left part of the graph reports the values referred to male participants when they justify their choices in front of male or female audience. Male participants in the male audience treatment choose, on average, 3.8 superior choices out of 6; yet, with a female audience the number of

superior choices goes down drastically to 1.7 out of 6. The difference between the two averages is statistically significant at the conventional level (p-value=0.0064). Focusing the attention on female participants, we do not find any significant difference between the two audience treatments (p-value=0.682). This result suggests that women put the same level of effort in solving the task regardless of the gender of the audience.

Although we find a significant difference in the number of correct answers chosen among male participants, we do not find any difference in the time employed to solve the task. The same result is valid also for the female sample.

In the control sessions, the ones without using the audience, we find that male participants choose on average 3.1 correct answers out of 6, whereas female participants 1.5 out of 6. These results clearly suggest that male subjects in the treatment sessions over-perform when are paired with a male audience and underperform when paired with a female audience. On the other hand, females underperform in the control sessions with respect to both the treatment with male and female audience.

We turn now our attention to the physiological measures. Figure 2.4 presents the average values of the systolic blood pressure taken in three different moments during the experiment: the first is measured at the beginning of the experiment and it is considered as baseline value; the second is referred to the moment just after the resolution of the task and before the explanation; the third is taken after the explanation. In Panels A and B are reported, respectively, the average values for male and female subjects paired with male or female audience. Among male subjects we find a significant difference on the second systolic blood pressure across variations of the audience's gender: the average value with a male audience is 122.08 mmHg but it increases to 130.4 mmHg with female a audience (p-value=0.09)[14]. Controlling for physical characteristics, such as weight and height, this difference becomes slightly larger and remains statistically significant at 10% level.

Since the second measure is taken just few seconds before the explanation, it should indicate if the participant feels stressed or not. Male subjects feel more stressed when they are paired with an audience of the opposite gender. This result seems to be in line with van der Meij, Buunk, van de Sande and Salvador (2008)'s study that shows that a short presence of a woman induces an hormonal reaction, such as in the level of salivary testosterone (T), in young men. Experiments conducted with animals show that an increase in level of T corresponds to an increase in the blood pressure (Fischer and Swain, 1977). On the other

[14] In the blood pressure graphs, we only report the systolic values because, as reported in the medical literature (see e.g. Musini and Wright 2009), the diastolic values typically display little variations.

hand, the effect of a change in T on the arterial pressure in individuals has not been well-detected, and existing medical research only shows that men with low levels of T tend to have high blood pressure (Svartberg et al. 2004).

As in the case of task performance, we do not find that the gender of the audience matters for female blood pressure. To identify how the gender of the audience influences the reaction of male and female participants to stress (i.e. providing an explanation in our context), we compare the LF/HF Ratio around the three above-mentioned markers (see Figure 2.1). In unreported analysis, we do not find any significant impact of gender pairings on the heart rate activity; neither males nor females vary significantly the LF/HF Ratio as the gender of the audience changes. This lack of significance is not surprising in light of the results obtained in the decision tasks. As reported in Section 2, the medical literature indicates that male responsiveness to a stressful situation mostly shows up in the blood pressure value and not in the heart rate. Because we have demonstrated that gender pairings matter for male only, once we examine the physiological measures we find a significant effect of gender pairings for men only on the blood pressure. Also consistent with the finding that women behave in accountable manner regardless of the gender of the audience, in the female sample physiological measures of stress do not vary depending on the gender pairing with the audience.

5.2. Regression analysis

Table 2.2, Panel A, shows the results of OLS regressions, separately for men (Columns 1 and 2) and women (Columns 3 and 4), in which the dependent variable is the sum of the superior simple event choices. The explanatory variables are the dummy treatment, equal to one when audience and participant's gender is the opposite, the player's age and her/his field of study (three categories are considered: social science, science and humanities; humanities is used as reference group). Our results confirm that being paired with the opposite gender has a negative and statistically significant effect on male's accountability. By contrast, the effect is quantitatively small and statistically insignificant for women.

We test the robustness of the findings above in several ways. Given the structure of our dependent variable, we estimate an ordered probit model (See Table 2.2, Panel B) rather than an OLS. Alternatively, we adopt a Poisson regression model. As a next step, in the OLS specification we control for the difference between audience and participant's age (coefficient=-2.510; p-value=0.01). We explore the validity of our results within a smaller

range of the dependent variable, excluding participants that choose all right or all wrong answers. Finally, we propose two alternative ways to compute standard errors; we cluster standard errors by the hour of the experiment, and we compute bootstrapped standard errors (using 500 replications) to deal with the limited sample size. Our results (unreported to save space) are in all cases qualitatively unchanged. Men's level of accountability is strongly affected by the gender of the audience; on the other hand the level of accountability in the decision process of the female participants is not affected by the gender of the audience.

As above-mentioned, a stressful situation can influence performance, though the direction of this influence is ambiguous. As argued in Kavanagh (2005), there can either be a negative relationship between stress and performance, or a positive relationship, at a moderate level of stress. Thus, not controlling for the stress in the form of physiological reaction may lead to an overestimation or underestimation of the impact of the audience's gender on the participant's decisions. To mitigate this concern, we show in Table 2.3[15] the results of OLS regressions including physiological measures; blood pressure (the difference between the first and the second systolic value) and heart rate variability (ratio between low and high frequency of the heartbeat[16]). In unreported regressions, we also controlled for smoking behavior, consumption of alcohol and sport activities before the experiment, and hypertension problems. Results are unchanged by the inclusion of these controls.

We have demonstrated that gender pairings matter for men only and once we control for the physiological activities we can confirm the previous findings. We find a statistically significant effect on the number of correct decisions only by the blood pressure. Comparing the coefficients in Tables 2.2 and 2.3, we notice that not controlling for the physiological measures lead to a slight overestimation of the impact of audience's gender on participants' decision-making performance. Also, consistent with the finding that women behave in accountable manner regardless of the gender of the audience, in the sample of female participants physiological measures of stress do not vary depending on the pairing with the audience.

To test the notion that irrationalities in choices may due to anchoring at the single-extraction probability and insufficient adjustment away from the probability (Vieider 2011), we examine how choices are influenced by the differences in probability of extracting a winning ball in a single extraction. We create a panel dataset using the choices in the six tasks

[15] The number of participants is not the same in each regression. In columns (2), (3), (5), and (6), in which we control for the heart rate, there are less observations due to technical problems with the equipment.

[16] The HRV indicators that best measure the activity of the Parasympathetic and Sympathetic Systems are LF and HF components. While the LF component expresses the sympathetic modulation, the HF assesses the parasympathetic activity.

entered as consecutive decisions, and estimate a random-effects probit analysis using as dependent variable a dummy equal to one for choices of the superior simple prospect event and zero otherwise. The results show that in the group of male participants the interaction between female audience treatment and the difference in probability of extracting a winning ball in a single-extraction increases significantly the probability of choosing the inferior compound event (Z=-2.57; p-value=0.01). By contrast, we do not find any significant difference on the sample of female players (Z=1.44; p-value=0.168). Overall, this result provides further support for our finding that men paired with a female audience display a lower accountability.

6. Concluding remarks

The accountability of decision makers is important for the functioning of modern institutions, such as corporations and political organizations. In this paper, we have examined for the first time whether males and females shape their degree of responsibility depending on the gender of an audience that calls for an explanation of each choice made.

The experimental results obtained are twofold. First, we found that male participants are strongly influenced by the gender of the audience: when men are paired with an audience of the opposite sex, their level of accountability goes down drastically compared to when they are paired with a male audience. Second, our evidence suggests that women's decisions are not affected by the gender of the audience: we find that the level of accountability does not differ between female or male audience; also female accountability is similar to male level when men are paired with a female audience. However, when men are paired with a male audience, their accountability level is significantly higher. A possible explanation is that male decision-makers may feel competitive, even though the situation does not involve real competition.

The physiological measures, which proxy for the individual stress during the resolution of the task, are in line with our results on the effects of gender pairings in the decision task; while men's blood pressure significantly increases with respect to the baseline value when they are in front of a female audience, there are not significant differences among female participants. Because stress tends to influence performance, this result might be in line with the finding that men perform worse when they are paired with women.

Overall, an interesting result of our paper is that significant differences in accountability show up only when participants are paired with people of the same or opposite gender. This finding supports the idea that examining gender differences in isolation does not offer an

adequate description of the existing differences between men and women; rather, in line with Sutter et al. (2009), it is necessary to analyze how individuals behave once they are paired with the opposite gender.

Future research is needed to improve our understanding of gender differences in accountability. While our paper is restricted to a type of accountability based on a passive audience, it would be important to employ an active audience, which assesses participants' performance and gives them feedback or judgments of their performance. Another extension could be represented by detecting the impact of gender pairings in a strong accountability setting, such as when the audience earns the payoffs from the subjects' choices and subjects must explain their decisions, or in an incentivized accountability setting, when the audience decides on a payoff -reward or punishment- to the subject based on her/his decisions. Finally, in future work we also plan to study the interaction of gender and audience using decision tasks different from the one used in this paper.

7. Bibliography

Ablon S.L., Brown D.P., Khantzian E.J. and Mack J.E. (2013). *Human feelings: Explorations in affect development and meaning*: Routledge.

Adams R.B. and Ferreira D. (2009) "Women in the boardroom and their impact on governance and performance", *Journal of financial economics* 94, 291-309.

Adelmann P.K. and Zajonc R.B. (1989) "Facial efference and the experience of emotion", *Annual review of Psychology* 40, 249-280.

Allen M.T., Stoney C.M., Owens J.F. and Matthews K.A. (1993) "Hemodynamic adjustments to laboratory stress: the influence of gender and personality", *Psychosomatic Medicine* 55, 505-517.

Andreoni J. and Vesterlund L. (2001) "Which is the fair sex? Gender differences in altruism", *The Quarterly Journal of Economics* 116, 293-312.

Arkes H.R., Christensen C., Lai C. and Blumer C. (1987) "Two methods of reducing overconfidence", *Organizational behavior and human decision processes* 39, 133-144.

Atkinson S.M., Baird S.B. and Frye M.B. (2003) "Do female mutual fund managers manage differently?", *Journal of Financial Research* 26, 1-18.

Bar-Hillel M. (1973) "On the subjective probability of compound events", *Organizational behavior and human performance* 9, 396-406.

Bohnet I., Herrmann B. and Zeckhauser R. (2006) "The requirements for trust in gulf and western countries", *Quarterly Journal of Economics*, 811-828.

Brody L.R. and Hall J.A. (2000) "Gender, emotion, and expression", *Handbook of emotions* 2, 338-349.

Buck R., Miller R.E. and Caul W.F. (1974) "Sex, personality, and physiological variables in the communication of affect via facial expression", *Journal of Personality and Social Psychology* 30, 587.

Buss D.M. (1999). *Evolutionary psychology: The new science of the mind*: Allyn & Bacon.

Cadsby C.B., Hamaguchi Y., Kawagoe T., Maynes E. and Song F. (2007) "Cross-national gender differences in behavior in a threshold public goods game: Japan versus Canada", *Journal of Economic Psychology* 28, 242-260.

Charness G. and Rustichini A. (2011) "Gender differences in cooperation with group membership", *Games and Economic Behavior* 72, 77-85.
Croson R. and Gneezy U. (2009) "Gender differences in preferences", *Journal of Economic Literature*, 448-474.
Croson R., Marks M. and Snyder J. (2008) "Groups work for women: Gender and group identity in social dilemmas", *Negotiation Journal* 24, 411-427.
Dwyer P.D., Gilkeson J.H. and List J.A. (2002) "Gender differences in revealed risk taking: evidence from mutual fund investors", *Economics Letters* 76, 151-158.
Eckel C.C. and Grossman P.J. (2001) "Chivalry and solidarity in ultimatum games", *Economic Inquiry* 39, 171-188.
Eckel C.C. and Grossman P.J. (2002) "Sex differences and statistical stereotyping in attitudes toward financial risk", *Evolution and Human Behavior* 23, 281-295.
Fehr-Duda H., De Gennaro M. and Schubert R. (2006) "Gender, financial risk, and probability weights", *Theory and decision* 60, 283-313.
Gagliarducci S. and Paserman M.D. (2012) "Gender interactions within hierarchies: evidence from the political arena", *The Review of Economic Studies* 79, 1021-1052.
Girdler S.S., Turner J.R., Sherwood A. and Light K.C. (1990) "Gender differences in blood pressure control during a variety of behavioral stressors", *Psychosomatic Medicine* 52, 571-591.
Gneezy U., Leonard K.L. and List J.A. (2009) "Gender differences in competition: Evidence from a matrilineal and a patriarchal society", *Econometrica* 77, 1637-1664.
Gneezy U., Niederle M. and Rustichini A. (2003) "Performance in competitive environments: Gender differences", *The Quarterly Journal of Economics* 118, 1049-1074.
Gneezy U. and Rustichini A. (2004) "Gender and competition at a young age", *American Economic Review*, 377-381.
Holtgraves T. and Skeel J. (1992) "Cognitive biases in playing the lottery: Estimating the odds and choosing the numbers", *Journal of Applied Social Psychology* 22, 934-952.
Janis I.L. and Mann L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*: Free Press.
Kavanagh J. (2005). Stress and Performance A Review of the Literature and its Applicability to the Military: DTIC Document.
Kruglanski A.W. and Freund T. (1983) "The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring", *Journal of Experimental Social Psychology* 19, 448-468.
Larson L.C. (2005). *When Girls Stop Competing Against Boys: An Experimental Analysis of the Competitive Behavior of Young Children.* Harvard University.
Lawler K.A., Wilcox Z.C. and Anderson S.F. (1995) "Gender differences in patterns of dynamic cardiovascular regulation", *Psychosomatic Medicine* 57, 357-365.
Lerner J.S. and Tetlock P.E. (1999) "Accounting for the effects of accountability", *Psychological bulletin* 125, 255.
Matsa D.A. and Miller A.R. (2013) "A female style in corporate leadership? Evidence from quotas", *American Economic Journal: Applied Economics* 5, 136-169.
Matthews K.A. and Stoney C.M. (1988) "Influences of sex and age on cardiovascular responses during stress", *Psychosomatic Medicine* 50, 46-56.
McAdoo W.G., Weinberger M.H., Miller J.Z., Fineberg N.S. and Grim C.E. (1990) "Race and gender influence hemodynamic responses to psychological and physical stimuli", *Journal of hypertension* 8, 961-967.

Miller P.M. and Fagley N.S. (1991) "The effects of framing, problem variations, and providing rationale on choice", *Personality and Social Psychology Bulletin* 17, 517-522.
Musini V.M. and Wright J.M. (2009) "Factors affecting blood pressure variability: lessons learned from two systematic reviews of randomized controlled trials", *PloS one* 4, e5673.
Niederle M. and Vesterlund L. (2007) "Do Women Shy Away From Competition? Do Men Compete Too Much?", *The Quarterly Journal of Economics* 122, 1067-1101.
Press C.U. (2008). *Cambridge advanced learner's dictionary*: Ernst Klett Sprachen.
Rivas M.F. (2008) "An experiment on corruption and gender", *University of Granada working*.
Schwieren C. and Sutter M. (2008) "Trust in cooperation or ability? An experimental study on gender differences", *Economics Letters* 99, 494-497.
Scott M.B. and Lyman S.M. (1968) "Accounts", *American sociological review*, 46-62.
Semin G.R. and Manstead A.S. (1983). *The accountability of conduct: A social psychological analysis*: Academic Press London.
Simpson B. (2003) "Sex, fear, and greed: A social dilemma analysis of gender and cooperation", *Social forces* 82, 35-52.
Steyer R., Schwenkmezger P., Notz P. and Eid M. (1997) "Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF) Hogrefe", *Göttingen, Germany*.
Stoney C.M., Matthews K.A., McDonald R.H. and Johnson C.A. (1988) "Sex differences in lipid, lipoprotein, cardiovascular, and neuroendocrine responses to acute stress", *Psychophysiology* 25, 645-656.
Sutter M., Bosman R., Kocher M.G. and van Winden F. (2009) "Gender pairing and bargaining—Beware the same sex!", *Experimental Economics* 12, 318-331.
Svartberg J., von Muhlen D., Schirmer H., Barrett-Connor E., Sundfjord J. and Jorde R. (2004) "Association of endogenous testosterone with blood pressure and left ventricular mass in men. The Tromso Study", *European Journal of Endocrinology* 150, 65-71.
Takemura K. (1993) "The effect of decision frame and decision justification on risky choice", *Japanese Psychological Research*.
Takemura K. (1994) "Influence of elaboration on the framing of decision", *The Journal of Psychology* 128, 33-39.
Tetlock P.E. (1983) "Accountability and complexity of thought", *Journal of Personality and Social Psychology* 45, 74.
Tetlock P.E. (1992) "The impact of accountability on judgment and choice: Toward a social contingency model", *Advances in experimental social psychology* 25, 331-376.
Tetlock P.E. and Boettger R. (1989) "Accountability: A social magnifier of the dilution effect", *Journal of Personality and Social Psychology* 57, 388.
Tversky A. and Kahneman D. (1974) "Judgment under uncertainty: Heuristics and biases", *Science* 185, 1124-1131.
Tversky A. and Kahneman D. (1981) "The framing of decisions and the psychology of choice", *Science* 211, 453-458.
van der Meij L., Buunk A.P., van de Sande J.P. and Salvador A. (2008) "The presence of a woman increases testosterone in aggressive dominant men", *Hormones and Behavior* 54, 640-644.
Vieider F.M. (2011) "Separating real incentives and accountability", *Experimental Economics* 14, 507-518.

Appendix A: Tables and Figures

Table 2.1 Summary statistics

Panel A: individual characteristics	Male	Female
Age participant (average)	22.71	24.30
Age audience (average)	33.50	32.44
Degree/Master in social science (%)	64.10	63.64
Degree/Master in science (%)	12.82	27.27
Degree/Master in humanities (%)	23.08	9.09
Panel B: physiological conditions		
Non- smoker (%)	79.49	81.82
No hypertension problems (%)	97.44	100
Physical activity before the experiment (%)	2.56	2.27
No alcohol consumption before the experiment (%)	73.23	59.66
Height (cm)	176	164

Table 2.2 Regression results

Panel A and B report respectively the results of the OLS and the Ordered Probit regressions. In both panels, Columns (1) and (2) report the results using the sample of male participants; Columns (3) and (4) report the results using the sample of female participants. The dependent variable in Columns (1)-(4) is an index built as a sum of the number of the superior simple prospect choices; the index ranges from 0 to 6 (each simple event choice is counted as 1, and each participant answers to six questions). The treatment represents the group in which the participant's gender is the opposite of the audience's one. The Education dummies are Social Science Degree, Science Degree, and Humanities is used as a reference group (the coefficients are not reported but available upon request). Columns (1) and (3) report the regressions using only as explanatory variable the gender of the audience; Columns (2) and (4) show the results of the regressions using in addition two control variables, participant's age and education. Robust standard errors are reported in parenthesis. *, ** and *** denote significance at 10%, 5% and 1% respectively

Dependent variable: Sum of simple superior prospect choices				
	Male participants		Female participants	
	[1]	[2]	[3]	[4]
Panel A: OLS regression				
Treatment	−2.113***	−2.600***	−0.301	−0.597
	(0.707)	(0.745)	(0.729)	(0.689)
Age participants		−0.077		−0.030
		(0.073)		(0.070)
Panel B: Ordered probit				
Treatment	−1.121**	−1.543 **	−0.229	−0.447
	(0.475)	(0.609)	(0.379)	(0.414)
Age participant		−0.046		−0.018
		(0.043)		(0.042)
Education dummies	No	Yes	No	Yes
Number of participants	28	28	33	33

Table 2.3 Physiological measures regressions

The table reports the results of the OLS regression using the physiological measures as control. Columns (1), (2), and (3) report the results using the sample of male participants; Columns (4), (5), and (6) report the results using the sample of female participants. The dependent variable in Columns (1)-(6) is an index built as a sum of the number of the superior simple prospect choices; the index ranges from 0 to 6 (each simple event choice is counted as 1, and each participant answers to six questions). The treatment represents the group in which the participant's gender is the opposite of the audience's one. Columns (1) and (4) report the regressions using as explanatory variable the gender of the audience and the difference between the second and the first systolic value of the blood pressure; Columns (2) and (5) report the regressions using, in addition to the gender of the audience, the ratio between the low and high frequency of the heart rate; Columns (3) and (6) show the results of the regressions using both the physiological measures, blood pressure and heart rate. Robust standard errors are reported in parenthesis. *, ** and *** denote significance at 10%, 5% and 1% respectively

Dependent variable: sum of simple superior prospect choices						
	Male participants			Female participants		
	[1]	[2]	[3]	[4]	[5]	[6]
Treatment	−2.423***	−2.043***	−2.390***	0.206	0.311	0.144
	(0.697)	(0.786)	(0.790)	(0.724)	(0.844)	(0.812)
Δ Systolic	0.051*		0.061*	−0.038		−0.053
	(0.029)		(0.030)	(0.042)		(0.045)
LF/HF Ratio		−0.00015	−0.00015		−0.004	−0.004
		(0.002)	(0.002)		(0.005)	(0.005)
Number of participants	28	24	24	33	30	30

Figure 2.1 Heart rate

The figure represents the measure of the heart rate during the experiment. On the time bar, the blue part in the lower part of the graph, there are three markers: the first, approximately at the eighth minute, marks the moment in which the participant sees the audience for the first time; the second is at the nineteenth minute and it is marked just before the explanation starts; the last one is at the twenty-fourth minute and it is marked at the end of the explanation.

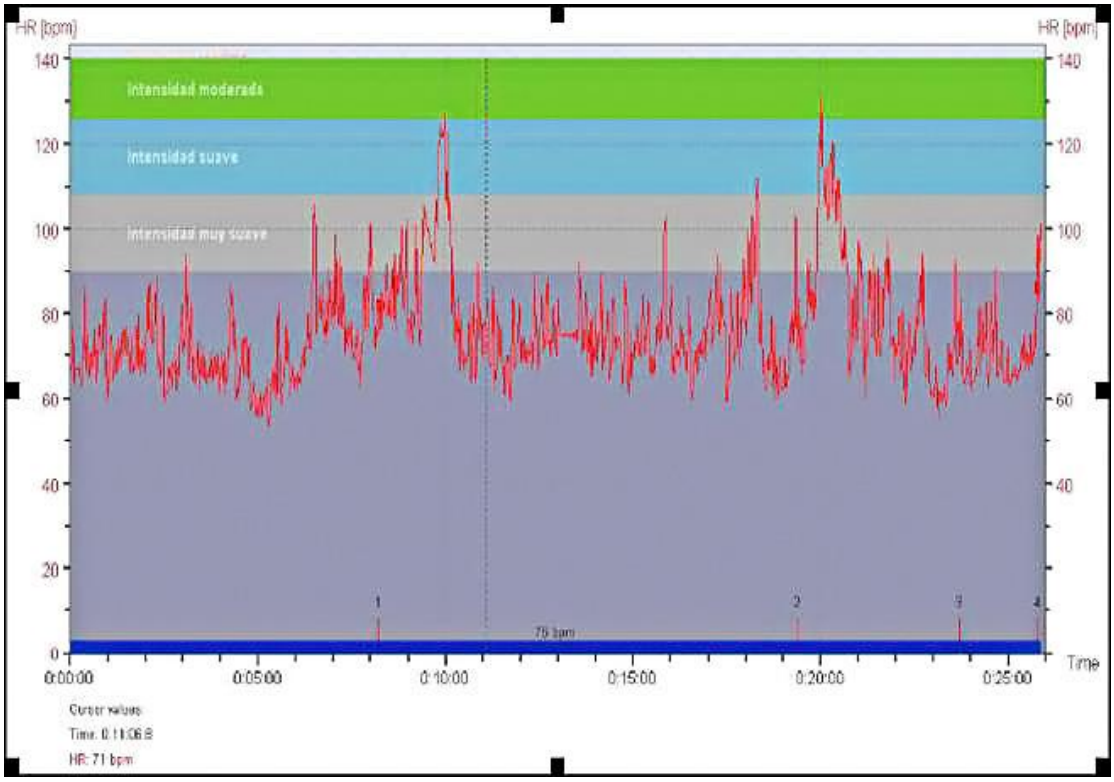


Figure 2.2 (A) Correct answers by audience's gender

The graphs report the mean number of choices of the superior simple event. The choice is encoded as 1 if it is simple prospect and 0 if compound prospect. Panel A reports the choices of the whole sample but distinguishing for the gender of the audience; Panel B reports the results of female and male participants without a distinction of the gender of the audience.

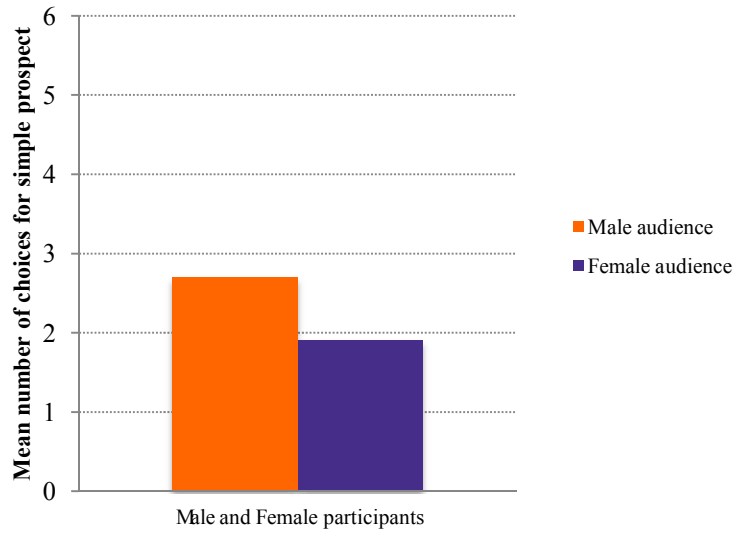


Figure 2.2 (B) Correct answers by participant's gender

The graphs report the mean number of choices of the superior simple event. The choice is encoded as 1 if it is simple prospect and 0 if compound prospect. Panel A reports the choices of the whole sample but distinguishing for the gender of the audience; Panel B reports the results of female and male participants without a distinction of the gender of the audience.

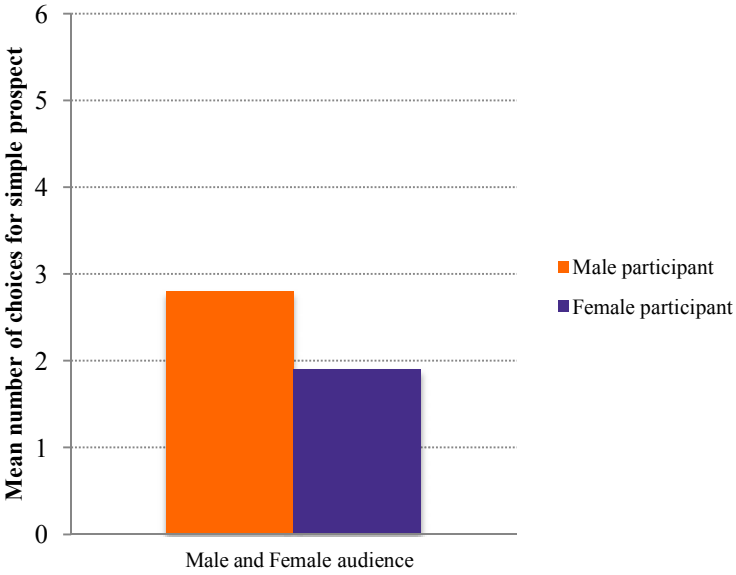


Figure 2.3 Correct answers by gender pairings

The left part of the graph reports the values referred to male participants when they justify their choices in front of male or female audience. The right part shows the results about female participants when paired with a male or female audience.

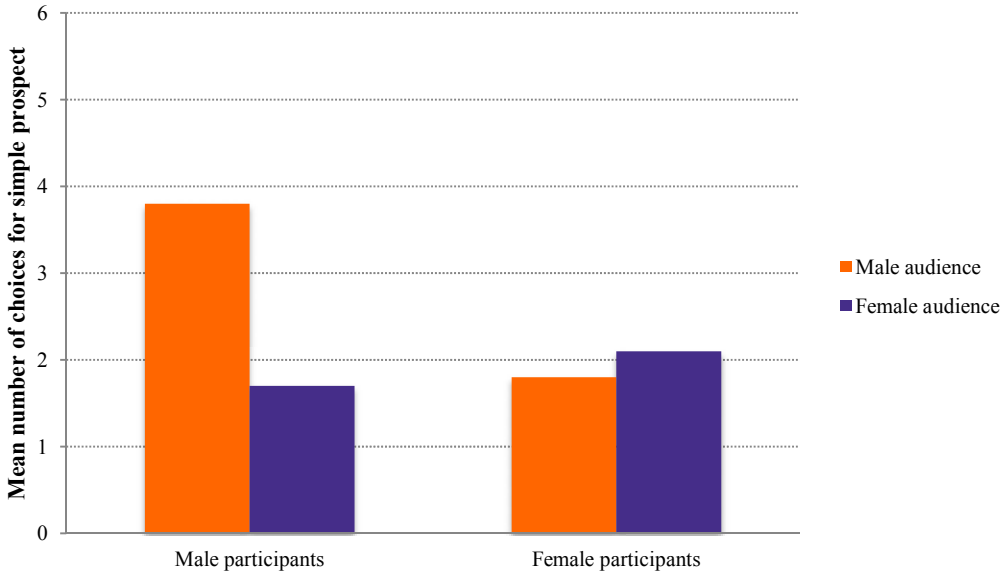


Figure 2.4 (A) Systolic values of male participants

The graph reports three values of systolic blood pressure of male participants when paired with male or female audience. The first value is taken at the beginning of the experiment; the second at the end of the task and before the explanation starts; the last after the explanation.

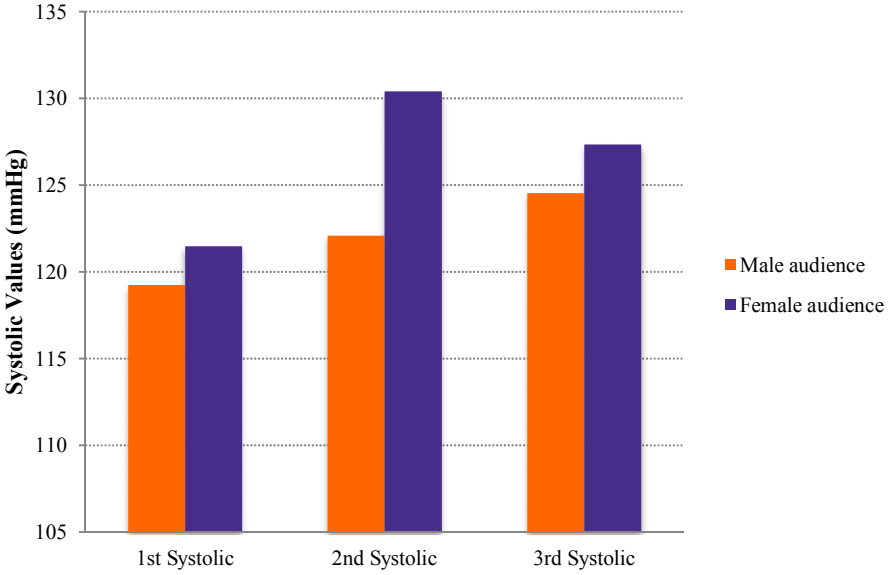
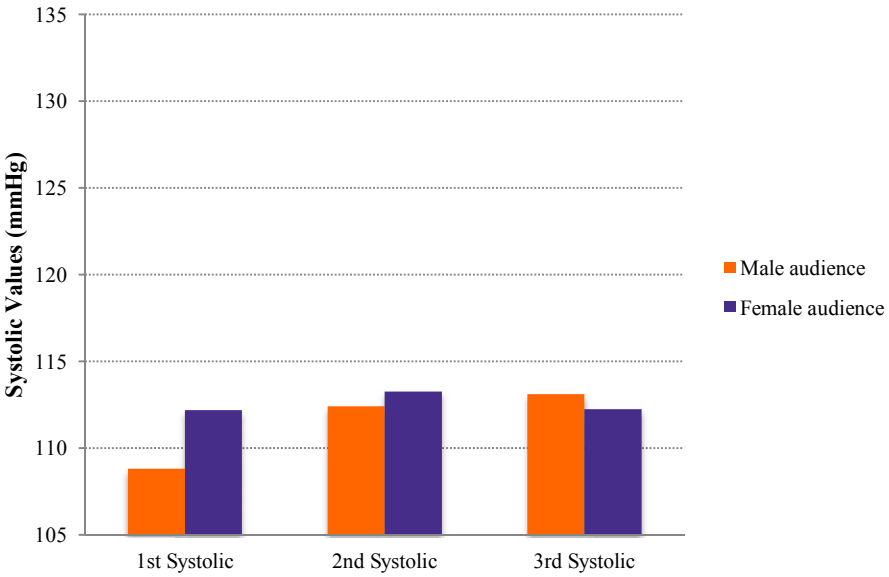


Figure 2.4 (B) Systolic values of female participants

The graph reports three values of systolic blood pressure of female participants when paired with male or female audience. The first value is taken at the beginning of the experiment; the second at the end of the task and before the explanation starts; the last after the explanation.



CHAPTER III

INCENTIVES, GENDER PAIRINGS, AND ACCOUNTABILITY

1. Introduction

In Chapter III, we use a laboratory experiment to study decision-makers' accountability in a setting where the choices are incentivized and where the decision-makers are asked to share their final payoffs with the members of an audience. This chapter includes a slight revision of Brandts and Garofalo (2013) with some extensions based on new treatments which enable us to isolate the accountability effect of payoff sharing and monetary incentives.

In modern societies, the accountability of decision-makers is becoming an increasingly important consideration in the design of policies aimed at improving the quality of various public services, e.g. health care, criminal justice, politics, and education. Understanding the drivers of "accountable behavior" of public decision-makers is thus stimulating academic research in various fields such as political economics, e.g. Nannicini, Stella, Tabellini and Troiano (2010), and psychology (Lerner and Tetlock 1999; Tetlock 1992)

According to the psychology literature, accountability refers to the implicit or explicit expectation that one may be called to justify one's belief, feelings, and actions to others (Tetlock 1992). It has been proved that this expectation significantly increases the likelihood that the individual makes the right choice (Vieider 2011; Lerner and Tetlock 1999).

An important question in these studies is whether, in an accountable setting whereby individuals have to justify their choices to an audience, the provision of effort by a player depends on the demographic characteristics of the player herself and the audience members. Gender is, in our opinion, one of the most potentially interesting variables. For example, in relation to the presence of women both in CEO positions and on the board of directors of corporations, it is crucial to understand whether e.g., a monitoring committee spurs accountability differently depending on the genders of participants and committee members.

In Chapter II, Brandts and Garofalo (2012) present an experiment to explore how the gender composition of the audience interacts with the gender of a decision-maker thereby shaping her/his level of effort. On one hand, we find that women's level of effort is not influenced by the gender of the audience. On the other hand, men's effort turns out to be strongly affected by the audience's gender and it increases when they are paired with a male audience. This article, however, studies accountability in an environment in which the payoffs of both decision-maker and audience do not depend on the performance in the decision task.

In this chapter, we extend Brandts and Garofalo (2012)'s experiment to a context in which individuals make incentivized decisions and share payoffs with the audience members. Thus, in our experiment we create an accountable setting and study how the effort is influenced by both the gender pairings between the decision-maker and the audience, and the compensation scheme. The advantage of this new context is that it is better suited to describe real-world situations, such as the above-mentioned example of CEOs and boards of directors, in which accountable or unaccountable decisions have consequences not only for the decision-maker but also for the counterpart.

In the main treatment of the present study, participants are paid based on their performance and the total amount earned is equally divided with the audience. On the one hand, the participants' effort may be incentivized by the variable payoffs scheme; on the other hand, their effort may be influenced by the payoff sharing component. In order to disentangle the effect of the two components of the compensation scheme, incentives and payoff sharing, we run two additional treatments in which (1) we pay participants a flat fee and keep the payoff sharing with the audience, and (2) we only incentivize participants' decision (without payoff sharing). Finally, we run a treatment without audience, i.e. in a context in which the participants solve the task but do not need to justify their choice, and in which we pay participants a variable amount depending on the number of correct answers.

What are our research hypotheses for the new study? The most basic economic intuition would suggest the presence of incentives should tend to equalize the performance of men and women called to justify their choices in front a female or male audience. In other words, in the absence of incentives it may be natural for various factors to influence behavior, but with incentives men and women will behave like *homo oeconomicus;* then they should behave equally in front a female or male audience. Brandts and Garofalo (2012) found no differences in female's behavior depending on the audience's gender, and important differences in men's behavior. The above intuition would imply that the introduction of incentives should eliminate any differences in behavior.[17]

However, we have to take into account the second new element of our experimental setting, i.e. the fact that the audience earns exactly the same amount as the decision-maker. Purely selfish individuals will not be affected by this fact, thus their effort should be mostly influenced by the *incentive component* of the compensation scheme.

[17] As such, this result should be consistent with Niederle and Vesterlund (2007), who examine whether men and women of the same ability differ in their selection into a competitive environment. As pre-treatment they run some sessions in which both male and female participants are paid with a piece-rate scheme in a non-competitive setting; their results indicate that there are no gender differences in performance.

On the contrary, individuals who are positively affected by others' payoffs are expected to be further motivated by the payoff sharing component. Hence, overall we expect no gender differences in the average number of correct answers in the decision task depending on the audience's gender.

Contrary to this expectation, our results show that women's cognitive effort is influenced by the audience's gender when their compensation scheme includes both incentives and payoff sharing; and that male participants in this new context do not behave differently depending on the audience's gender. The results of the additional treatments show that the women's cognitive effort level is not influenced by the audience gender when in their compensation scheme is present only one of the two elements, either incentives or payoff sharing, but only when both elements are included in the compensation scheme. On the contrary, male's effort provision is influenced by the audience's gender when the compensation scheme includes only incentives.

The chapter is organized as follows. Section 2 illustrates the experimental design. Section 3 presents the data and the results. Section 4 concludes.

2. Design of experiment

2.1. Experimental procedure

We conducted the experiment at Universitat Autònoma de Barcelona. The participants were recruited by email using a database (ORSEE) of student who voluntary registered in the database to participate to previous experiment. The subjects can participate only in one session and they are randomly allocated to a male or female audience, and to one of the four treatments. The final profits of each subject are equal to the sum of the earnings she makes during the experiment (either variable or flat) plus a 5€ show-up fee. Subjects are paid privately and in cash right after the experiment.

2.2. Experimental design

Each session takes place in two rooms: the audience's room and the resolution room. The audience's room is a semi-empty room containing only four chairs and a big table in which the participants sit on one side and the audience on the opposite side. The resolution room is a big room in which there are approximately ten desks used for the resolution of the task. The audience is composed by three members[18] (either three women or three men), and they are recruited from PhD students at the Business Department of Universitat Autònoma de

[18] The audience members are not aware of the purpose of the experiment.

Barcelona. We exclude from the experiment the Master student of the Business Department because they may know in person the audience members.

The session is divided into several parts and it follows the same structure independently of the treatment (except for the treatment without audience).

First, the experimenter hands to the participant the first part of the instructions in which the more general conditions of the experiment are explained, and then the experimenter accompanies the participant to the audience's room for the first meeting with the audience. For the entire duration of the experiment, the audience stays in the room waiting for the arrival of each participant. Each meeting with the participant is held in private.

Second, during the private meeting, the audience has the duty to hand in to the participant the second part of the instructions, a document to be filled in with personal information, the task sheets, and to carefully explain the compensation scheme to the participant[19]. The audience is asked to exhibit a neutral behavior and to pay attention to the participant during the explanation. We asked to the members of the audience to look professional but not to directly stress or make pressure on the participants. Before starting the resolution of the task, the participant is asked to answer to two comprehensive questions in order to be sure she/he understood the compensation scheme.

Third, the participant has to move to the resolution room where she/he has to solve the task, six math problems (see Appendix), without any time limit. During the resolution of the task, the experimenter takes the time each participant spends in the resolution room. Once the participants solved the six problems, they have to come back to the audience's room, again one by one, and justify their answers in front of the audience, the same one that explained them the rules of the experiment. Finally, the audience communicates to the participant the number of correct answers and the amount gained. The participant leaves the audience's room and goes to the check-in desk to sign the receipt for the payment and being paid.

The task consists of six problems in which the participant has to choose between simple and compound prospects (as in Brandts and Garofalo 2012; Vieider 2011; Holtgraves and Skeel 1992; Bar-Hillel 1973). The simple event[20] represents always the correct answer because of the higher probability to draw the winning ball with respect to the compound event[21]. The structure of the six problems is kept the same for all of them; the differences consist of the level of analytical difficulty, number of extractions, and composition of the

[19] The compensation scheme changes depending on the treatment.
[20] The simple event consists of drawing a single ball from a bag containing a certain number of balls.
[21] The compound event consists of drawing more than one ball with replacement from a bag composed by a certain number of balls.

bag.[22]

This task has been used in previous research to study also the anchoring and adjustment effect. As explained in Tversky and Kahneman (1974), the anchoring effect is a cognitive bias that influence the way people intuitively assess probabilities. People may be anchored at an initial value that may be provided externally in the formulation of the problem, derive from incomplete calculations of a number, or consist in a "natural reference point". The bias consists in insufficient adjustment away from the anchor value. The main feature of the task used in the experiment is that just reading or not calculating the success probability lead to incorrect answers; on the other hand, solving all the calculations likely gives the correct answers. Then, a low cognitive effort leads to an insufficient adjustment from the reference point.

The stated probability of the elementary event (successful at any stage) provides a natural starting point for the estimation of the probabilities of the compound event. Since the adjustment from the starting point is typically insufficient, the final estimates remain too close to the probability of the elementary events (Tversky and Kahneman 1981). Given these findings, making the simple lottery the correct choice allows for a better detection of the effort exerted in making the decision.

We run four treatments in which we manipulate the presence of the audience, the gender of the audience[23], and the compensation scheme. In the first three treatments the participants have to justify their choices in front of the audience; on the contrary, in the last treatment there is no audience, and the participants are asked only to solve the task.

In the main treatment, *incentives and payoff sharing*[24], the participants are paid a variable amount based on their performance, and they are forced to share equally their final payoffs with the audience (see Appendix). Hence, in the main treatment both the participants and the audience are paid based on the participants' performance. In the first control treatment,

[22] Bar-Hillel (1973) reported an example in which subjects were asked to choose between one of two events. Three types of events were used: (1) simple events, such as drawing a red ball from an urn containing 50 red balls and 50 black balls; (2) conjunctive events, such as drawing a red ball seven times in succession, with replacement, from a bag containing 90 red balls and 10 black balls; and (3) disjunctive events, such as drawing a red ball at least once in seven successive tries, with replacement, from a bag containing 10 red balls and 90 black balls. A significant majority of subjects chose the conjunctive event (probability of 0.48) rather than the simple event (probability of 0.50); they also preferred the simple event rather than the disjunctive event (probability of 0.52). These biases can be explained as effects of an anchoring and adjustment process.

[23] Female audience paired either with female or male participant, and male audience paired either with female or male participant.

[24] The participants gain 4€ for each correct answer but this amount has to be shared equally with the audience. E.g. the participant answers correctly to one answer, he/she gets 4€ but half of this amount is for the audience. So, at the end of the experiment the participant has gained 2€ and the audience as well. Another example, the participant scores 6 correct answers, she/he gets 24€ but half of this amount is for the audience, at the end of the experiment the participant has earned 12€ and the audience as well.

incentives (*only)*[25], participants are paid based on their performance and they do not have to share their final payoffs with the audience. Audience members are paid a flat amount per hour. In the second control treatment, *payoff sharing* (*only)*[26], participants are paid a flat fee while the audience is paid based on participants' performance. Finally, to control for the mere presence of the audience, we run a fourth treatment, *incentives and no audience.* Here, participants are paid based on their performance and they do not have to justify their choices in front of the audience. This is the only treatment without audience.

Depending on the treatment, the audience plays a *passive* or an *active* role even if the members of the audience are not asked to interact with the participant during the resolution of the task. In particular, the active role is given by the payoff sharing rule, and the passive role is given by the absence of any sharing rule.

3. Data and results

The dataset contains personal details about each participant and audience member, the six choices made by the participants during the resolution of the task, and the time spent for the resolution of the task. In addition, we have information about gender, age, education (field) of the participants, and gender and age of the audience.

To measure the effort level provided during the resolution of the task, we created a simple index composed by the sum of the correct answers obtained; the index ranges from zero (lowest number of correct answers) to six (highest number of correct answers).

During the resolution of the task, the experimenters took the time each participant spent in the resolution room. We find that in each treatment male spent less time solving the task when they were paired with a female audience, but this difference is statistically different only in the main treatment.[27] We obtain a similar result for female participants in front a female audience in the main treatment.

Participants are in total 472, 177 in the *Incentives and payoff sharing* treatment, 107 participants in the *payoff sharing* (*only)* treatment, 130 in *incentives* (*only)* treatment, and 58 in the (*only) incentives and no audience* treatment. The sample is composed by 47% male and 53% female, and all are students (from bachelor, master or PhD[28]): about 50% from social science, 36% from science, and 12% from humanities studies. Most of the participants are

[25] The participants are paid 2€ for each correct answer.
[26] The participants are paid a flat fee of 9€, and the audience is paid based on the participant's performance, 2€ each correct answer.
[27] Summary statistics not reported but available upon request.
[28] In the dataset there are only few PhD students and they do not belong to the same department of the audience members, so they do not know each other.

between 18 and 34 years old (just few participants are older than 34). Table 3.1 reports the individual characteristics of participants and audience members. The average age of participants and audience members is, respectively, about 24 and 30.

3.1. Methodology

In this paragraph, we describe the methodology used for the statistical analysis. Our goal is to estimate the effect of the gender pairing and the compensation scheme on the cognitive effort provided when the participants are called to justify their choices in front of an audience.

Our dependent variable is the above-discussed *Effort Index* (ranging from zero to six), given by the sum of the correct (superior simple event) answers. In Table 3.2 we report the summary statistics of the Effort index by treatment and gender pairings.

Given the structure of the dependent variable, we use a count data model. One option could be the *Poisson* model but a key requirement for its validity is that the conditional mean should equal the conditional variance. We run the goodness of fit test of the Poisson model and in our dataset the assumption about the conditional mean equaling the conditional variance is violated. Hence, we decided to use a similar model, *Negative Binomial* model that does not rely on the same restrictive assumption of the Poisson model.

We estimate the following model by participant's gender and treatment:

$$\text{Effort Index}_i = \beta_0 + \beta_1(\text{Audience gender}) + \delta_t + \varepsilon_{it}$$

in which the main explanatory variable, Audience gender, is a dummy variable equal to 1 if the gender of the audience and the participant is the opposite. The model includes also a number of controls such as the age of the audience, the age of the participants and dummies for their field of study (social science, science, ad humanities).

3.1.1. Incentives and payoff sharing

In this paragraph, we show how the gender pairing affects the cognitive effort provision when the participants' compensation scheme includes both incentives and payoff sharing.

Table 3.3 reports the results of the negative binomial regression, separately for female (Column 1) and male (Column 2), in which the dependent variable is the sum of the superior simple event choices (Effort index). The results show that for female participants being paired with the audience of the opposite gender has a positive and statistically significant effect on

the sum of the correct answers (p-value=0.040), the Effort index increases by 0.48 points. By contrast, the effect is quantitatively small and statistically insignificant for men (p-value=0.838).

From these results we cannot infer whether the effect is driven by the incentives, or by the payoff sharing. In order to disentangle the two effects, we run two additional treatments. In the first one, we only incentivize the decision and we do not allow any payoff sharing. In the second one, we pay the participants a flat fee and pay the audience a variable amount depending on the participants' performance. In the next paragraphs we present the results obtained.

3.1.2. Payoff sharing only

In this paragraph we present the results about the treatment in which the participants are paid a flat fee and the audience's final payoffs totally depend on the participants' performance. In particular, participants are paid 9€, while the audience gets 2€ for each correct answer the participant chooses. Table 3.4 illustrates that for both female and male participants being paired with an audience of the opposite gender does not influence their cognitive effort in an accountable setting if the compensation scheme includes only payoff sharing.

From these results we can infer that the significant effect of gender pairing on female's cognitive effort we found in the main treatment[29] is not driven by the payoff sharing component.

3.1.3. Incentives only

In this paragraph, we present the results about the treatment in which the participants' choices are incentivized and the audience is paid a flat fee per hour. Table 3.5 shows that in a setting where the participants' compensation scheme is only based on incentives, being paired with an audience of the opposite gender has no effect on female's cognitive effort provided during the resolution of the task (Column 1). By contrast, when male participants are paired with a female audience (Column 2) we find that the Effort index decreases by 0.387 points and the coefficient is statistically significant at the 5% level, thus male's cognitive effort is negatively affected by the audience's gender.

[29] In the main treatment, the participants are paid based on their performance and their final payoffs are shared with the audience

This result suggests that female's behavior is positively influenced only by the combination of the two elements, incentives and payoff sharing.

Finally, in order to control for the presence of the audience, we run an additional treatment without audience. We pay the participants a variable amount depending on the number of correct answers and they do not need to justify their choices. The results show that there are not statistical differences between the average number of correct answers in the treatment with incentives and audience and in the one without audience.

3.2. Comparison with Chapter II

In this paragraph we compare the findings of Chapter II with the ones of Chapter III. On one hand, in Brandts and Garofalo (2012), presented in Chapter II, we find that men's effort provision is influenced by the gender of the audience. They provide a statistically significantly lower effort when paired with an audience of the opposite gender. By contrast, women's effort provision is not influenced by the audience gender. On the other hand, in Chapter III we introduce a new element, e.g. different kind of compensation scheme, and we find that both male and female are influenced by the gender of the audience. Men's effort is negatively influenced when paired with an audience of the opposite gender and when their compensation scheme includes only incentives. Concerning female participants, we find that their level of effort is positively influenced when paired with a male audience and when their compensation scheme includes both incentives and payoff sharing.

4. Conclusion

We have conducted an experiment to extend Brandts and Garofalo (2012) and further examine the relationship between gender and accountability. We use a context-based accountability design in which participants are paid depending on the performance and/or they have to share their payoffs with the members of an audience (either all men or all women).

We found that the audience's gender influences women's level of cognitive effort provided during the resolution of the task; by contrast, men's behavior is not affected by the audience's gender. Our results are unexpected since they reject our intuitive hypothesis that in the presence of monetary incentives and revenue sharing the gender pairings of decision-maker and audience would have no effect on cognitive effort. At the same time we can say that the differences we find (and the absence thereof) are in line with some of our original ideas on the interaction between gender pairings and accountability. The main initial motivation for

studying this topic stems from thinking about situations where top female executives have to justify themselves in front of all male boards. Our initial hypothesis was that women would be affected by the gender of the audience while men would not, as we find in the data reported above.

However, what is puzzling is the comparison between the gender pairing effects on men's behavior that we found in Brandts and Garofalo (2012) and the effect for women found in the present study. One possibly relevant factor is that, in our previous study, we measured blood pressure and heart rate for each participant involved in the experiment. We did this to detect potential gender differences in stress during the justification in front of the audience. We found that blood pressure and heart rate variations only had minor effects on behavior and, therefore, we decided to exclude these measurements from the experiment for this chapter. However, the blood pressure and heart rate measurements could itself have altered participants' alertness and have lead to differences in the behavior of men. In the defense of the design used in this chapter we can say that both the presence of the incentives and payoff sharing, and the absence of the blood pressure and heart rate measurements seem to us to be the most natural environment.

5. Bibliography

Bar-Hillel M. (1973) "On the subjective probability of compound events", *Organizational behavior and human performance* 9, 396-406.

Brandts J. and Garofalo O. (2012) "Gender pairings and accountability effects", *Journal of Economic Behavior & Organization* 83, 31-41.

Brandts J. and Garofalo O. (2013) "11 Does sharing payoffs affect gender differences in accountability?", *Neuroscience and the Economics of Decision Making* 5, 197.

Holtgraves T. and Skeel J. (1992) "Cognitive biases in playing the lottery: Estimating the odds and choosing the numbers", *Journal of Applied Social Psychology* 22, 934-952.

Lerner J.S. and Tetlock P.E. (1999) "Accounting for the effects of accountability", *Psychological bulletin* 125, 255.

Nannicini T., Stella A., Tabellini G. and Troiano U. (2010) "Social capital and political accountability".

Niederle M. and Vesterlund L. (2007) "Do Women Shy Away From Competition? Do Men Compete Too Much?", *The Quarterly Journal of Economics* 122, 1067-1101.

Tetlock P.E. (1992) "The impact of accountability on judgment and choice: Toward a social contingency model", *Advances in experimental social psychology* 25, 331-376.

Tversky A. and Kahneman D. (1974) "Judgment under uncertainty: Heuristics and biases", *Science* 185, 1124-1131.

Tversky A. and Kahneman D. (1981) "The framing of decisions and the psychology of choice", *Science* 211, 453-458.

Vieider F.M. (2011) "Separating real incentives and accountability", *Experimental Economics* 14, 507-518.

Appendix A: Tables

Table 3.1 Summary statistics

Individual characteristics	Male	Female
Age participants (average)	23.55	23.78
Age audience (average)	31.70	28.18
Degree/Master in social science (%)	50.68	51.64
Degree/Master in science (%)	37.56	35.25
Degree/Master in humanities (%)	11.31	13.11

Table 3.2 Effort index by treatment and gender pairings

This table reports the summary statistics of the Effort index by treatments and gender pairings. The Effort index is built as sum of the simple superior prospect choices; the index ranges from zero to six (each simple event choice is counted as 1, and each participant answers to six questions).

Effort index (average)				
	Incentives and payoff sharing	Incentives only	Payoff sharing only	Incentives and no audience
Male participants and male audience	3.47	3.31	3.85	3.17
Male participants and female audience	3.02	3.12	3.69	
Female participants and male audience	1.78	2.51	2.38	2.38
Female participants and male audience	1.38	1.82	1.86	

Table 3.3 Incentives and payoff sharing

This table reports results from the Negative Binomial regressions. The dependent variable is the Effort index, built as sum of the simple superior prospect choices; the index ranges from zero to six (each simple event choice is counted as 1, and each participant answers to six questions). Column (1) reports the results using only female participants, whereas Column (2) shows the results using only male participants. Treatment is a dummy equal to one if the gender of the audience is opposite to the participants' gender, and zero otherwise. Education reports the field of study: social science, science, and humanities. Standard errors are clustered by states of headquarter. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

Dependent variable: Effort index		
	Female participants	Male participants
	[1]	[2]
Treatment	0.478**	-0.031
	(0.233)	(0.153)
Age participants	-0.015	0.006
	(0.029)	(0.016)
Age audience	-0.056	0.035
	(0.036)	(0.022)
Education	✓	✓
Number of observations	85	90

Table 3.4 Payoff sharing

This table reports results from the Negative Binomial regressions. The results refer to the *payoff sharing (only)* treatment. The dependent variable is the Effort index, built as sum of the simple superior prospect choices; the index ranges from zero to six (each simple event choice is counted as 1, and each participant answers to six questions). Column (1) reports the results using only female participants, whereas Column (2) shows the results using only male participants. Treatment is a dummy equal to one if the gender of the audience is opposite to the participants' gender, and zero otherwise. Education reports the field of study: social science, science, and humanities. Standard errors are clustered by states of headquarter. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

Dependent variable: Effort index	Female participants	Male participants
	[1]	[2]
Treatment	0.165	0.269
	(0.336)	(0.187)
Age participants	0.012	-0.024
	(0.033)	(0.021)
Age audience	0.022	0.104***
	(0.062)	(0.039)
Education	✓	✓
Number of observations	60	46

Table 3.5 Incentives

This table reports results from the Negative Binomial regressions. The results refer to the *incentives (only)* treatment. The dependent variable is the Effort index, built as sum of the simple superior prospect choices; the index ranges from zero to six (each simple event choice is counted as 1, and each participant answers to six questions). Column (1) reports the results using only female participants, whereas Column (2) shows the results using only male participants. Treatment is a dummy equal to one if the gender of the audience is opposite to the participants' gender, and zero otherwise. Education reports the field of study: social science, science, and humanities. Standard errors are clustered by states of headquarter. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

Dependent variable: Effort index	Female participants	Male participants
	[1]	[2]
Treatment	0.318	-0.387**
	(0.504)	(0.192)
Age participants	-0.061*	0.005
	(0.034)	(0.016)
Age audience	0.022	-0.082**
	(0.087)	(0.033)
Education	✓	✓
Number of observations	70	56

Appendix B: Task

Date _____ Name and Surname_____Code _1_

Choice between Simple and Compound Prospects

In the following pages you find 6 problems. In each problem you are asked to choose between two options. One option allows you to extract only one ball from a bag; the other one allows a "multiple extraction" from a different bag. In the multiple extraction option, the ball you have extracted will be placed back in the bag and the balls in the bag will be mixed before you extract again, so as to keep the composition of the bag constant.
These rules are valid for all the problems below. Please pay attention to the composition of the bags and the number of extractions because both vary across problems.
Please, mark one of the two options for each problem. Choose the option you prefer in this way:

__X__Option # (extract ## time from a bag with ## red and ## green balls, win if red)
_____Option # (extract ## time from a bag with ## red and ## green balls, win if red)

or

_____Option # (extract ## time from a bag with ## red and ## green balls, win if red)
__X__Option # (extract ## time from a bag with ## red and ## green balls, win if red)

Please, take into account that the amount of money you will win (additional to the 5€ show-up fee) depends only on your choices.

For each amount you win the audience will receive the same amount. For example:

-if after solving all the problems you win 0 €: 0€ will be for the audience and 0€ for you.
-if after solving all the problems you win 4 €: 2€ will be for the audience and 2€ for you.
-if after solving all the problems you win 8 €: 4€ will be for the audience and 4€ for you.
-if after solving all the problems you win 12€: 6€ will be for the audience and 6€ for you.
-if after solving all the problems you win 16€: 8€ will be for the audience and 8€ for you.
-if after solving all the problems you win 20€: 10€ will be for the audience and 10€ for you.
-if after solving all the problems you win 24€: 12€ will be for the audience and 12€ for you.

Code _1_

Comprehension questionnaire

Please mark the correct answer.

1. If you answer correctly to one question and obtain 4 €, the audience gain:
    a. 0 € and you gain 4 €
    b. 4 € and you gain 4 €
    c. 2 € and you gain 2 €

2. If you answer correctly to two questions and obtain 8 €, the audience gain:
    a. 4 euros y tú ganas 4 euros
    b. 8 euros y tú ganas 4 euros
    c. 6 euros y tu ganas 2 euros

Problem 1

You were given a choice between two options to win 4€.
Option A allows you to extract one ball from a bag containing 10 red and 10 green balls. If you extract a red ball, you win 4€; if you extract a green ball, you win nothing.
Option B allows you to extract 7 balls in sequence with replacement from a bag containing 18 red balls and 2 green balls. If all 7 balls extracted are red you win 4€; if one or more of the balls extracted are green, you win nothing.

What would you choose?

____Option A (extract 1 time from a bag with 10 red and 10 green balls, win if red)
____Option B (extract 7 times from a bag with 18 red and 2 green balls, win if 7 times red)

Problem 2

You were given a choice between two options to win 4€.
Option A allows you to extract 5 balls in sequence with replacement from a bag containing 15 red balls and 5 green balls. If all 5 balls extracted are red you win 4€; if one or more of the balls extracted are green, you win nothing.
Option B allows you to extract one ball from a bag containing 5 red and 15 green balls. If you extract a red ball, you win 4€; if you extract a green ball, you win nothing.

What would you choose?

____Option A (extract 5 times from a bag with 15 red and 5 green balls, win if 5 times red)
____Option B (extract 1 time from a bag with 5 red and 15 green balls, win if red)

Code _1_

Problem 3

You were given a choice between two options to win 4€.
Option A allows you to extract 7 balls in sequence with replacement from a bag containing 16 red balls and 4 green balls. If all 7 balls extracted are red you win 4€; if one or more of the balls extracted are green, you win nothing.
Option B allows you to extract one ball from a bag containing 5 red and 15 green balls. If you extract a red ball you win 4€; if you extract a green ball, you win nothing.

What would you choose?

____Option A (extract 7 times from a bag with 16 red and 4 green balls, win if 7 times red)
____Option B (extract 1 time from a bag with 5 red and 15 green balls, win if red)

Problem 4

You were given a choice between two options to win 4€.
Option A allows you to extract one ball from a bag containing 2 red and 18 green balls. If you extract a red ball, you win 4€; if you extract a green ball, you win nothing.
Option B allows you to extract 4 balls in sequence with replacement from a bag containing 10 red balls and 10 green balls. If all 4 balls extracted are red you win 4€; if one or more of the balls extracted are green, you win nothing.

What would you choose?

____Option A (extract 1 time from a bag with 2 red and 18 green balls, win if red)
____Option B (extract 4 times from a bag with 10 red and 10 green balls, win if 4 times red)

Code _1_

Problem 5

You were given a choice between two options to win 4€
Option A allows you to extract 6 balls in sequence with replacement from a bag containing 15 red balls and 5 green balls. If all 6 balls extracted are red you win 4€; if one or more of the balls extracted are green, you win nothing.
Option B allows you to extract one ball from a bag containing 4 red and 16 green balls. If you extract a red ball, you win 4€; if you extract a green ball, you win nothing.

What would you choose?

____Option A (extract 6 times from a bag with 15 red and 5 green balls, win if 6 times red)
____Option B (extract 1 time from a bag with 4 red and 16 green balls, win if red)

Problem 6

You were given a choice between two options to win 4€.
Option A allows you to extract one ball from a bag containing 6 red and 14 green balls. If you extract a red ball, you win 4€; if you extract a green ball, you win nothing.
Option B allows you to extract 2 balls in sequence with replacement from a bag containing 10 red balls and 10 green balls. If all 2 balls extracted are red you win 4€; if one or more of the balls extracted are green, you win nothing.

What would you choose?

____Option A (extract 1 time from a bag with 6 red and 14 green balls, win if red)
____Option B (extract 2 times from a bag with 10 red and 10 green balls, win if 2 times red)

CHAPTER IV

GENDER INTERACTIONS WITHIN THE FAMILY FIRM

1. Introduction

Motivated by the observation that women are dramatically under-represented in leadership positions (see e.g. Bertrand and Hallock 2001), policy makers around the world are devoting considerable attention to the participation of women in the labor market. This attention has led to such major policy interventions as the introduction of political quotas in India and of board quotas in Norway.

Although experimental works have suggested that women generally exhibit less willingness than men to engage in competitive activities (Niederle and Vesterlund 2007) and worse performance when subject to competitive pressures (Gneezy and Rustichini 2004; Gneezy, Niederle and Rustichini 2003), the effect of female leaders on corporate performance remains unclear. For instance, Adams and Ferreira (2009) report an ambiguous effect of female directorship on firm profitability whereas Dezsö and Ross (2012) find that female management can benefit performance in some contexts. We contribute to this literature by analyzing whether gender interactions in the top layers of a firm's hierarchy affect corporate profitability. Using a comprehensive data set of family-controlled firms in Italy over the period 2000–2010, we show that companies led by female chief executive officers (CEOs) perform significantly better with increasing numbers of women on the board of directors.

Taking advantage of the longitudinal dimension of our data, we establish our main result controlling for company fixed effects (which absorb the constant unobserved heterogeneity) and for a wide array of governance and financial factors. We mitigate endogeneity problems by employing a triple-difference approach to assess the effect of executive transitions. In particular, we compare family firm profitability before and after transitions from male to female CEOs with a control sample of transitions from male to male CEOs (Huang and Kisgen 2013); we then analyze how the resulting difference in profitability is shaped by the presence of female directors. Transition analyses can be problematic if past profitability or omitted factors affect the choice of the incoming CEO. We address this problem in two ways. First, we use a matching procedure that enables better isolation of the change in profitability after CEO succession by comparing male–male and male–female transitions in firms that are most alike in terms of pre-succession characteristics. Second, we employ instrumental variables based on (a) the gender composition of the pool of potential family heirs, and (b)

geographic variations in gender stereotypes. Our results confirm that the profitability effect of female CEO transitions is increasing in the proportion of female directors on the board.

It is well established that CEOs exert a significant influence on corporate policies and performance (Bennedsen, Pérez-González and Wolfenzon 2012, 2011; Bertrand and Schoar 2003). Moreover, the board of directors provides important services to companies. As summarized by Adams, Hermalin and Weisbach (2010), directors hire, evaluate, and fire executives as well as contribute to corporate strategy—for example, by providing expert advice. Executive and governance officers clearly do not work in isolation, and the literature has underscored the importance of interactions between CEO and organizational characteristics. For instance, Adams and Ferreira (2007) argue that the flow of firm-specific information from CEOs to directors affects the quality of advice and monitoring that boards provide to CEOs. Drawing on social psychology studies, Westphal and Zajac (1995) suggest that the effectiveness of interactions between directors and CEOs can be shaped by similarities in demographic characteristics such as age and education. But why, in particular, should interactions between female CEOs and directors affect corporate outcomes?

On the one hand, women are often considered to be their own worst enemies—a dynamic that may be exacerbated in ruthless corporate environments. On the other hand, female interactions have been shown to benefit women in a variety of contexts. A field study conducted by Greig and Bohnet (2009) in a developing country shows that women cooperate more with female groups than with mixed-gender groups. Eckel and Grossman (2001) find that "women paired with women almost never fail to reach an agreement" in an ultimatum game. Dasgupta and Asgari (2004) document that the frequency of classroom encounters with female faculty reduces the activation of gender stereotypes among female students. In a business context, the findings of Matsa and Miller (2011) suggest that female directors help other women to reach management positions.

Drawing on these insights, we posit that the presence of female directors improves the profitability of companies led by female CEOs. First, female directors may enhance the self-esteem of a female CEO in an environment, such as corporate leadership, that is typically viewed as male oriented (Koenig, Eagly, Mitchell and Ristikari 2011) and is thus capable of inducing psychic costs (Blau, Ferber and Winkler 2002) that could lead to female underperformance (Rey-Biel and Iriberri 2013). Second, a female-friendly corporate culture arising from the presence of female directors can encourage cooperation and information exchange at the top and thereby improve the quality of board advice, which in turn improves the task performance of female CEOs. In line with this view, our fixed-effects estimates

indicate that—although lone female CEOs tend to depress profitability—a large presence of female directors significantly improves the profitability of family firms that are led by a woman: the profitability effect of female CEOs working with a female-dominated board of directors is 0.9 percentage points. That being said, there is wide variation in this effect; in particular, such female interactions are less effective in geographic areas that are characterized by a relatively conservative view of women's role in society and also in large firms, where it may be more difficult for women to affect corporate policies.

Our study is closely related to recent research on female leadership. Several works analyze the characteristics of female directorship and its influence on a variety of firm outcomes (Matsa and Miller 2013; Adams and Funk 2012; Adams, Gray and Nowland 2011; Bianco, Ciavarella and Signoretti 2011; Adams and Ferreira 2009). Other research investigates the corporate policies undertaken by female CEOs (Faccio, Marchica and Mura 2012), chief financial officers (Huang and Kisgen 2013), and owners (Matsa and Miller 2012) as well as the effect of female managers on the gender pay gap (Tate and Yang 2014) and firm performance (Dezsö and Ross 2012). Yet another strand of the literature analyzes how the gender composition of teams affects group performance (Hoogendoorn, Oosterbeek and Van Praag 2013; Apesteguia, Azmat and Iriberri 2012). We differ from these studies in that, whereas they focus solely on the leader's gender and/or the gender composition of teams, we focus on the importance of female interactions involving governance and CEO positions.

Our paper also contributes to the emerging stream of studies that examine gender interactions in the field. Gagliarducci and Paserman (2012) find that female-headed municipalities in Italy are more likely to experience early termination of the legislature when the council consists only of men. Our main finding—that female CEOs perform significantly better when there are more women on the board of directors—extends to the business sector the result of Gagliarducci and Paserman that gender interactions play a critical role in the functioning of organizations. Delfgaauw, Dur, Sol and Verbeke (2013) study the effect of tournament incentives on the sales growth of Dutch stores. These authors find that tournament incentives increase growth only for those stores in which the manager and a majority of workers are of the same gender. Our work here differs from these previous papers in that we shed light on the profitability effect of gender interactions at the top of the corporate hierarchy. This focus is of particular importance because these top-level gender interactions can have a significant effect not only on corporate profitability but also on such corporate behavior as risk taking in financial and investment policies.

Finally, by assessing the corporate consequences of women in CEO and governance positions, we complement extant research on the "glass ceiling" (Matsa and Miller 2011) and gender stereotypes (Price 2012), which has focused mostly on policies regarding the appointment and compensation of women in leadership positions.

2. Data

2.1. Sources

Our empirical analysis is based on the population of medium and large family-controlled firms (hereafter family firms) in Italy (see also Miller, Minichilli and Corbetta 2013; Amore, Minichilli and Corbetta 2011). Family firms represent the most common ownership form worldwide (Faccio and Lang 2002; Porta, Lopez- de- Silanes and Shleifer 1999). Even among large publicly traded US companies, which are typically thought to have dispersed ownership, families hold large equity stakes in around a third of S&P 500 firms and half of the largest 2000 industrial firms (Anderson, Duru and Reeb 2009; Villalonga and Amit 2006; Anderson and Reeb 2003).

We follow the literature in considering as family firms private companies in which one family owns an absolute majority (i.e., at least 50%) of the shares. That level of ownership is generally needed to achieve control given that private firms are usually owned by few shareholders (Bennedsen and Wolfenzon 2000). However, as in other studies on European companies (e.g., Andres 2008), we reduce this threshold to 25% for firms listed on the Milan stock exchange while assuming de facto control due to collective action problems and/or the family's use of control-enhancing mechanisms.[30]

Each firm covered in our data set had sales exceeding €50 million in 2009, a threshold that corresponds to a typical large or medium-sized family firm in Italy. Such companies make for a more appropriate setting in which to investigate leadership and governance structures than do smaller firms, which are often characterized by simple and mainly informal governance structures. Large and medium-sized family firms constitute a major segment of the Italian economy. In our data, they account for 57.1% of all similar-sized firms—followed by local subsidiaries of multinational companies (21.3%), coalitions of owners (8.7%), cooperatives and consortia (5.8%), state-owned firms (5.1%), and companies controlled by a bank or private equity firm (2.0%). Moreover, descriptive evidence suggests that medium-sized and

[30] Because of differences in the average equity share and in stock market characteristics, the ownership thresholds used by scholars to define a US "family firm" are typically lower than those used in European countries Bar-Hillel (1973).

large Italian family firms account for approximately 20% of the country's gross domestic product and employ a similar portion of the workforce.

Our data set is assembled from two main sources. Ownership and governance data are hand-collected from official public filings obtained from the Italian Chamber of Commerce. These filings track changes over time in ownership, governance, and leadership structures. Starting with the population of companies with sales greater than the aforementioned threshold, we reconstruct the ownership structure of all firms for the period 2000–2010. Overall, we identify around 2,400 family-controlled firms per year.[31] The gender of each CEO and board member is coded via a fiscal identifier that allows us to identify several demographic characteristics. Family membership of CEOs and directors is identified by surname affinity with that of the controlling family.[32]

Accounting data come from AIDA, which is the Italian provider of the Bureau van Dijk European Databases. We merge our data sources for the period under study and then drop observations with missing values in the key explanatory variables, with negative or zero book value of assets and revenues, or in cases where there is no formal CEO.

Although our final data set is limited to one organizational form, it has a number of important advantages. First, it covers not only listed firms but also a vast majority (some 93%) of privately held companies, which are widespread in the Italian economy. Second, its longitudinal dimension and large set of firm characteristics allows us to adopt a fixed-effects methodology and control for potentially confounding factors. Third, and similarly to previous research (e.g., Bennedsen, Nielsen, Pérez-González and Wolfenzon 2007), it allows the use of family characteristics to generate exogenous variations (in organizational structure) that are useful for mitigating endogeneity concerns.

2.2. Summary statistics

In Table 4.1, Panel A presents summary statistics for the main firm characteristics. We adopt the return on assets (ROA) as our measure of operating profitability. Computed as the earnings before interest and taxes (EBIT) divided by the book value of total assets, ROA is

[31] Specifically, we started with a population of about 6,800 firms with sales of more than €50 million in 2009. After a first screening, we identified 4,221 family firms. Among those, we sampled the parent company for firms in business groups operating in the same 2-digit industry (because the boards and leadership structures of controlled companies are almost always identical to those of the parent firm); in such cases, we used consolidated financial information for the whole group. For business groups (usually holding companies) that operate in more than one 2-digit industry, we analyzed the controlled companies separately.

[32] Since this approach fails to capture some family ties—for example, marriages in which the wife does not assume her husband's surname—it should be viewed merely as a proxy for blood ties.

widely used to assess firm performance (see e.g. Bennedsen et al. 2007; Pérez-González 2006). In our sample, the average ROA is 0.054.

Leadership and governance characteristics are tabulated next. First, in line with existing evidence, we find that ownership is highly concentrated (an average of 91% of shares are held by family members). Firms in our sample have, on average, two CEOs; this reflects the possibility in Italian firms that there may be more than one CEO with the same duties and legal responsibilities—a special leadership structure that renders our results not directly comparable to studies of firms in countries (e.g., the United States and the United Kingdom) where the single-CEO model is more typical. In order to reduce the number of dropped observations, we conduct our main analysis on the full sample formed by both firm types; however, we establish that our results hold also on the subsample of firms with a single CEO. In our data, CEOs are members of the controlling family in 74% of the cases. The typical firm in our sample has 5 board members, although this number ranges from 2 to 23. Directors are 54 years old on average, and 65% of them are part of the controlling family.

Panel B of the table shows that about 20% of the 10,000+ observations report at least one female CEO. Overall, nearly 11% of CEOs are female; this portion is 9% in the subsample of firms with only one CEO (approximately 45% of the observations). On average, the portion of female directors is 17% (11% if female CEOs are excluded from the count).[33] About 45% of observations have no female directors, 32% have one female director, 15% have two female directors, and the remaining 8% have three or more female directors. There are 4,140 observations involving female directors in addition to female CEOs (if any), of which 53% involve one female director; 32%, two female directors; and 16%, three or more female directors. As expected, the average percentage of female directors is higher in the subsample of firms with at least one female CEO (38%, versus 12% in companies without a female CEO), and in 32% of those cases the board consists of at least 50% female members.

3. Empirical analysis

We begin this section by presenting our main results. We then show that our findings are not affected by other firm characteristics. Finally, we perform a number of robustness checks.

[33] In the United States, women held 14.8% of the Fortune 500 board seats in 2007 (see Adams and Ferreira 2009 and the references therein). Among Italian listed companies, Bianco et al. (2011) find that women accounted for 6.8% of all board seats at the end of 2010. This is a lower percentage than that reported here, which is likely a consequence of their data set including only listed firms; such firms are typically among the largest, and (as documented by Bianco et al.) female directors are more common in smaller companies.

3.1. Preliminary evidence

Our goal is to estimate how the interaction of women in CEO and governance positions affects firm profitability. Toward this end, we estimate the following model separately for firms with at least one female CEO and firms without a female CEO:

$$\text{ROA}_{it} = \beta_0 + \beta_1(\text{Female directors}_{it}) + \delta_t + \alpha_j + \text{X}_{it-1}\eta + \varepsilon_{it} \quad (1)$$

Here ROA_{it} is our measure of firm profitability for firm *i* at time *t*. The main explanatory variable, $\text{Female directors}_{it}$, is the ratio of female directors to all directors. The model includes year dummies δt to control for common shocks as well as 2-digit industry dummies αj to control for sectoral differences in profitability. The vector Xit−1 captures firm characteristics. These include the (log of) firm age and total assets, which control for a firm's different stages of development; the ratio of cash holdings to total assets, which controls for differences in liquidity; and the ratio of capital expenditures to total assets, which controls for differences in investment. Finally, we control for the presence of the family in governance, executive, and ownership positions; this is motivated by previous studies showing that family involvement at different levels has strong implications for firm performance (see e.g. Bennedsen et al. 2007; Pérez-González 2006; Villalonga and Amit 2006; Anderson and Reeb 2003). To avoid simultaneity problems, all firm-level controls are lagged by one year.

Results reported in Table 4.2 show that, in the absence of female CEOs, the percentage of female directors has no effect on firm profitability (column [1]). When there is at least one female CEO, however, the effect is positive and significant at the 5% level (column [2]).[34]

Next, we estimate the following interaction model on the full sample:

$$\begin{aligned}\text{ROA}_{it} = {} & \beta_0 + \beta_1(\text{Female CEO presence}_{it}) + \beta_2(\text{Female directors}_{it}) + \\ & + \beta_3(\text{Female CEO presence}_{it} \times \text{Female directors}_{it}) + \delta_t + \alpha_j + \text{X}_{it-1}\eta + \varepsilon_{it}\end{aligned} \quad (2)$$

Here Female CEO presence it measures the presence of women in CEO positions. To ensure enough variation, we operationalize that presence using an indicator variable set equal to 1 (resp., to 0) if, at time t, any (resp., none) of the CEOs of firm i is female. The coefficient

[34] Taken together, these results suggest that the profitability improvement is not simply driven by a reduction in overall diversity and instead is specific to female interactions.

β1 measures the effect (on profitability) of female CEOs as the fraction of female directors tends to zero; β2 measures the effect of female directors on the profitability of companies without female CEOs; and β3 measures how the profitability of companies with at least one female CEOs varies with the fraction of female directors. Thus, for any given value of Female $directors_{it}$, the effect on firm profitability of female CEOs is given by β1 + β3(Female $directors_{it}$).

Results reported in column [3] of Table 4.2 indicate that, whereas the coefficient for the female CEO dummy is not statistically different from zero, the interaction term is positive and significant at the 5% level. We can interpret the economic magnitude of female CEOs evaluated when at least half of the board members are women: in this case, the effect is 0.000 + 0.018(0.501) = 0.009, or 0.9 percentage points (statistically significant at the 1% level).

3.2. Fixed-effects estimates

In Section 3.1, we offered descriptive evidence that the interaction between women in CEO and board positions improves firm profitability. We now provide more accurate estimates by augmenting equation (2) with firm fixed effects αi in order to reduce unobservable time-invariant heterogeneity at the firm level. And because many of the controls adopted in Table 4.2 could be endogenous, here we adopt a more parsimonious vector Xit−1 that only includes the one-year lagged logarithms of firm age and total assets in addition to industry trends.[35]

Columns [1] and [2] of Table 4.3 report that the profitability effect of female CEOs and of female directors is negative, although only the coefficient for female directors is statistically significant at the 10% level. In column [3], we include the interaction between female CEOs and female directors. The results indicate that female directors have a significantly negative effect on the profitability of companies without a female CEO. Moreover, the results indicate that as the fraction of female directors tends to zero, female CEOs have a significantly negative effect on profitability; the coefficient is -0.016— that is, 1.6 percentage points. Thus, when women are present only as CEO(s), the result is reduced firm profitability (compared with the case when no women CEOs are present). Yet the coefficient for the interaction term is positive and statistically significant at the 1% level; in other words, the underperformance

[35] In our main specification, we control for industry trends by interacting year and 1-digit industry dummies; however, the results are robust to excluding this control. Our results hold also if we control for industry trends in different ways—for example, by interacting year and 2-digit industry dummies or by subtracting from firm ROA the annual median or mean industry ROA (computed at the 1- or 2-digit level).

of companies led by female CEOs is less when the fraction of female directors increases. The economic magnitude of female CEOs when at least half the board members are women is −0.016 + 0.050(0.501) = 0.009—that is, 0.9 percentage points (significant at the 5% level).[36]

In the regressions described so far, Female CEO presence$_{it}$ is operationalized as an indicator variable set equal to 1 if any of the family firm's CEOs is female (and set to 0 otherwise); however, in column [4] of Table 4.3 we show that our results are robust to replacing this dummy with a continuous measure corresponding to the fraction of female CEOs. These results, which are similar to those reported in column [3], reveal that female CEOs have a negative effect on profitability when the fraction of female directors tends to zero (although the p-value is only 0.19). Likewise, female directors have a significantly negative effect on profitability as the fraction of female CEOs tends to zero. However, the interaction term is positive and statistically significant at the 1% level.

Overall, our results indicate that, while lone female leadership is detrimental to family firm performance, this negative effect is significantly mitigated by the interaction between women in CEO and board positions.[37]

3.2.1. Confounding factors

One concern with our estimates is that of bias stemming from omitted factors. Given that female interactions do not occur randomly, they may be correlated with unobserved characteristics that affect firm profitability and thus bias our estimates. An advantage of our specification is that, by exploiting the data's longitudinal dimension, we can effectively control for firm-level unobserved heterogeneity that is invariant over time. We further mitigate omitted factor concerns in two ways. First, we augment our fixed-effects specification with controls for a number of governance and financial characteristics. Second, we show that our estimates do not mirror the interaction of other demographic characteristics.

In Panel A of Table 4.4, we sequentially control for the family involvement variables used in Table 4.2. Specifically, in column [1] we include the lagged fraction of family CEOs; in column [2], the fraction of family directors; and in column [3], the fraction of family owners. In column [4] we control for the presence of female owners in order to mitigate concerns

[36] In unreported regressions, we replace the continuous fraction of female directors with dummies indicating whether that fraction is low, medium or high (corresponding to the lowest, medium, or highest tercile of the empirical distribution in female-led firms). Our findings indicate that female CEOs have a significantly negative profitability effect when there is a low fraction of female directors. This effect, which is close to zero for a medium fraction of female directors, becomes positive and significant for a high fraction of female directors. These results suggest that, in order for female interactions to induce a significant effect on profitability, the fraction of female directors has to be large.

[37] It should be stressed here that we do not refer to the *total* effect of female interactions on profitability (i.e. the sum of all coefficients), but rather to how the effect of female CEOs varies with the presence of female directors.

that—since female leadership may be correlated with female ownership—the effect on profitability may arise from female owners and not from the interaction between female CEOs and female directors. Finally, in column [5] we control for the fraction of nonexecutive directors as a proxy for the quality of board oversight. This panel confirms that none of these additional controls affects the statistical and economic significance of the interaction term. In unreported analyses, we establish that our findings remain unchanged when these controls are all included in a single regression and also when we include additional controls such the number of board members.

Although our results indicate that female interactions affect firm profitability, it could be that we are capturing not a gender-specific effect but rather a generic effect stemming from demographic similarities. However, we can show that no profitability effect arises from interactions of another demographic characteristic such as age. This is done, in Panel B of Table 4.4, by replacing our gender variables with the average age of CEOs and directors and with the interaction between these variables.[38]

3.2.2. Robustness

First, we confirm our main result by excluding CEO seats on the board of directors from the calculation of Female directors$_{it}$.[39] As shown in column [1] of Table 4.5, using this alternative measure of female directors results in an interaction term that is significant both statistically and economically.

In Section 2, we mentioned that it is common for Italian firms to have multiple CEOs with similar duties. Column [2] restricts the analysis to firms that have but a single CEO. The values reported in this column indicate that our results are not driven by the number of a firm's CEOs.

Next we establish that our findings hold irrespective of how firm profitability is measured. In column [3], we change the dependent variable from ROA (i.e., EBIT divided by total assets) to EBITDA (earnings before interest, taxes, depreciation, and amortization) divided by total assets; in column [4], we use the ratio of net income to total assets. In untabulated regressions we adopt still other profitability measures: return on sales and (log of) sales to assets; the latter measure captures the efficiency with which the company uses its assets to

[38] This result remains insignificant when we use the logarithms of CEOs' and directors' ages.

[39] For example, a firm with one female CEO and a board consisting of one female and one male director has a fraction of female directors equal to 50%. But if the CEO sits on the board then there is no true female interaction because the same woman is serving both as CEO and director. We accommodate that possibility by excluding the number of female CEO board positions when calculating the fraction of female directors. In the example given here, that fraction would then be 0% rather than 50%.

generate revenues. Finally, our results hold when the dependent variable is a measure of total factor productivity, computed as the residuals obtained from estimating, via ordinary least squares (OLS), a log-linear Cobb–Douglas production function in which the dependent variable is the logarithm of sales and the explanatory variables are the logs of employment and capital together with industry and year dummies.

One criticism of using firm fixed-effects in this context is that the identification exploits only within-company changes in the presence of women among executives and directors, and such changes may well be limited in number. In column [5], we show that, even for a specification based on random effects, female interactions have a positive effect on profitability (with a significance level that approaches 5%). Column [6] controls for the interaction between regional and year dummies and so accommodates any effect of time-varying geographic factors. We next mitigate concerns that our results are driven by particularly influential observations. In column [7], we trim 1% of the observations on both the left and right tails of the ROA distribution (in unreported regressions, we also check that our findings are robust to trimming 2.5% and 5% of these observations). Our results are robust as well to taking transformations of the dependent variable—for example, the natural logarithm of 1 + ROA. Column [8] details the results when an alternative computation of the standard errors, one based on firm clustering, is used; this procedure is valuable when we seek to account for the potential correlation of residuals by firms (in unreported tests we also cluster residuals by firm and by year). Next, we augment our specification with the squared fraction of female directors in case the profitability effect of female interactions is a nonlinear one; column [9] shows that, although the quadratic term is itself weakly significant (unreported), the main interaction term remains positive and significant at the 5% level.

Another concern with our analysis is that successful firms may be more likely to conform to gender diversity norms (Dezsö and Ross 2012) and thus appoint female executives. To address the likelihood of such reverse causality, we follow Dezso and Ross in augmenting the specification with firm profitability lagged by one year. We can thus establish the effect of gender interactions after controlling for differences in past profitability across firms. It is well known that fixed-effects estimates are inconsistent when lags of the dependent variable are included as explanatory variables—a problem that arises by construction because lags are correlated with the residuals. We therefore adopt the dynamic generalized method of moments developed by Arellano and Bond (1991). This method first differentiates the model (in order to eliminate fixed effects) and then performs an instrumental variables regression of the resulting model using lags of the explanatory variables as instruments. As shown in column

[10], which reports one-step heteroskedasticity robust estimates, the interaction term remains statistically significant at the 10% level as well as economically close to our baseline estimates.[40] We further assess the causal direction of our results by estimating a non-contemporaneous specification. Specifically, we re-estimate our baseline regression model using one-year lags of the main explanatory variables and interaction term. The results (untabulated) of this estimation indicate that past female interactions have a positive and significant effect on current firm profitability ($\beta 3 = 0.045$, $p = 0.005$).

4. Evidence from executive transitions

Although our findings are robust to several checks, the concern of endogeneity remains. In this section we estimate a triple-difference model of executive transitions. An important advantage of this approach is that, by absorbing trends and other factors that are unobserved but common to all firms, it better isolates the profitability effect of female interactions. First, we present estimates obtained on the full sample of CEO transitions. Second, we present a number of robustness checks. Third, we report results based on propensity score matching and instrumental variables.

4.1. Main results

We start by identifying the subsample of approximately 3,300 observations involving family firms that experienced CEO successions during the period under study. Our treatment group is formed by CEO transitions in which the departing CEO is male and the incoming CEO is female (i.e., male–female transitions); the control group consists of male–male transitions.[41] The treatment group accounts for 13% of our observations of CEO succession. Comparing the profitability effect of these male–female transitions with that of the control group is a difference-in-differences (DD) approach similar to that employed by Huang and Kisgen (2013). One condition for this approach to be valid is that firms in the treatment and control group exhibit similar pre-transition trends relative to the outcome of interest. We establish the satisfaction of this condition in the study undertaken here by comparing the mean profitability of treatment and control groups prior to transition. Two-sample t-tests confirm that both our treatment and control groups are similar in terms of pre-transition ROA (difference = −0.0001, $p = 0.985$). This lack of a significant difference should mitigate

[40] When employing many instruments, the Arellano–Bond method can suffer from poor performance of asymptotic results. Therefore, in unreported regressions we check that our results are robust to restricting the instruments to the first available lag.

[41] We exclude the few cases of female–male transitions.

concerns about reverse causality, a topic that is addressed more thoroughly in Sections 4.2 and 4.3.

Next we adopt the fraction of female directors as an additional (continuous) treatment. The resulting interaction constitutes a triple-difference (DDD) estimator, which assesses how the profitability effect of male–female versus male–male CEO transitions varies depending on the presence of female directors. We estimate the following model:

$$\begin{aligned} \mathrm{ROA}_{it} &= \beta_0 + \alpha_i + \delta_t + \mathrm{X}_{it-1}\eta + \beta_1(\mathrm{Post}_{it} \times \mathrm{Female}_i) + \beta_2(\mathrm{Post}_{it} \times \mathrm{Female}_i \times \mathrm{Female\ directors}_{it}) + \\ &+ \beta_3(\mathrm{Post}_{it} \times \mathrm{Female\ directors}_{it}) + \beta_4(\mathrm{Female}_i \times \mathrm{Female\ directors}_{it}) + \beta_5(\mathrm{Female\ directors}_{it}) + \varepsilon_{it} \end{aligned} \quad (3)$$

Here Post_{it} is an indicator variable set equal to 1 only for the years after transition, and Female_i is an indicator set equal to 1 if the company experiences a male–female transition. The coefficient β1 for the $\mathrm{Post}_{it} \times \mathrm{Female}_i$ interactions estimates the profitability effect of male–female transitions as compared with male–male transitions. Our estimate of interest is given by β2—that is, the coefficient for the triple interaction $\mathrm{Post}_{it} \times \mathrm{Female}_i \times \mathrm{Female\ directors}_{it}$—which measures how the profitability effect of male–female versus male–male transitions varies depending on the fraction of female directors. The dependent variable and controls are similar to those in the regressions reported by Table 4.3.

The results are reported in Table 4.6. First, we find that male–female CEO transitions have a positive effect on firm profitability that is significant at the 10% level (columns [1] and [2]). Second, male–female transitions have an economically and statistically significant effect on profitability as the fraction of female directors increases (columns [3] and [4]). These results support our previous finding that female interactions improve the profitability of family firms.

4.2. Robustness

With regard to sample composition, we show that our results are robust to excluding the few firms that undertake more than one succession during the period considered (column [5] of Table 4.6) and to including only single-CEO firms (column [6]). In the latter case, the triple-interaction term is positive and close to our main estimate in column [4], although the statistical significance is not achieved at conventional levels (owing, perhaps, to the small sample size). In column [7], we check that our results hold when female CEOs are excluded from the count of female directors.

As already mentioned, firms undergoing male–male versus male–female leadership transitions do not differ in terms of profitability prior to the transition. This sameness is

crucial because it indicates that reverse causality is highly improbable in our setting. However, we further mitigate that concern by using "placebo" successions two years prior to the actual succession year. If the coefficient for such placebo succession turned out to be statistically significant, then our previous estimates might merely reflect diverging pre-transition trends in the treatment and control groups rather than a treatment effect. In column [8], we show that the triple-interaction term is neither statistically nor economically significant. This result further validates our approach and confirms that our findings are indeed driven by executive transitions and not by diverging trends.

In unreported regressions, we also check that the triple-interaction term remains statistically and economically significant when residuals are clustered by firms, and also when we trim 1% of the observations in both tails of the ROA distribution to reduce the effect of outliers. Finally, we check that our main finding remains significant when either the variable Female directors$_{it}$ or the interaction term Post$_{it}$ × Female directors$_{it}$ is excluded. In fact, the result holds even when we include only the terms Post$_{it}$ × Female$_{i}$ × Female directors$_{it}$ and Post$_{it}$ × Female$_{i}$.

4.3. Propensity score matching and 2SLS estimates

Propensity score matching ensures that firms experiencing male–male successions and those experiencing male–female successions are as similar as possible with respect to an array of observable characteristics. Specifically, we match each treatment firm, i.e., a firm in which a male–female transition occurs, with a firm in which a male–male transition occurs. The matching is performed—without replacement—on the basis of a propensity score estimated using a probit regression in which the dependent variable is the succession treatment dummy and the controls are the following variables: pre-succession averages of firm age, firm size, ROA, and fraction of female directors; and dummies for year, region and industry. Note that employing pre-succession ROA ensures that the compared firms exhibited similar profitability prior to the CEO succession.

Results are reported in columns [1] and [2] of Table 4.7 The sample size for these regressions is much smaller because (1) for each firm that undergoes a male–female transition, we consider the most similar firm undergoing a male–male transition while discarding observations for all other firms, and (2) male–male transitions are much more frequent than male–female transitions. The coefficient reported in column [2] for the triple-interaction term is economically relevant and statistically significant at the 5% level.

Propensity score matching addresses problems of reverse causality, but it may still be that unobservable variables are correlated with the incoming CEO's gender. We deal with this concern by using instrumental variables. Recent works suggest that family firms offer a unique setting in which to address endogeneity: family characteristics have a strong influence on the organization of the family firm, yet they are unlikely to affect corporate outcomes directly.[42] In exploiting family-specific characteristics, our first instrument is an indicator variable set equal to 1 if the CEO has a female child (and set to 0 otherwise).[43] This variable is potentially a good instrument because—though it is (nearly) randomly determined and unlikely to be affected by firm profitability at the time of succession—it can be a significant predictor of the incoming CEO's gender: when there is a female child within the controlling owner's family, it is (expected to be) more likely that a female CEO will be appointed. Indeed, results from a probit regression indicate that the likelihood of appointing a female successor in this case is 9% higher ($p = 0.000$) after we control for firm size, firm age, and year and industry dummies. As a second instrument, we exploit a unique feature of the Italian context. Namely, there are marked differences across regions in terms of gender roles. Southern regions of the country are characterized by a traditional view of the family in which the woman is the homemaker and the man is the breadwinner. This belief, which is exhibited less often in central Italy, is even less prevalent in northern regions of the country. We therefore employ an indicator variable set equal to 1 if the firm is located in northern Italy (and to 0 otherwise). Although geographic variations in gender roles are likely to affect an incoming CEO's gender (in fact, firms located in northern regions are approximately 9% more likely to appoint a female CEO), it is unlikely that such variations have a direct effect on changes in profitability subsequent to a CEO succession.

The main endogenous variable in our framework is the indicator variable for incoming female CEO. Thus, equation (3) contains three endogenous variables: $Post_{it} \times Female_i$, $Post_{it} \times Female_i \times Female\ directors_{it}$, and $Female_i \times Female\ directors_{it}$. We estimate a two-stage least-squares (2SLS) model in which OLS are first used to regress each endogenous variable on the (similarly operationalized) instruments and the other controls, after which firm ROA is regressed on the predicted values of the endogenous variables together with the same

[42] Bennedsen, Kongsted and Nielsen (2008) use family size as an instrument when estimating the effect of board size on performance; Bennedsen et al. (2007) use the gender of a CEO's firstborn child to establish the performance effect of family CEO successions.

[43] Data on potential heirs were hand-collected while considering the family nucleus of the main physical shareholder at the year of succession. In the presence of controlling legal entities, we identified the main shareholder among the physical shareholders of the controlling holding. We collected information on the gender of the main owner's children from a variety of sources, including company websites, press releases, family blogs, administrative documents, newspapers, and obituaries. Unfortunately, this information is available for only about 60% of the observations used in columns [1]–[4] of Table 6.

first-stage controls. The results of this procedure, which are reported in columns [3] and [4] of Table 4.7, indicate that the triple-interaction term is statistically significant (at 8%) and economically larger than the OLS estimates reported in Table 4.7.[44]

5. Variations

5.1. Gender stereotypes

So far we have established that, on average, family firms led by female CEOs perform significantly better as the fraction of female directors increases. We now explore how gender stereotypes shape this result. On the one hand, female interactions could be particularly valuable in the presence of stereotypes against women by helping to alleviate the negative effect of such stereotypes on CEO task performance. On the other hand, female productivity may be lower in workplaces characterized by stereotypes against women. In Table 4.8 we adopt two empirical tests to disentangle these hypotheses.

First, we present results separately for firms located in the southern, central, and northern regions of Italy; as discussed in Section 4.3, these regions traditionally differ in terms of gender roles. Our results indicate that lone female CEOs have a large negative impact on profitability in southern regions (column [1]) and have a negative but relatively smaller effect in central and northern regions (columns [2] and [3], respectively). The interaction term is positive and significant for companies operating in southern regions, and it is positive but relatively smaller in northern regions. Assessing the effect of female CEOs evaluated for firms whose boards are predominantly female, we find a negative effect in southern regions (−0.023, p = 0.073), an insignificant effect in central regions (0.009, p = 0.294), and positive and significant effect in northern regions (0.012, p = 0.025).[45]

Second, we follow the approach in Guiso and Rustichini (2012) that is based on data from the World Value Survey (WVS). The WVS is a collection of comprehensive individual-level surveys carried out in several countries around the world,[46] and it contains many questions on

[44] In order to assess the change in economic magnitude, we estimate an OLS regression on the subsample of firms used for the 2SLS regression. We find that the interaction term is nearly 3 times larger when estimated with 2SLS.

[45] We use the baseline specification, but our results are robust to excluding overlaps between female directors and CEOs and also to restricting the analysis to single-CEO firms.

[46] The countries covered by the WVS increased from 20 in the first wave (1981–1984, 25,000 respondents) to 54 in the last wave (2005–2008, 77,000 respondents). For Italy, the 2005 wave was addressed to around 1,000 statistical units selected from the population of age 18–74. Overall, 80 municipalities were selected from all 20 Italian regions while accounting for the distribution of population by region and the degree of urbanization (small, medium, medium-large and large municipalities). The WVS then created a proportional sample stratified by region, age, and gender. The WVS website (http://www.worldvaluessurvey.org) documents that there are no significant differences between the real and theoretical sample in terms of age, gender, demographic size of municipalities, and geographic areas. In terms of geographic areas, for instance, 45.65% of WVS statistical units are in northern regions of Italy (versus 45.80% in the theoretical sample), 19.17% (versus 19.35%) are in central regions, and 35.18% (versus 34.79%) are in southern regions.

human values related to religion, political participation, attitudes toward women and minorities, and family values; the survey also reports standard demographic characteristics of the respondents. For the purpose of our analysis, we extract data from the 2005 wave and use responses to four questions concerning the role of women in the society. Respondents were asked to indicate the extent of their agreement—on a 4-point scale ranging from "strongly agree" to "strongly disagree"—to the following statements: (1) "Being a housewife is just as fulfilling as working for a pay"; (2) "On the whole, men make better political leaders than women do"; (3) "A university education is more important for a boy than for a girl"; (4) "On the whole, men make better business executives than women do". We used the survey to devise two indexes of female emancipation: one based on the first principal component of the responses to statements (1)–(4) and another based solely on responses to question (4). Finally, we compute the regional averages of these two indexes based on the respondents' geographic location and then assign the relevant value to each firm in our data set as a function of where its headquarters is located.

Columns [4]–[6] report the results when observations are sorted by terciles of the first index.[47] Although the negative coefficient for lone female CEOs is slightly more negative when female emancipation is high, in this case the interaction coefficient is both positive and much larger. A similar result is obtained when we use the second index of female emancipation; see columns [7]–[9].

5.2. Firm size, director identity, and economic conditions

In this section, we show how our main finding varies in terms of firm characteristics and business cycles. We start by focusing on firm size, which may shape our results in two opposing ways. On the one hand, it may be easier for women to leave a personal imprint on corporate policies when the firm is small. On the other hand, large companies might be less subject to gender stereotypes and so may enact policies that favor the careers of women. In Table 4.9, columns [1] and [2] present estimates separately for small and large firms—that is, for firms below and above (respectively) the median with respect to size of assets. The reported values reveal that the negative effect (on profitability) of lone female CEOs is slightly more pronounced for large firms; also, the interaction term is positive and significant (albeit relatively smaller) for large firms. Thus, the effect of female CEOs heading firms with female-dominated boards is 10% significant and greater in small firms than in large ones.

[47] Our findings are substantially unchanged if the sample is split at the median instead of into terciles.

Next we test how the family affiliation of female directors affects our result. Toward this end, we measure the fractions of family-affiliated and family-unaffiliated female directors[48], and then interact the two resulting variables with our indicator variable for female CEO. The results, which are reported in columns [3] and [4], indicate that lone female CEOs in each case have a significantly negative effect on firm profitability.[49] The interaction term is positive and significant for both related and unrelated female directors, but the effect is much larger in the latter case. When interpreting this result, one should bear in mind that female CEOs are almost always appointed from within the controlling family; the implication is that column [3] is de facto analyzing female interactions between family members while column [4] is analyzing female interactions between family-affiliated CEOs and outside directors. In other words, the female interactions that generate the most improvement in profitability are those between a related female CEO and unrelated female directors. A possible explanation for this finding is that, because family members typically know each other and share the same values and objectives, gender is less likely to influence CEO–board interactions. This result could also be interpreted from a governance perspective: the value of a woman in a governance position (as documented by Adams and Ferreira 2009) may well be amplified when she is unrelated to the controlling family, given that outside directors mitigate expropriation by minority stakeholders (Anderson and Reeb 2004). Moreover, the companies we analyze are in particular need of effective directors because family ties are often associated with worse management practices (Bloom and Van Reenen 2007). So improvements in firm profitability that are due to better governance should be greater when female interactions are between CEOs and family-unaffiliated directors—and, indeed, this is what we find. Overall, these special circumstances argue for caution when seeking to extend our results to nonfamily firms.

Finally, we explore the effect of variations due to business cycles. In particular, we estimate our main specification for the period ending in 2007 (i.e., for "normal" economic times) and also for the "financial crisis" period 2008–2010. One reason why the profitability effect of female interactions could differ in crisis versus normal economic circumstances is risk aversion. If women are more risk averse than men, then companies with women in CEO and board positions may be better positioned financially in the aftermath of an economic crisis and thus better able to respond and adapt to that circumstance. In contrast, if women are

[48] Around 80% of all female directors are family affiliated.

[49] Results also show that only nonfamily female directors have a negative and significant effect on profitability—perhaps because they allocate too many resources to governance issues when the CEO is male.

less risk averse than men (or if there are no gender differences in risk taking) then the profitability effect of female interactions during the 2008–2010 period should be negative (or, respectively, merely negligible). Unfortunately, the evidence on gender differences in risk aversion is ambiguous. There is some research indicating that women are the more risk-averse gender (for a recent contribution, see Charness and Gneezy 2012); yet Adams and Funk (2012) find that female directors take more risk than do their male counterparts. According to our results, the female-related stand-alone and interaction terms are significant in normal times (column [5]) but not in times of financial crisis (column [6]).

6. Discussion and conclusion

Motivated by the introduction of policies that aim to increase the presence of women in political and corporate positions, a growing literature has explored gender differences in economic behavior and their implications for organizations.

Experimental research in the field of economics has stressed the role of environmental factors in explaining female competitiveness. For instance, Gneezy, Leonard and List (2009) find that women compete more than men in a matrilineal society but that the opposite is true in a patriarchal society. Booth and Nolen (2012) show that girls from single-gender schools are more competitive than girls from coeducational schools, and other researchers (Rey-Biel and Iriberri 2013; Shurchkov 2012) argue that women tend to underperform on tasks that are perceived to be male oriented. A number of authors in the field of finance have investigated the effect of female involvement in various corporate positions (e.g., Huang and Kisgen 2013; Dezsö and Ross 2012; Faccio et al. 2012; Adams and Ferreira 2009).

Bridging these two strands of research, in this paper we explore how the interactions between women in governance and CEO positions affect corporate profitability. Our working hypothesis is that, though lone female CEOs may underperform because of the psychic costs (Blau et al. 2002) induced by a company's pervasive male-oriented context (Koenig et al. 2011), their performance may improve in response to interaction with women in the board of directors. In particular, female CEOs may feel less inhibited when operating with female peers in governance positions. Such interactions between female CEOs and female directors may also serve to reduce the risk of communication breakdowns, improve cooperation, and facilitate information exchange—effects that should result in higher-quality board performance and thus in more efficient managerial decision making.

Using a panel of family firms from Italy, we find that an increasing fraction of women on the board of directors significantly improves the profitability of firms led by female CEOs.

One challenge to a causal interpretation of our findings is the nonrandom occurrence of female interactions. We exploit the longitudinal nature of our data to control for unobserved firm heterogeneity and several other possible sources of bias, including reverse causality. We further mitigate endogeneity concerns by taking a triple-differences approach to the analysis of executive transitions. Finally, we identify several factors that reduce the positive effect (on profits) of interactions between female CEOs and female directors; in particular, such reduction is evident when the firm is located in geographic areas characterized by a conservative view of women's role in society, when the firm is large, and in times of financial crisis.

Our results have several implications. Existing evidence suggests that women are more sensitive than men to context (Croson and Gneezy 2009). By showing that the interactions with a board in which women are largely represented mitigate the observed underperformance of lone female CEOs, our results augment recent discussions (e.g., Dezsö and Ross 2012) of the contexts under which female leadership can be more effective. Our findings also suggest that the criteria by which women gain seats on the board may shape the performance effect. That is, the female interactions that are most beneficial to firm profitability are those involving nonfamily female directors, who are more likely to have been selected via a meritocratic process. This notion is reinforced by our finding that female interactions are less effective in areas with strong traditional family values, where companies more likely appoint executives on the basis of family ties. Our results thus complement those of Ahern and Dittmar (2012), who document a decline in the market value of firms subsequent to the passage in Norway of gender quotas for boards (a finding that may reflect the tendency of those companies to appoint relatively inexperienced female directors).

Our results are relevant also to the ongoing debate over the governance of private unlisted firms. Although traditionally ignored by policy makers, unlisted companies are now increasingly affected by the adoption of "codes of good governance".[50] A recent example is the Buysse Code in Belgium, which offers detailed suggestions on how private firms should be governed in order to secure their survival and long-term success. We believe that similar codes could be used to promote female representation on boards of directors; even in the absence of specific gender quotas, such representation can yield female complementarities that could improve firm performance.

[50] See "Corporate Governance Guidance and Principles for Unlisted Companies in Europe", a report prepared by the European Confederation of Directors' Association (ecoDa), and the discussion "Good Governance in Unlisted Firms" by Van den Berghe (2010).

7. Bibliography

Adams R., Gray S. and Nowland J. (2011) "Does gender matter in the boardroom? Evidence from the market reaction to mandatory new director announcements".

Adams R.B. and Ferreira D. (2007) "A theory of friendly boards", *The Journal of Finance* 62, 217-250.

Adams R.B. and Ferreira D. (2009) "Women in the boardroom and their impact on governance and performance", *Journal of financial economics* 94, 291-309.

Adams R.B. and Funk P. (2012) "Beyond the glass ceiling: Does gender matter?", *Management Science* 58, 219-235.

Adams R.B., Hermalin B.E. and Weisbach M.S. (2010) "The Role of Boards of Directors in Corporate Governance: A Conceptual Framework and Survey", *Journal of Economic Literature* 48, 58-107.

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

Amore M.D., Minichilli A. and Corbetta G. (2011) "How do managerial successions shape corporate financial policies in family firms?", *Journal of Corporate Finance* 17, 1016-1027.

Anderson R.C., Duru A. and Reeb D.M. (2009) "Founders, heirs, and corporate opacity in the United States", *Journal of financial economics* 92, 205-222.

Anderson R.C. and Reeb D.M. (2003) "Founding- family ownership and firm performance: evidence from the S&P 500", *The Journal of Finance* 58, 1301-1327.

Anderson R.C. and Reeb D.M. (2004) "Board composition: Balancing family influence in S&P 500 firms", *Administrative Science Quarterly* 49, 209-237.

Andres C. (2008) "Large shareholders and firm performance—An empirical examination of founding-family ownership", *Journal of Corporate Finance* 14, 431-445.

Apesteguia J., Azmat G. and Iriberri N. (2012) "The impact of gender composition on team performance and decision making: Evidence from the field", *Management Science* 58, 78-93.

Arellano M. and Bond S. (1991) "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *The Review of Economic Studies* 58, 277-297.

Bar-Hillel M. (1973) "On the subjective probability of compound events", *Organizational behavior and human performance* 9, 396-406.

Bennedsen M., Kongsted H.C. and Nielsen K.M. (2008) "The causal effect of board size in the performance of small and medium-sized firms", *Journal of Banking & Finance* 32, 1098-1109.

Bennedsen M., Nielsen K.M., Pérez-González F. and Wolfenzon D. (2007) "Inside the family firm: The role of families in succession decisions and performance", *The Quarterly Journal of Economics* 122, 647-691.

Bennedsen M., Pérez-González F. and Wolfenzon D. (2011) "Do CEOs Matter?".

Bennedsen M., Pérez-González F. and Wolfenzon D. (2012). Estimating the value of the boss: Evidence from CEO hospitalization events.

Bennedsen M. and Wolfenzon D. (2000) "The balance of power in closely held corporations", *Journal of financial economics* 58, 113-139.

Bertrand M. and Hallock K.F. (2001) "Gender Gap in Top Corporate Jobs, The", *Indus. & Lab. Rel. Rev.* 55, 3.

Bertrand M. and Schoar A. (2003) "Managing with style: The effect of managers on firm policies", *The Quarterly Journal of Economics* 118, 1169-1208.
Bianco M., Ciavarella A. and Signoretti R. (2011) "Women on boards in Italy", *Rome: Consob*.
Blau F.D., Ferber M.A. and Winkler A.E. (2002). *The economics of women, men, and work* (Vol. 4): Prentice Hall Upper Saddle River, NJ.
Bloom N. and Van Reenen J. (2007) "Measuring and explaining management practices across firms and countries", *The Quarterly Journal of Economics* 122, 1351-1408.
Booth A. and Nolen P. (2012) "Choosing to compete: How different are girls and boys?", *Journal of Economic Behavior & Organization* 81, 542-555.
Charness G. and Gneezy U. (2012) "Strong evidence for gender differences in risk taking", *Journal of Economic Behavior & Organization* 83, 50-58.
Croson R. and Gneezy U. (2009) "Gender differences in preferences", *Journal of Economic Literature*, 448-474.
Dasgupta N. and Asgari S. (2004) "Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping", *Journal of Experimental Social Psychology* 40, 642-658.
Delfgaauw J., Dur R., Sol J. and Verbeke W. (2013) "Tournament incentives in the field: Gender differences in the workplace", *Journal of Labor Economics* 31, 305-326.
Dezsö C.L. and Ross D.G. (2012) "Does female representation in top management improve firm performance? A panel data investigation", *Strategic Management Journal* 33, 1072-1089.
Eckel C.C. and Grossman P.J. (2001) "Chivalry and solidarity in ultimatum games", *Economic Inquiry* 39, 171-188.
Faccio M. and Lang L.H. (2002) "The ultimate ownership of Western European corporations", *Journal of financial economics* 65, 365-395.
Faccio M., Marchica M.-T. and Mura R. (2012) "CEO gender, corporate risk-taking, and the efficiency of capital allocation", *Available at SSRN 2021136*.
Gagliarducci S. and Paserman M.D. (2012) "Gender interactions within hierarchies: evidence from the political arena", *The Review of Economic Studies* 79, 1021-1052.
Gneezy U., Leonard K.L. and List J.A. (2009) "Gender differences in competition: Evidence from a matrilineal and a patriarchal society", *Econometrica* 77, 1637-1664.
Gneezy U., Niederle M. and Rustichini A. (2003) "Performance in competitive environments: Gender differences", *The Quarterly Journal of Economics* 118, 1049-1074.
Gneezy U. and Rustichini A. (2004) "Gender and competition at a young age", *American Economic Review*, 377-381.
Greig F. and Bohnet I. (2009) "Exploring gendered behavior in the field with experiments: Why public goods are provided by women in a Nairobi slum", *Journal of Economic Behavior & Organization* 70, 1-9.
Guiso L. and Rustichini A. (2012) "What drives women out of entrepreneurship? The joint role of testosterone and culture".
Hoogendoorn S., Oosterbeek H. and Van Praag M. (2013) "The impact of gender diversity on the performance of business teams: Evidence from a field experiment", *Management Science* 59, 1514-1528.
Huang J. and Kisgen D.J. (2013) "Gender and corporate finance: Are male executives overconfident relative to female executives?", *Journal of financial economics* 108, 822-839.
Koenig A.M., Eagly A.H., Mitchell A.A. and Ristikari T. (2011) "Are leader stereotypes masculine? A meta-analysis of three research paradigms", *Psychological bulletin* 137, 616.

Matsa D.A. and Miller A.R. (2011) "Chipping Away at the Glass Ceiling: Gender Spillovers in Corporate Leadership", *American Economic Review* 101, 635-639.
Matsa D.A. and Miller A.R. (2012). Workforce reductions at women-owned businesses in the United States: Working paper.
Matsa D.A. and Miller A.R. (2013) "A female style in corporate leadership? Evidence from quotas", *American Economic Journal: Applied Economics* 5, 136-169.
Miller D., Minichilli A. and Corbetta G. (2013) "Is family leadership always beneficial?", *Strategic Management Journal* 34, 553-571.
Niederle M. and Vesterlund L. (2007) "Do women shy away from competition? Do men compete too much?", *The Quarterly Journal of Economics* 122, 1067-1101.
Pérez-González F. (2006) "Inherited control and firm performance", *The American Economic Review*, 1559-1588.
Porta R., Lopez- de- Silanes F. and Shleifer A. (1999) "Corporate ownership around the world", *The Journal of Finance* 54, 471-517.
Price C.R. (2012) "Gender, competition, and managerial decisions", *Management Science* 58, 114-122.
Rey-Biel P. and Iriberri N. (2013) "Let's (not) talk about sex: The effect of information provision on gender differences in performance under competition".
Shurchkov O. (2012) "Under pressure: gender differences in output quality and quantity under competition and time constraints", *Journal of the European Economic Association* 10, 1189-1213.
Tate G. and Yang L. (2014) "Female leadership and gender equity: Evidence from plant closure", *Journal of financial economics*.
Villalonga B. and Amit R. (2006) "How do family ownership, control and management affect firm value?", *Journal of financial economics* 80, 385-417.
Westphal J.D. and Zajac E.J. (1995) "Who shall govern? CEO/board power, demographic similarity, and new director selection", *Administrative Science Quarterly* 40.

Appendix A: Tables

Table 4.1 Summary statistics

Panel A. Firm characteristics

	Number of observations	Mean	Median	s.d.
ROA	10,154	0.054	0.045	0.069
Ln firm age	10,154	3.108	3.219	0.694
Ln assets	10,154	11.420	11.296	1.175
Family ownership	10,123	0.915	1	0.151
Number of CEOs	10,154	1.979	2	1.224
Family CEOs	10,154	0.742	1	0.376
Number of directors	10,154	4.996	4	2.422
Directors' age	10,120	53.780	54	7.253
Family directors	10,153	0.650	0.667	0.300

Panel B. Women in CEO and board positions

	Number of observations	Mean	Median	s.d.
Female CEO presence	10,154	0.204	0	0.403
Female CEOs	10,154	0.107	0	0.242
Female directors	10,154	0.174	0.143	0.199

Notes. ROA is the ratio of earnings before interest and taxes (EBIT) divided by the book value of total assets. Ln firm age and Ln assets are, respectively, the natural logarithm of a firm's age measured in years, and of the book value of total assets. Family ownership is the fraction of equity held by the family. Number of CEOs is the number of a firm's CEOs. Family CEOs is the ratio of family CEOs to the total number of CEOs. Number of directors is the number of a firm's board members. Directors age is the average age of board members. Family directors is the ratio of family directors to the total number of board members. Female CEO presence is a dummy equal to one if at least one of the firm CEOs is woman. Female CEOs is the ratio of female CEOs to the total number of a firm's CEOs. Female directors is the ratio of female directors to the total number of board members.

Table 4.2 Preliminary evidence

Dependent variable: ROA			
	No female CEOs	At least one female CEO	Full sample
	[1]	[2]	[3]
Female CEO presence			0.000
			(0.004)
Female CEO presence × Female directors			0.018**
			(0.009)
Female directors	−0.006	0.020**	−0.008
	(0.005)	(0.010)	(0.005)
Ln firm age	0.003***	0.003	0.003***
	(0.001)	(0.002)	(0.001)
Ln assets	−0.005***	−0.005***	−0.005***
	(0.001)	(0.002)	(0.001)
Cash holdings	0.159***	0.271***	0.191***
	(0.013)	(0.029)	(0.012)
Capex	−0.019***	−0.028***	−0.019***
	(0.005)	(0.010)	(0.005)
Family CEOs	−0.001	−0.007	−0.002
	(0.003)	(0.010)	(0.003)
Family directors	−0.004	0.010	−0.004
	(0.004)	(0.007)	(0.003)
Family owners	0.011***	−0.009	0.009**
	(0.004)	(0.010)	(0.004)
Industry fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Number of observations	5,764	1,637	7,401

Notes. Robust standard errors are reported in parentheses below the coefficients.
*, **, and *** denote significance at (respectively) 10%, 5%, and 1%.

Table 4.3 Fixed-effects regressions

Dependent variable: ROA				
	Main effects		Interactions	
	[1]	[2]	[3]	[4]
Female CEO presence	−0.002	−0.002	−0.016***	−0.011
	(0.003)	(0.003)	(0.006)	(0.008)
Female CEO presence × Female directors			0.050***	0.047***
			(0.015)	(0.018)
Female directors	−0.021*	−0.023*	−0.033**	−0.031**
	(0.012)	(0.012)	(0.014)	(0.013)
Ln firm age		0.009**	0.009**	0.009**
		(0.004)	(0.004)	(0.004)
Ln assets		−0.001	−0.001	−0.001
		(0.002)	(0.002)	(0.002)
Firm fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Industry trends	No	Yes	Yes	Yes
Number of observations	10,154	10,154	10,154	10,154

Notes. Robust standard errors are reported in parentheses below the coefficients.
*, **, and *** denote significance at (respectively) 10%, 5%, and 1%.

Table 4.4 Confounding factors

Panel A. Additional controls

Dependent variable: ROA

	[1]	[2]	[3]	[4]	[5]
Female CEO presence	−0.017***	−0.015***	−0.023***	−0.025***	−0.016***
	(0.005)	(0.005)	(0.006)	(0.007)	(0.006)
Female CEO presence × Female directors	0.049***	0.047***	0.054***	0.057***	0.052***
	(0.015)	(0.015)	(0.017)	(0.018)	(0.015)
Female directors	−0.035**	−0.039***	−0.039**	−0.043**	−0.034**
	(0.014)	(0.015)	(0.018)	(0.018)	(0.014)
Family CEOs	0.014**				
	(0.006)				
Family directors		0.014			
		(0.009)			
Family owners			0.013*		
			(0.007)		
Female owners				0.010*	
				(0.006)	
Nonexecutive directors					0.004
					(0.005)
Firm controls	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Industry trends	Yes	Yes	Yes	Yes	Yes
Number of observations	10,154	10,153	7,613	7,531	10,154

Panel B. Age interactions

Dependent variable: ROA

CEOs' age	0.001
	(0.001)
CEOs' age × Directors' age	−0.000
	(0.000)
Directors' age	0.001
	(0.001)
Firm controls	Yes
Firm fixed effects	Yes
Year fixed effects	Yes
Industry trends	Yes
Number of observations	10,120

Notes. Robust standard errors are reported in parentheses below the coefficients. *, **, and *** denote significance at (respectively) 10%, 5%, and 1%.

Table 4.5 Robustness

	No CEO–board overlaps	Single-CEO firms	EBITDA to assets	Net income to assets	Random effects	Regional trends	Trimmed ROA	Firm-clustered residuals	Nonlinear effects	Lagged ROA
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Female CEO presence	−0.012***	−0.028**	−0.015***	−0.010**	−0.004	−0.016***	−0.016***	−0.016**	−0.014**	−0.016*
	(0.004)	(0.012)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.008)	(0.007)	(0.009)
Female CEO presence × Female directors	0.039**	0.067**	0.042***	0.048***	0.027**	0.048***	0.040***	0.050**	0.044**	0.036*
	(0.016)	(0.033)	(0.015)	(0.015)	(0.013)	(0.015)	(0.012)	(0.022)	(0.018)	(0.021)
Female directors	−0.035**	−0.058*	−0.028**	−0.031**	−0.010	−0.030**	−0.018***	−0.033	−0.042*	−0.016
	(0.014)	(0.034)	(0.012)	(0.015)	(0.008)	(0.013)	(0.007)	(0.022)	(0.022)	(0.017)
Firm controls	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	10,154	4,550	10,154	10,154	10,154	10,154	9,952	10,154	10,154	6,527

Notes. Robust standard errors (clustered by firm in column [8]) are reported in parentheses below the coefficients.
*, **, and *** denote significance at (respectively) 10%, 5%, and 1%.

Table 4.6 Triple-difference estimates

Dependent variable: ROA								
	DD estimates		Interaction: DDD estimates		Unique transitions	Single-CEO firms	No CEO–board overlaps	Placebo transitions
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Post × Female incoming CEO	0.010*	0.010*	−0.014	−0.012	−0.011	−0.022	0.001	0.007
	(0.006)	(0.006)	(0.009)	(0.008)	(0.014)	(0.017)	(0.006)	(0.008)
Post × Female incoming CEO × Female directors			0.087***	0.076**	0.076*	0.091	0.073**	−0.007
			(0.033)	(0.032)	(0.041)	(0.067)	(0.034)	(0.008)
Female directors	−0.036	−0.038	−0.037	−0.043	−0.079*	−0.132	−0.029	−0.060*
	(0.029)	(0.029)	(0.027)	(0.027)	(0.044)	(0.097)	(0.031)	(0.031)
Post × Female directors			−0.018	−0.012	−0.017	−0.039	−0.022	0.013
			(0.020)	(0.021)	(0.024)	(0.052)	(0.025)	(0.016)
Female incoming CEO × Female directors			0.001	0.005	0.065	0.183	0.012	0.048
			(0.033)	(0.033)	(0.059)	(0.134)	(0.043)	(0.033)
Firm controls	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry trends	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Number of observations	3,343	3,343	3,343	3,343	2,832	1,566	3,343	3,343

Notes. Robust standard errors are reported in parentheses below the coefficients.
*, **, and *** denote significance at (respectively) 10%, 5%, and 1%.

Table 4.7 Triple-difference estimates: Alternative estimation methods

Dependent variable: ROA				
	Matched sample		2SLS	
	[1]	[2]	[3]	[4]
Post × Female incoming CEO	−0.013	−0.017	−0.082	−0.054
	(0.015)	(0.015)	(0.108)	(0.091)
Post × Female incoming CEO × Female directors	0.076*	0.093**	0.386*	0.345*
	(0.044)	(0.043)	(0.217)	(0.197)
Female directors	−0.014	−0.017	−0.128**	−0.123**
	(0.027)	(0.032)	(0.064)	(0.060)
Post × Female directors	−0.024	−0.034	−0.110**	−0.122**
	(0.027)	(0.027)	(0.056)	(0.059)
Female incoming CEO × Female directors	−0.013	−0.014	0.238	0.282
	(0.044)	(0.049)	(0.181)	(0.187)
Firm controls	No	Yes	No	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Industry trends	No	Yes	No	Yes
Number of observations	851	851	1,973	1,973

Notes. Robust standard errors are reported in parentheses below the coefficients.
*, **, and *** denote significance at (respectively) 10%, 5%, and 1%.

Table 4.8 Gender stereotypes

Dependent variable: ROA									
	Geographic areas			Overall emancipation			Emancipation of women in business		
	South	Central	North	Low	Medium	High	Low	Medium	High
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Female CEO presence	−0.064***	−0.014	−0.014**	−0.019**	0.008	−0.028***	−0.027***	0.005	−0.038***
	(0.020)	(0.015)	(0.006)	(0.008)	(0.013)	(0.009)	(0.008)	(0.010)	(0.012)
Female CEO presence × Female directors	0.082*	0.047	0.053***	0.048**	−0.023	0.083***	0.058***	0.013	0.097***
	(0.047)	(0.038)	(0.018)	(0.022)	(0.031)	(0.027)	(0.022)	(0.023)	(0.033)
Female directors	−0.029	0.004	−0.045***	−0.000	0.006	−0.076***	−0.002	−0.008	−0.087***
	(0.038)	(0.020)	(0.017)	(0.016)	(0.017)	(0.027)	(0.017)	(0.013)	(0.033)
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	587	1,551	8,016	3,386	2,089	4,679	3,139	3,458	3,557

Notes. Robust standard errors are reported in parentheses below the coefficients.
*, **, and *** denote significance at (respectively) 10%, 5%, and 1%.

Table 4.9 Firm and business cycle characteristics

Dependent variable: ROA	Firm size		Directors' affiliation		Business cycle	
	Small	Large	Family	Nonfamily	Normal times	Financial crisis
	[1]	[2]	[3]	[4]	[5]	[6]
Female CEO presence	−0.014	−0.018**	−0.010**	−0.006**	−0.014*	0.007
	(0.009)	(0.007)	(0.005)	(0.003)	(0.007)	(0.013)
Female CEO presence × Female directors	0.050**	0.041**	0.027**	0.095**	0.046**	−0.009
	(0.025)	(0.018)	(0.013)	(0.042)	(0.022)	(0.036)
Female directors	−0.044*	−0.009	−0.005	−0.099***	−0.054***	0.027
	(0.027)	(0.011)	(0.010)	(0.037)	(0.021)	(0.032)
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry trends	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	5,077	5,077	10,141	10,141	6,187	3,967

Notes. Robust standard errors are reported in parentheses below the coefficients.
*, **, and *** denote significance at (respectively) 10%, 5%, and 1%.

CHAPTER V

HOW DO FEMALE BANKERS REACT TO COMPETITIVE PRESSURES?

1. Introduction

An established strand of experimental economics research has analyzed gender differences in competitiveness and risk-taking behavior. Some central findings of this literature are that women tend to be more risk-averse than men (see e.g., Charness and Gneezy 2012) and that, relative to male performance, competitive pressures are detrimental to female performance (Gneezy and Rustichini 2004; Gneezy, Niederle and Rustichini 2003), especially in high-pressure and stereotypically male-oriented tasks (Rey-Biel and Iriberri 2013; Shurchkov 2012; Günther, Ekinci, Schwieren and Strobel 2010). These findings are important as they may help explain the typical under-representation of women in top highly competitive corporate jobs (Bertrand and Hallock 2001) as well as the potential firm implications of recent policies aimed at increasing female representation in corporate positions.

However, as some studies have recently noted (see e.g., Adams and Ragunathan 2013; Adams and Funk 2012), gender differences in competitiveness and risk aversion have been established primarily using samples of students or individuals from the general population, and the general validity of these findings in the context of firm executives is unclear. Because women at the top of companies have self-selected themselves into performance-oriented environments and have successfully gone through highly competitive recruiting processes, they are expected to be dissimilar from women from the general population in their attitudes toward competition and risk, whereas the differences with men could weaken or even disappear.

A number of empirical findings are consistent with this notion. For instance, Guiso and Rustichini (2012) show that women who break the glass ceiling and manage to become entrepreneurs display more masculine traits and are as equal as men in translating their ability into firm outcomes. Analyzing individual traits, Adams and Ragunathan (2013) suggest that women who choose finance careers are less risk averse than other women (and thus more similar to men). These similarities suggest that we still know little about whether, when focusing on top corporate executives, men and women display any difference in task performance under competition or in risk-taking attitudes.

In this study, we focus on legislative changes in the US banking industry to test whether women and men executives respond differently to an exogenous increase in competitive

pressures. Our main results indicate female leadership in times of tough competition is conducive of lower profitability and, at the same time, lower corporate risk.

The US banking sector in the late 20th century represents an ideal laboratory for our study. While historical regulations severely limited the geographic expansion of banks, US states gradually lifted these restrictions starting in the 1970s.[51] Specifically, our identification is based on the legal roadblocks US states passed to limit the scope of deregulation of interstate branching activities introduced in 1994 with the Interstate Banking and Branching Efficiency Act (IBBEA). The staggered introduction and removal of these restrictions determined both temporal and geographic variations in the intensity of competition from the mid 1990s to the early 2000s, variations that are useful for mitigating endogeneity concerns (Cornaggia et al. 2013; Rice and Strahan 2010; Johnson and Rice 2008).

Using data on listed US banking institutions from 1997 to 2006, our main results indicate that, relative to male-led companies, banks led by female executives underperform when competitive pressures increase. This finding holds when adopting both accounting and market-based measures of performance as well as controlling for individual- and firm-specific factors to account for differences between male- and female-led banks. Moreover, we validate our results by the following: (1) mitigating the concern that stronger competition changes the likelihood of having a woman at the top, (2) showing that no female underperformance emerges for companies operating in industries unaffected by the regulatory changes or (3) using placebo reform years to rule out anticipation effects and strengthen the causal interpretation, and (4) controlling for the interaction of competitive pressures with various bank-specific variables (to allow for the fact that deregulation affected differently banks with certain observable characteristics that are correlated with executives' gender).

Our results are broadly consistent with the notion that, relative to male executives, even female executives may underperform when competitive pressures increase. However, another interpretation is that prior to regulatory changes used in our analysis, female-led banks were enjoying extra-profits reflecting rent, e.g., if they operated in rural areas relatively more protected from existing within-state competition. In this case, the profitability drop experienced by female-led banks may arise because increased competition eroded those noncompetitive extra-profits, whereas male-led banks were already gaining normal profits at a more competitive level. We rule out this alternative interpretation by determining that the

[51] This approach has been used in several studies that analyze the effect of banking structure characteristics on aggregate and firm-level outcomes (see e.g., Amore, Schneider and Žaldokas 2013; Chava, Oettl, Subramanian and Subramanian 2013; Cornaggia, Mao, Tian and Wolfe 2013; Rice and Strahan 2010; Cetorelli and Strahan 2006; Black and Strahan 2002, 2001)

profitability of male- and female-led banks does not differ on the first year they enter the sample and by verifying that our regression results are robust to controlling for early profitability differences and for location type.

After having established the profitability effect, we test whether any gender difference exists in corporate risk-taking following the increase in competitive pressures. Although the general effect of competition on banks' risk-taking is ambiguous (see e.g., Martinez-Miera and Repullo 2010; Boyd and De Nicolo 2005), Dick (2006) provides some indication that interstate branching deregulation led to an increase in banks' portfolio credit risk. Using various proxies of banks' risk-taking, we find significant evidence that female-led banks were less subject to an increase in risk after competition increased. Thus, our analysis indicates that, despite lower profitability, female leadership generates superior bank capitalization and more stable streams of profits.

We broadly relate to a recent research on the impact of women in politics (Clots-Figuerasa 2012; Gagliarducci and Paserman 2012; Chattopadhyay and Duflo 2004) and in companies (Amore, Garofalo and Minichilli 2014; Tate and Yang 2014; Matsa and Miller 2013; Ahern and Dittmar 2012; Adams and Ferreira 2009). A number of recent studies in this area report that, even at the top of corporate ladders, women differ from men in some important dimensions and that these differences may lead to different corporate policies. For instance, investigating corporate outcomes, Huang and Kisgen (2013) provide evidence that male and female CFOs differ in their financial decisions; the latter make fewer high-quality acquisitions and are less debt-oriented, possibly due to different risk preferences. In a similar vein, Faccio, Marchica and Mura (2012) argue that female-led firms have less volatile streams of profits and a higher chance of survival. While these works try to gain some understanding of how possible gender differences in preferences translate into corporate policies, we identify the effect of competitive pressures on executives task performance, as mirrored by firm profitability[52], and on risk taking. The identification of these effects is complicated by the fact that, even within the competitive world of corporate jobs, women managers may self-select into more protected market niches because of a possible intrinsic tendency to shy away from competition (Booth and Nolen 2012; Niederle and Vesterlund 2007). We are able to significantly mitigate this concern because we employ exogenous variations to competitive pressures for given executives-firm observations within the same sector.

[52] Existing works indicate that firm profitability is a close function of CEO task performance. For instance, (Bennedsen, Pérez-González and Wolfenzon 2012, 2011) show that individual CEO shocks translate into firm profitability.

Our work is also close to recent works that employ real-world data to analyze gender differences in response to changing competition. Delfgaauw, Dur, Sol and Verbeke (2013) find that the introduction of tournament competition affects firm sales depending on the gender composition of work teams, whereas Morin (2013) and Ors, Palomino and Peyrache (2013) use educational data to show that men respond more effectively than women to a higher level of competition. We extend these works in a number of significant directions. While these works analyze reactions to competition in a non-corporate context or in small and young organizations (firms analyzed in Delfgaauw et al. 2013 have, on average, 11 employees that are 25 years old), we focus on the profitability and risk profile of listed US banking institutions. This context is especially interesting as we expect the selection channel to be most pronounced: women that managed to get to the top of large corporations in a stereotypically masculine sector should be most likely to display preferences for competitive environments and risk-taking. Yet, our results highlight the presence of significant gender differences in response to increased competition and thus suggest that existing experimental findings do extend to the corporate arena above and beyond the mitigating role of selection. Identifying such differences in the context of banks is also important because of the negative externalities that excessive risk may generate for the whole economy. Regarding this point, our results highlight a trade-off whereby, under increased competitive pressures, female-led companies experience lower profits but at the same time lower risk-taking.

Section 2 illustrates the regulatory changes used to generate exogenous variations in competitive pressures in the US banking industry. Section 3 describes the data used in the empirical analysis and provides summary statistics. Section 4 presents the main profitability results, together with a number of placebo checks and extensions. Section 5 presents our findings on corporate risk. Section 6 concludes.

2. Competitive pressure in the US banking sector

A number of historical regulations such as the McFadden Act of 1927 severely limited the expansion of banks across and within the US states. However, starting in the late 1960s, states started deregulating within-state branching activity, allowing the creation of new branches via M&A of existing banks and via de novo branching. Moreover, starting in the early 1980s, states started deregulating interstate ownership of banks, which was prohibited by the Douglas Amendment to the 1956 Bank Holding Company Act. While this wave of deregulation had a profound impact on the structure and efficiency of the US banking system (see Jayaratne and Strahan 1998; 1996), interstate branching was still prohibited.

This restriction was removed with the Interstate Banking and Branching Efficiency Act (IBBEA) passed by the US Congress in 1994. The economic effects of the IBBEA began in 1997 and continued until the mid 2000s. This relaxation of interstate branching restrictions led to a significant development of banking activities across state borders. For instance, Johnson and Rice (2008) show that, while only 64 out-of-state banks existed in a few states in 1994, by 2005 this number had increased to more than 24,000, and more than 6,000 de novo out-of-state branches were opened over the same period. These figures point to a substantial increase in competitive pressures among banking institutions.

However, while initiating the deregulation of interstate branching activities, the IBBEA also granted US states the right to erect some barriers for the branches of out-of-state banks by allowing them to pass a law at any time between the IBBEA passage in September 1994 and its trigger date in July 1997. In particular, states were allowed to erect barriers regarding four provisions of the IBBEA, namely: (1) the minimum age of target institutions; (2) de novo interstate branching; (3) the acquisition of individual branches; and (4) a statewide gap on the amount of deposits.[53]

Similar to existing works (e.g., Cornaggia et al. 2013; Rice and Strahan 2010), we exploit the changes in the number of these state-level roadblocks to the interstate branch expansion in order to establish variations in the increase of competitive pressures across US states. Specifically, we sum the number of barriers to the four above-described provisions. More precisely, following Rice and Strahan (2010), we add one if a state: (1) imposes a minimum age of 3 or more years on target institutions of interstate acquirers; (2) does not permit de novo interstate branching; (3) does not permit the acquisition of individual branches by an out-of-state bank; (4) imposes a deposit cap lower than 30%. The resulting index (RS index hereafter) is equal to zero for states that are most open to interstate branching (e.g., Ohio and Pennsylvania) and to four for states that are most closed to interstate branching (e.g., Arkansas and Colorado).

Given that different states passed the above-mentioned roadblocks at different points in time, the index displays both temporal and geographic variation. Table A1 (Appendix B) illustrates all changes by state and year. As shown, the first changes are enacted in 1995 and 1996. Given that our sample spans from 1997 (when data on executives' gender becomes available) to 2006, we do not have any control period prior to changes of the RS index. However, US states keep revising interstate branching barriers until the early 2000s and our

[53] See Johnson and Rice (2008) and Rice and Strahan (2010) for a detailed discussion of each of these provisions.

sample period covers around 25% of all changes in the RS index (corresponding to around 22% of observations). In our sample, the index takes an average of 1.8 and a standard deviation of 1.3 (values similar to those reported by Rice and Strahan 2010).

3. Data and variables

Information on executives' gender comes from the Investor Responsibility Research Center (IRRC) and the RiskMetrics dataset. We construct a dummy equal to one if a firm executive is female, and zero otherwise. Using the same data source, we construct the natural logarithm of executives age measured in years.

To obtain firm-level accounting information, we match the IRRC dataset with the Compustat dataset of publicly listed firms headquartered in the US. Given our focus on the banking industry, we restrict the analysis to standard industrial classification (SIC) codes between 6000 and 6200. We also exclude observations for which the net sales or book value of assets are either missing or negative. Our sample period ranges from 1997 through 2006 and consists of 168 unique firms and 854 firm-year observations (even if missing values in explanatory variables reduce the number of observations used in the regression analysis). We assign the RS index to companies on the basis of the state of headquarter. Given that the state of headquarter is typically chosen at the beginning of a firm's operations, it can be thought to be predetermined to deregulation events.[54]

Common to the literature (see e.g., Fahlenbrach and Stulz 2011), our main measure of accounting performance is the return on assets (ROA), computed as the ratio of earnings before interest, taxes, depreciation, and amortization (EBITDA) to the book value of total assets. To mitigate concerns about outliers, we drop 1% of the firm-year observations from each tail of the ROA distribution. As an alternative performance measure, we employ the market to book ratio, computed by dividing the market value of each firm by the book value of its common equity. Following Baker and Wurgler (2002), the market to book ratio is limited to the interval between 0 and 10.

Following existing works (e.g., Houston, Lin, Lin and Ma 2010; Laeven and Levine 2009) our main measure of bank risk is the z-score, computed as ROA plus the capital to assets ratio divided by the standard deviation of ROA.[55] Because the z-score is the inverse of probability

[54] The average and median years the firms in our sample have been in Compustat are 23 and 20, respectively. However, one concern is that because Compustat only reports the last state of headquarter we cannot control for headquarter relocations potentially driven by the passage of restrictions to IBBEA restrictions. This concern is mitigated by the infrequent nature of headquarter relocations. For instance, Pirinsky and Wang (2006) find that, of more than 4000 firms over the period 1992-1997, only 118 relocated for reasons other than M&A and major restructuring events.

[55] Standard deviations in the risk analysis are computed over three-year periods (1997-1999, 2000-2002 and 2003-2006).

that firms' losses exceed equity, it can be considered a measure of the distance from insolvency. By construction, higher values indicate more bank stability. Given that the z-score is highly skewed, we employ its natural logarithm in the empirical analysis. We then focus on the individual components of the z-score. First, we adopt two volatility measures, namely, the standard deviation of ROA and market returns. Second, we use the capital to assets ratio and the risk-adjusted (Tier1 and Tier2) capital ratio as measures of banks' stability, where higher values correspond to less fragile banks.

We construct a number of firm-specific variables used as controls in the regression analysis. In particular, we take the logarithm of firm revenues, the ratio of liquid holdings to total assets, the ratio of financial debt to the book value of total assets and, lastly, the ratio of tangible assets to total assets. Appendix A2 reports a detailed description of how each variable was constructed.

Summary statistics (for the subsample of observations with non-missing information on executives' gender) are reported in Table 5.1. As shown, approximately 14% of the observations report the presence of a female executive, whereas the median is zero. Focusing on performance measures, we find an average ROA of 3.5% and a market to book ratio of 2.5.

4. Profitability

4.1. Identification and main results

Our identification relies on the combination of (1) across-state variations in competitive pressures, as generated by the adoption and removal of roadblocks against the IBBEA deregulation of interstate branching, and (2) within-state differences in executives' gender across companies.

To exemplify our identification, we can consider two treatment banks in a given state: they are hit by the same increase in competitive pressures but differ in their executives' gender. Taking the difference in performance between these two banks around the competition increase allows us to establish the differential performance effect of competitive pressures by gender. At the same time, exploiting the staggered change in competitive pressures, we can employ in the analysis two additional banks in a different state, which again differ in their executives' gender but for which the competitive pressures have not changed. These control banks are useful to absorb the effect of general macroeconomic changes as well as differences that are specific to female-executives banks.

Generalizing this example to the full sample, we estimate the following model:

$$ROA_{itk} = \alpha + \beta_1 \text{Female Executives}_{it} + \beta_2 RS\ index_{tk} + \beta_3 \text{Female Executives}_{it} \times RS\ index_{tk} + X_{it-1}\delta + \theta_k + \tau_t + \varepsilon_{itk}$$

in which the dependent variable is the time-t ROA of a firm i headquartered in state k. The female executives dummy is equal to one if the executives of firm i at time t is female, and zero otherwise. The RS index captures variations in competitive pressures (with higher values indicating more protection from interstate branching and thus lower competitive pressures).

The interaction between the RS index and the female executives dummy measures how the profitability effect of female executives varies with changing competitive pressures. A positive and significant interaction coefficient β_3 would indicate that increased banking competition has a more negative effect on the performance of banks led by women.

The specification includes for a host of firm-level variables, which are included in the X vector. These variables are important to control for the fact that male- and female-led banks may differ along a number of observable characteristics. Specifically, we consider the logarithm of revenues to control for differences in size and asset tangibility to control for differences in the asset base. Furthermore, we consider the ratios of total debt and cash holdings to total assets to control for the fact that more indebted companies may be worst at responding to an adverse shock (Khanna and Tice 2000; Zingales 1998). Finally, we include the logarithm of executives age to further mitigate concerns of omitted factor bias at the executives level.

We take a one-year lag of all controls to avoid simultaneity problems between RS index and firm controls; however, our results are robust to using contemporaneous rather than lagged controls.[56]

Our specification also includes year dummies, τ_t, to absorb shocks common to all banks and state fixed effects, θ_k, to control for geographic time-invariant heterogeneity. Although our firms belong to the same sector, we can include SIC fixed effects and their interaction with year dummies to control for the possibility that, e.g., commercial banks (SIC 602) were differently affected than saving institutions (SIC 603) by the increase in competitive pressures.[57]

Residuals are clustered by a firm's state of headquarter, which is the appropriate level of clustering given that the RS index affected companies at the state level. However, in

[56] In unreported regressions, we have also verified that our findings are robust to including of quadratic terms of CEO age, as well as to controlling for size in alternative ways, e.g. by taking the logarithm of total assets.

[57] We use SIC dummy times year dummies to control for SIC specific trends given recent works suggesting that including the industry-year average of the dependent variable as a control or taking an industry adjustment of the dependent variable may lead to misleading inference (Gormley and Matsa 2014) While we prefer to control for SIC-specific trends in the baseline specification, our main findings are largely unaffected by the exclusion of these controls.

additional checks we confirm the robustness of the main results computing standard errors in alternative ways.

Results obtained using ROA as a dependent variable are reported in Table 5.2, Columns (1)-(3), in which we first estimate the model only controlling for firm size; second, we add the additional firm-level controls; and third, we control for executives age. Focusing on the result obtained using the full set of controls (Column 3) we find that the coefficient of Female Executives is negative and significant at the 10% level. In other words, female-led banks experience lower profitability than their male counterpart when banking competition is at its highest level (i.e. RS index equal to zero). The coefficient of RS index is insignificant in itself, possibly owing to the fact that, pushed by the increased competition, banks improved existing service quality and introduced new (more sophisticated) services that enabled them to maintain existing revenues (Dick 2006) but that this activity was conducted primarily by male-led banks. The interaction between Female Executives and RS index is positive and significant at the 5% level; thus, as competition increases, female-led banks experience profitability rates significantly lower than male-led banks.

To interpret the economic magnitude of female leadership under competitive pressures, we can consider the maximum relaxation of interstate branching barriers: when the state enjoys the highest barriers to competition, the total economic effect of a Female Executives is given by -0.0041+(0.0029×4), i.e., 0.0075. However, when the RS index drops to zero (i.e., highest competitive pressures), the effect of Female Executives is given by the stand-alone coefficient of the Female Executives dummy, i.e., -0.0041 (significant at the 10% level).

These estimates indicate that an increase in competitive pressures is detrimental to the operating performance of female-led banks. One possible interpretation of our finding so far is that female-led firm profits do not significantly change, but ROA decreases because female executives invest more and thus increase a firm's asset base (i.e., the ROA denominator) as competition increases. We rule out this interpretation by determining that the interaction term does not have any significant effect on the logarithm of total assets or the logarithm of property, plants and equipment expenses.

In Table 5.2, Columns (4) and (5), we use the market to book ratio as a dependent variable. The results are in line with our findings on ROA. This additional evidence is important because it shows that our findings are not specific to using accounting measures of performance but rather that the effect is incorporated in the stock prices.

4.2. Validation of casual effect

Female leadership may be more present in banks with given bank characteristics (e.g. large banks) and thus our results may reflect an interaction with such characteristics rather than with executives' gender. In Table 5.3 we verify that our results hold including the interaction of the RS index with each of the firm characteristics previously used as controls, thus allowing for heterogeneous effects of changing competition on profits by firm characteristics.

One challenge to our identification is that the RS index itself may change likelihood of having a female executives. For instance, the boards of directors may have a lower tendency to appoint and/or retain women executive when competition is high. Alternatively, women may shy away from competition (Niederle and Vesterlund 2007) thereby shrinking the pool of female executives candidates, which in turn makes their recruitment less likely, or, lastly, women executives may have a higher likelihood of exiting a competitive contest (Hogarth, Karelaia and Trujillo 2012).

To test for the importance of these concerns, we estimate a linear probability model in which the dependent variable is the Female Executives dummy and the main explanatory variables are the RS index together with one-year lagged firm size, state fixed effects and SIC-year dummies. The results, tabulated in Table 5.4, indicate that the RS index does not have any significant effect on the likelihood of having a female executive. In an unreported check, we further mitigate this concern by estimating our baseline models in Table 5.2 after excluding banks that experienced a change in executives' gender in the period following a change of the RS index; the exclusion of these few banks does not materially affect any of our results.

We further validate the causal interpretation of our findings using a number of placebo tests. First, we check whether our findings are indeed specific to banks (and to the increase in banking competition they were subject to) by estimating our main specification on a sample of manufacturing firms (SIC codes between 2000 and 4000). Given that these firms were not directly affected by any change in competitive pressures, one should not find any differential effect in performance by executives' gender. The results in Table 5.5, Column (1), indeed show that the interaction term is statistically and economically insignificant. In unreported regressions, we also confirm this lack of significance using firms in all SIC codes except those related to banking institutions.

Second, we construct changes of the RS index in placebo years. This test is important to rule out concerns of pre-existing trends and confounding events, thus confirming that our findings are indeed attributable to the RS index. The lack of significance in the interaction

term (Table 5.5, Column 2) offers confirming evidence that the effect is specific to the state-level changes of the RS index. Third, we check that our findings are specific to having a female executive. To this end, we randomly construct executives' gender. As shown in Table 5.5, Column (3), the interaction of placebo executives' gender and RS index is not statistically significant.

4.3. Robustness

In this section, we address a number of empirical concerns to further assess the validity of our findings. The results, reported in Table 5.6, broadly confirm the statistical and economic significance of our main result. While we present only the ROA results to save space, we have verified that each of these checks delivers economically and statistically significant results when also using the market to book ratio as a dependent variable.

As discussed in the introduction, our findings may be consistent with the notion that, prior to the increase in competitive pressures, banks led by women were enjoying extra profits reflecting non-competitive rent. In this case, the larger profitability drop we estimate for female-led banks may be due to increased competition eroding these noncompetitive extra-profits, whereas male-led banks were already gaining normal profits at a more competitive level. In Column (1), we rule out this alternative interpretation by showing that our results are robust to controlling for early profitability differences (i.e. ROA at the beginning of the sample period). In unreported checks, we also used t-tests to verify that the performance of male- and female-led banks at the beginning of the sample was not significantly different.

As discussed in Section 2, from the late 1960s to the mid 1990s US states adopted a number of regulatory changes, particularly concerning intrastate branching and interstate banking activities. Thus, one may be concerned that these alternative policy changes confound our main finding. We mitigate this concern in two ways. First, we check that at the beginning of our sample almost all states (accounting for approximately 94% of observations) had passed interstate banking and intrastate branching deregulation, thus confirming that these deregulation events do not overlap with changes in the roadblocks to the IBBEA provision. Second, in Column (2) we control for the number of years since a state has passed intrastate branching and interstate banking deregulation.

Increases in competitive pressures may have affected firm entry and exit, which in turn can influence our estimates. We mitigate this concern by estimating our main specification on the subsample of banks that are present for at least 5 years in the sample. The results are reported in Column (3).

Next, we address the computation of the standard errors. In the baseline regressions, we cluster standard errors by the state of headquarter, which is the level at which regulatory changes affected banks. However, we confirm that our findings are robust to cluster standard errors by firm (Column 4) or by computing heteroskedasticity-robust (un-clustered) standard errors (Column 5).

Although in our main specification we already drop 1% of observations from the right and left tails of the ROA distribution, we confirm that outliers do not affect our results by dropping a further 1% from the tails of the ROA distribution (Column 6), or 2% (unreported), or by taking a transformation of the dependent variable – the logarithm of one plus ROA (Column 7) – which is less sensitive to extreme observations.

Next, we confirm our findings using a more comprehensive set of controls to further mitigate omitted factor bias concerns. In particular, we add firm age, the capital to labor ratio, the logarithm of board size (measured as the number of board members), the logarithm of executives seniority measured in years, the ratio of independent board members to total directors and banks' location type, i.e. urban vs. other locations (Column 8).[58] These controls are apt to capture differences in corporate governance quality and other firm characteristics between male- and female-led banks that may also correlate with different response to deregulation.

We also address the concern of confounding policies. Over the same time period considered, a few US states changed the tax rate on banks' income, which in turn reduced banks' financial constraints (Farre-Mensa and Ljungqvist 2013). By exploiting within state differences provided by differences in executives' gender, our specification already takes into account these confounding policies and other shocks over the time period considered. However, we control for this concern in a more complete way by adding the tax rate changes as an explanatory variable (Column 9).

We lastly check the robustness to excluding Delaware and South Dakota[59] (Column 10) and to controlling for such state-level characteristics as the natural logarithm of GDP per capita and population size (Column 11).

[58] Urban banks are those headquarted in one of the ten largest metropolitan areas in the US as of 2000, namely Baltimore, Boston, Dallas, Detroit, Chicago, Houston, Los Angeles, New York City, Philadelphia and San Francisco.

[59] This exclusion is common to the banking deregulation literature (see e.g., Black and Strahan 2002), and it is motivated by the fact that these two states experienced a specific development of credit card banks.

5. Corporate risk

Our results so far indicate that female-led companies experienced a decrease in profitability as competitive pressures increased. In this section, we test for the presence of gender differences in banks' risk-taking behavior following the increase in competition.

The general effect of competition on bank risk is highly debated (see e.g., Martinez-Miera and Repullo 2010; Boyd and De Nicolo 2005). On the one hand, there are arguments suggesting that competition should increase survival rates by enhancing banks' efficiency. Focusing on the early stage of US deregulation, there is evidence that intrastate branching deregulation improved the quality of banks' loan portfolios (Jayaratne and Strahan 1996). On one other hand, some works have highlighted that a decrease in banks' market power increases asset risk (see e.g., Matutes and Vives 2000; Keeley 1990). Focusing on interstate branching deregulation, similar to our empirical analysis, Dick (2006) provides evidence that the competition induced by the IBBEA passage increased portfolio credit risk.

We argue that the extent to which banks increased risk following the IBBEA-induced increase in competitive pressures may also depend upon executives' attitude toward risk. The experimental economics literature has pointed to the presence of strong gender differences in risk-taking, whereby women take on less financial risk (see Charness and Gneezy 2012 and references therein). These differences, however, turn out to be weaker once one considers samples of executives. For instance, Adams and Funk (2012) use survey data to show that female directors are not more risk averse than male directors. Croson and Gneezy (2009) suggest that, due to selection or learning reasons, women managers may behave more similar to men in terms of risk-taking. Thus, whether female-led banks changed their attitude toward risk-taking following the increase in competition is theoretically unclear.

Addressing this question empirically, we start by using as a dependent variable the natural logarithm of z-score, a common measure of bank risk (see Houston et al. 2010; Laeven and Levine 2009).[60] We control for one-year lagged firm size, cash holdings and tangibility ratios as well as executives age. In addition, we include one-year lagged ROA to control for differences in profitability positions.

Given that higher z-score values indicate lower bank stability, the negative and 10% significant interaction term reported in Table 5.7, Column (1), suggests that female-led banks take on less risk when competitive pressures increase.

[60] We follow Laeven and Levine (2009) and use an accounting-based version of the *z*-score. In unreported analyses, however, we confirm that our results remain significant to the use of a market-based version (as e.g., Forssbæck 2011).

Next, we focus on the separate components of the z-score. In Column (2), we use as dependent variable the denominator of the z-score, namely the standard deviation of ROA. The results confirm that, when competition increases, female-led banks have less volatile streams of profits. In Column (3), we confirm this finding using a market-based measure of volatility, as proxied by the standard deviation of the market to book ratio. In Column (4), we show the results obtained using the capital to assets ratio as a dependent variable, where higher values indicate greater bank stability. As shown, the interaction term of interest is negative and 10% significant – again indicating a superior stability of female-led banks in periods of increased competition. This latter result is also confirmed using as dependent variable the risk-adjusted capital ratio, computed as the sum of Tier1 and Tier2 capital divided by total assets (Column 5).

Higher capitalization and more stable streams of profits during periods of increased competition may result in a lower likelihood of exit. Looking at the fraction of banks that exited the sample before 2006 (the last year covered in our dataset), we find that this is indeed the case: despite a lower profitability rate, female-led banks are just as likely as male-led banks to exit.[61]

Overall, the results of this section suggest that following an increase in competitive pressures, female-led banks experienced lower profits but, at the same time, lower corporate risk-taking.

6. Conclusion

Policy-makers around the world are implementing policies to increase the representation of women in the political and corporate arena. These interventions are generating an intense debate among academics and practitioners on what would be the effects of female leadership on shareholder returns and policy adoptions.

In this study, we use real-world data to analyze whether men and women executives respond differently to competitive pressures. Our empirical strategy is based on a wave of deregulation in the US banking industry, which generated exogenous variations in competitive pressures between the mid 1990s and the early 2000s. Exploiting the staggered introduction and the removal of barriers to interstate branching activities, our identification exploits increases in competition within a given state, and then establishes the differential

[61] Firms typically exit the Compustat sample due to mergers and acquisitions, bankruptcies, liquidations and delistings.

response by female- and male-led banks while controlling for state-level unobserved heterogeneity, common shocks and firm-level characteristics.

We find that when competitive pressures increase, female-led banks experience lower performance than their male counterparts. This finding holds even after mitigating various threats to our identification strategy, such as selection, reverse causality and confounding policy effects. This finding thus supports the view that experimental results on gender differences in competitiveness do extend to the corporate field above and beyond the selection effect.

Focusing on risk-taking behavior, we find that, relative to male-led banks, banks led by women experience a significantly lower increase in corporate risk. This result is important because it highlights that, despite lower profitability, female leadership may mitigate the potential exacerbating effect of competition on risk-taking and thus improve stability and survival rates.

7. Bibliography

Adams R.B. and Ferreira D. (2009) "Women in the boardroom and their impact on governance and performance", *Journal of financial economics* 94, 291-309.

Adams R.B. and Funk P. (2012) "Beyond the glass ceiling: Does gender matter?", *Management Science* 58, 219-235.

Adams R.B. and Ragunathan V. (2013) "Lehman Sisters".

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

Amore D.M., Garofalo O. and Minichilli A. (2014) "Gender Interactions Within the Family Firm", *Management Science* 0, null.

Amore M.D., Schneider C. and Žaldokas A. (2013) "Credit supply and corporate innovation", *Journal of financial economics* 109, 835-855.

Baker M. and Wurgler J. (2002) "Market timing and capital structure", *The Journal of Finance* 57, 1-32.

Bennedsen M., Pérez-González F. and Wolfenzon D. (2011) "Do CEOs Matter?".

Bennedsen M., Pérez-González F. and Wolfenzon D. (2012). Estimating the value of the boss: Evidence from CEO hospitalization events.

Bertrand M. and Hallock K.F. (2001) "Gender Gap in Top Corporate Jobs, The", *Indus. & Lab. Rel. Rev.* 55, 3.

Black S.E. and Strahan P.E. (2001) "The division of spoils: rent-sharing and discrimination in a regulated industry", *American Economic Review*, 814-831.

Black S.E. and Strahan P.E. (2002) "Entrepreneurship and bank credit availability", *The Journal of Finance* 57, 2807-2833.

Booth A. and Nolen P. (2012) "Choosing to compete: How different are girls and boys?", *Journal of Economic Behavior & Organization* 81, 542-555.

Boyd J.H. and De Nicolo G. (2005) "The theory of bank risk taking and competition revisited", *The Journal of Finance* 60, 1329-1343.

Cetorelli N. and Strahan P.E. (2006) "Finance as a barrier to entry: Bank competition and industry structure in local US markets", *The Journal of Finance* 61, 437-461.
Charness G. and Gneezy U. (2012) "Strong evidence for gender differences in risk taking", *Journal of Economic Behavior & Organization* 83, 50-58.
Chattopadhyay R. and Duflo E. (2004) "Women as policy makers: Evidence from a randomized policy experiment in India", *Econometrica* 72, 1409-1443.
Chava S., Oettl A., Subramanian A. and Subramanian K.V. (2013) "Banking deregulation and innovation", *Journal of financial economics* 109, 759-774.
Clots-Figuerasa I. (2012) "Are Female Leaders Good for Education?", *American Economic Journal: Applied Economics* 4, 212-244.
Cornaggia J., Mao Y., Tian X. and Wolfe B. (2013) "Does banking competition affect innovation", *Journal of financial economics*.
Croson R. and Gneezy U. (2009) "Gender differences in preferences", *Journal of Economic Literature*, 448-474.
Delfgaauw J., Dur R., Sol J. and Verbeke W. (2013) "Tournament incentives in the field: Gender differences in the workplace", *Journal of Labor Economics* 31, 305-326.
Dick A.A. (2006) "Nationwide Branching and Its Impact on Market Structure, Quality, and Bank Performance*", *The Journal of Business* 79, 567-592.
Faccio M., Marchica M.-T. and Mura R. (2012) "CEO gender, corporate risk-taking, and the efficiency of capital allocation", *Available at SSRN 2021136*.
Fahlenbrach R. and Stulz R.M. (2011) "Bank CEO incentives and the credit crisis", *Journal of financial economics* 99, 11-26.
Farre-Mensa J. and Ljungqvist A. (2013). Do Measures of Financial Constraints Measure Financial Constraints? : National Bureau of Economic Research.
Forssbæck J. (2011) "Ownership structure, market discipline, and banks' risk-taking incentives under deposit insurance", *Journal of Banking & Finance* 35, 2666-2678.
Gagliarducci S. and Paserman M.D. (2012) "Gender interactions within hierarchies: evidence from the political arena", *The Review of Economic Studies* 79, 1021-1052.
Gneezy U., Niederle M. and Rustichini A. (2003) "Performance in competitive environments: Gender differences", *The Quarterly Journal of Economics* 118, 1049-1074.
Gneezy U. and Rustichini A. (2004) "Gender and competition at a young age", *American Economic Review*, 377-381.
Gormley T.A. and Matsa D.A. (2014) "Common errors: How to (and not to) control for unobserved heterogeneity", *Review of Financial Studies* 27, 617-661.
Guiso L. and Rustichini A. (2012) "What drives women out of entrepreneurship? The joint role of testosterone and culture".
Günther C., Ekinci N.A., Schwieren C. and Strobel M. (2010) "Women can't jump?—An experiment on competitive attitudes and stereotype threat", *Journal of Economic Behavior & Organization* 75, 395-401.
Hogarth R.M., Karelaia N. and Trujillo C.A. (2012) "When should I quit? Gender differences in exiting competitions", *Journal of Economic Behavior & Organization* 83, 136-150.
Houston J.F., Lin C., Lin P. and Ma Y. (2010) "Creditor rights, information sharing, and bank risk taking", *Journal of financial economics* 96, 485-512.
Huang J. and Kisgen D.J. (2013) "Gender and corporate finance: Are male executives overconfident relative to female executives?", *Journal of financial economics* 108, 822-839.
Jayaratne J. and Strahan P.E. (1996) "The finance-growth nexus: Evidence from bank branch deregulation", *The Quarterly Journal of Economics* 111, 639-670.

Jayaratne J. and Strahan P.E. (1998) "Entry restrictions, industry evolution, and dynamic efficiency: Evidence from commercial banking", *JL & Econ.* 41, 239.
Johnson C.A. and Rice T. (2008) "Assessing a decade of interstate bank branching", *Wash. & Lee L. Rev.* 65, 73.
Keeley M.C. (1990) "Deposit insurance, risk, and market power in banking", *American Economic Review* 80, 1183-1200.
Khanna N. and Tice S. (2000) "Strategic responses of incumbents to new entry: the effect of ownership structure: capital structure, and focus", *Review of Financial Studies* 13, 749-779.
Laeven L. and Levine R. (2009) "Bank governance, regulation and risk taking", *Journal of financial economics* 93, 259-275.
Martinez-Miera D. and Repullo R. (2010) "Does competition reduce the risk of bank failure?", *Review of Financial Studies* 23, 3638-3664.
Matsa D.A. and Miller A.R. (2013) "A female style in corporate leadership? Evidence from quotas", *American Economic Journal: Applied Economics* 5, 136-169.
Matutes C. and Vives X. (2000) "Imperfect competition, risk taking, and regulation in banking", *European Economic Review* 44, 1-34.
Morin L.-P. (2013) "Do Men and Women Respond Differently to Competition? Evidence from a Major Education Reform", *Journal of Labor Economics* Forthcoming.
Niederle M. and Vesterlund L. (2007) "Do Women Shy Away From Competition? Do Men Compete Too Much?", *The Quarterly Journal of Economics* 122, 1067-1101.
Ors E., Palomino F. and Peyrache E. (2013) "Performance Gender Gap: Does Competition Matter?", *Journal of Labor Economics* 31, 443-499.
Pirinsky C. and Wang Q. (2006) "Does corporate headquarters location matter for stock returns?", *The Journal of Finance* 61, 1991-2015.
Rey-Biel P. and Iriberri N. (2013) "Let's (not) talk about sex: The effect of information provision on gender differences in performance under competition".
Rice T. and Strahan P.E. (2010) "Does Credit Competition Affect Small- Firm Finance?", *The Journal of Finance* 65, 861-889.
Shurchkov O. (2012) "Under pressure: gender differences in output quality and quantity under competition and time constraints", *Journal of the European Economic Association* 10, 1189-1213.
Tate G. and Yang L. (2014) "Female leadership and gender equity: Evidence from plant closure", *Journal of financial economics*.
Zingales L. (1998) "Survival of the Fittest or the Fattest? Exit and Financing in the Trucking Industry", *The Journal of Finance* 53, 905-938.

Appendix A: Tables

Table 5.1 Summary statistics

Return on assets (ROA) is the ratio of earnings before interest, taxes, depreciation, and amortization (EBITDA) to the book value of total assets. Market to book is the ratio of the market value of equity to the book value of equity. *z*-score is computed as the sum of ROA and capital to assets divided by the standard deviation of ROA. Capital to assets is the ratio of equity to total assets. Ln (Sales) is the logarithm of firm revenues. Cash holdings represent the ratio of cash and marketable securities to total assets. Debt to assets is the ratio of financial debt to total assets. Tangibility is one minus the ratio of tangible assets to total assets. Female EXECUTIVES is a dummy equal to one for female EXECUTIVESs, and zero otherwise. Ln (EXECUTIVES age) is the natural logarithm of EXECUTIVES age. Additional details on the construction of each variable are reported in Table A2.

	Number of observations	Average	Median	Standard deviation
ROA	847	0.0357	0.0297	0.0271
Market to book	853	2.5312	1.8899	1.9914
z-score	809	3.8999	3.9536	0.9047
σ (ROA)	812	0.0042	0.0023	0.0062
Capital to assets	851	0.0986	0.0845	0.0714
Risk-adjusted capital ratio	708	0.1302	0.1243	0.0240
Ln (Sales)	853	7.2438	7.0281	1.3320
Cash holdings	852	0.0659	0.0411	0.0826
Debt to assets	588	0.1584	0.1249	0.1421
Tangibility	779	0.9765	0.9850	0.0289
Female EXECUTIVES	854	0.1416	0	0.3489
Ln (EXECUTIVES age)	854	4.1066	4.1271	0.1195

Table 5.2 Competitive pressures and female-led firm performance

This table reports results from the OLS regressions. The dependent variable in Columns (1)-(3) is the return on assets (ROA) measured as the ratio of EBITDA to the book value of total assets, whereas in Columns (4) and (5) it is the ratio of market value of equity to the book value of equity. Female Executives is a dummy equal to one if a firm's executive is a woman, and zero otherwise. RS index is a variable (ranging between zero and four) measuring the number of roadblocks erected by a state in a given year against the IBBEA provisions (see Table A1 for details on the construction of the index). Other explanatory variables are described in Table 1. Each regression includes state and year fixed effects as well as the interaction between SIC and year dummies. Standard errors are clustered by states of headquarter. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

Dependent variable:	ROA			Market to book	
	(1)	(2)	(3)	(4)	(5)
Female Executives	-0.0021	-0.0031	-0.0041*	-0.6271	-0.7910*
	(0.0032)	(0.0019)	(0.0022)	(0.5131)	(0.4588)
Female Executives ×RS index	0.0029*	0.0029***	0.0029**	0.3501*	0.4295**
	(0.0015)	(0.0009)	(0.0011)	(0.2002)	(0.2012)
RS index	-0.0011	-0.0008	-0.0011	-0.1415	-0.1137
	(0.0012)	(0.0011)	(0.0013)	(0.1247)	(0.1687)
Ln (Sales)	0.0014	0.0013	0.0018	-0.0906	-0.0634
	(0.0017)	(0.0014)	(0.0015)	(0.2009)	(0.2274)
Cash holdings		0.0557*	0.0503	7.7644**	7.3647*
		(0.0306)	(0.0327)	(3.5627)	(3.7275)
Debt to assets		-0.0315	-0.0352	2.2989	1.9543
		(0.0257)	(0.0289)	(1.4257)	(1.5549)
Tangibility		-0.0533	-0.0616	13.6547*	15.0096*
		(0.1150)	(0.1308)	(8.0551)	(6.9167)
Ln (Executives age)			0.0078		-1.5109
			(0.0139)		(1.2945)
Year fixed effects	✓	✓	✓	✓	✓
State fixed effects	✓	✓	✓	✓	✓
SIC×year fixed effects	✓	✓	✓	✓	✓
Number of observations	847	506	440	512	446

Table 5.3 Allowing for heterogeneous effects

This table reports results from the OLS regressions. The dependent variable in Column (1) is the return on assets (ROA) measured as the ratio of EBITDA to the book value of total assets, whereas in Column (2) is the ratio of market value of equity to the book value of equity. Female Executives is a dummy equal to one if a firm's executive is a woman, and zero otherwise. RS index is a variable (ranging between zero and four) measuring the number of roadblocks erected by a state in a given year against the IBBEA provisions (see Table A1 for details on the construction of the index). Other explanatory variables are described in Table 1. Each regression includes state and year fixed effects as well as the interaction between SIC and year dummies. Standard errors are clustered by states of headquarter. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

Dependent variable:	ROA	Market to book
	(1)	(2)
Female Executives	-0.0049**	-0.9470**
	(0.0023)	(0.4044)
Female Executives ×RS index	0.0033***	0.5115**
	(0.0010)	(0.2303)
RS index	-0.0141	-9.7760*
	(0.1302)	(5.6789)
Ln (Executives age)	0.0049	-1.1155
	(0.0136)	(1.1032)
Ln (Sales)	0.0002	-0.1159
	(0.0014)	(0.1288)
Cash holdings	0.0294	2.9918
	(0.0199)	(2.1413)
Debt to assets	0.0117	1.5352
	(0.0152)	(1.2219)
Tangibility	0.0074	10.1224*
	(0.1257)	(5.6506)
Ln (Sales)×RS index	0.0016	0.2003
	(0.0017)	(0.2187)
Cash holdings×RS index	0.0029	1.1243
	(0.0305)	(3.0019)
Debt to assets×RS index	-0.0590	-1.1117
	(0.0436)	(2.7363)
Tangibility×RS index	-0.0881	-4.4272
	(0.1500)	(13.7372)
Year fixed effects	✓	✓
State fixed effects	✓	✓
SIC×year fixed effects	✓	✓
Number of observations	440	446

Table 5.4 Effect of competitive pressures on female executives appointments

This table reports results from the OLS regressions. The dependent variable is a dummy equal to one if a firm's executive is female. RS index is a variable (ranging between zero and four) measuring the number of roadblocks erected by a state in a given year against the IBBEA provisions (see Table A1 for details on the construction of the index). Other explanatory variables are described in Table 1. The regression includes state and year fixed effects as well as the interaction between SIC and year dummies. Standard errors are clustered by states of headquarter. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

Dependent variable: Female Executives	
RS index	-0.0038
	(0.0421)
Ln (Sales)	0.0559*
	(0.0281)
Year fixed effects	✓
State fixed effects	✓
SIC×year fixed effects	✓
Number of observations	854

5.5 Placebo tests

This table reports results from the OLS regressions. The dependent variable in Columns (1)-(3) is the return on assets (ROA) measured as the ratio of EBITDA to the book value of total assets. Female Executive is a dummy equal to one if a firm's executive is a woman, and zero otherwise. RS index is a variable (ranging between zero and four) measuring the number of roadblocks erected by a state in a given year against the IBBEA provisions (see Table A1 for details on the construction of the index). Other explanatory variables (coefficients unreported) are that same as those included in Table 2, Column (3), and are described in Table 1. Each regression includes state and year fixed effects as well as the interaction between the SIC and year dummies. Column (1) provides results obtained using firms in SIC codes between 2000 and 4000. Column (2) provides results obtained using changes of the RS index in random years. Column (3) provides results randomly assigning executives' gender to firms. Standard errors are clustered by states of headquarter. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

Dependent variable: ROA			
	Manufacturing Firms	Placebo RS years	Placebo Executives gender
	(1)	(2)	(3)
Female Executives	-0.0103	0.0011	0.0011
	(0.0102)	(0.0042)	(0.0017)
Female Executives ×RS index	-0.0004	0.0009	-0.0011
	(0.0058)	(0.0021)	(0.0009)
RS index	0.0025	0.0002	-0.0000
	(0.0034)	(0.0007)	(0.0010)
Year fixed effects	✓	✓	✓
State fixed effects	✓	✓	✓
SIC×year fixed effects	✓	✓	✓
Controls	✓	✓	✓
Number of observations	3449	390	459

Table 5.6 Robustness

This table reports results from the OLS regressions. Unless specified otherwise, the dependent variable is the return on assets (ROA) measured as the ratio of EBITDA to the book value of total assets. Female Executive is a dummy equal to one if a firm's executive is a woman, and zero otherwise. RS index is a variable (ranging between zero and four) measuring the number of roadblocks erected by a state in a given year against the IBBEA provisions (see Table A1 for details on the construction of the index). Other explanatory variables (coefficients unreported) are the same as those included in Table 2, Column (3), and are described in Table 1. Each regression includes state and year fixed effects as well as the interaction between the SIC and year dummies. Unless specified otherwise, standard errors are clustered by states of headquarter. As an additional control, Column (1) includes the ROA in the first year the firm enters our sample. Column (2) controls for the (logarithm of) years since a state deregulated interstate banking and intrastate branching. Column (3) restricts the analysis to firms that are present in the sample for at least 5 years. Column (4) clusters residuals by firm. Column (5) provides results using heteroskedasticity-adjusted (unclustered) residuals. Column (6) drops an additional 1% of observations on the right and left tails of the ROA distribution. Column (7) uses the logarithm of 1 plus ROA as a dependent variable. Column (8) includes one-year lagged firm age, capital to labor ratio, board size, executive seniority, a dummy equal to one for urban banks (i.e. located in one of the ten largest US metropolitan areas as of 2000) and the ratio of independent to total directors as additional explanatory variables. Column (9) controls for a number of state-year changes in the income tax that affected banks and other financial institutions over the period considered (data from Farre-Mensa and Ljungqvist 2013). Column (10) excludes firms headquartered in Delaware and South Dakota. Column (11) includes state-level macroeconomic controls such as the logarithm of GDP per capita and population size. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Female Executives	-0.0014	-0.0035	-0.0030	-0.0041*	-0.0041**	-0.0044*	-0.0039*	-0.0027	-0.0043*	-0.0034	-0.0033*
	(0.0020)	(0.0022)	(0.0018)	(0.0023)	(0.0020)	(0.0025)	(0.0021)	(0.0026)	(0.0023)	(0.0026)	(0.0026)
Female Executives ×RS index	0.0019**	0.0029**	0.0021***	0.0029*	0.0029**	0.0029**	0.0028**	0.0027**	0.0030**	0.0026**	0.0029***
	(0.0007)	(0.0011)	(0.0007)	(0.0016)	(0.0014)	(0.0012)	(0.0011)	(0.0011)	(0.0011)	(0.0013)	(0.0013)
RS index	-0.0012	-0.0005	-0.0003	-0.0011	-0.0011	-0.0012	-0.0010	-0.0019	-0.0011	-0.0006	-0.0000
	(0.0010)	(0.0010)	(0.0009)	(0.0013)	(0.0011)	(0.0014)	(0.0013)	(0.0012)	(0.0013)	(0.0013)	(0.0010)
Year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
State fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SIC×year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Number of observations	435	440	373	440	440	436	440	399	440	419	440

Table 5.7 Competitive pressures and female-led firm risk

This table reports results from the OLS regressions. The dependent variable in Column (1) is the *z*-score, computed as the sum of ROA and capital to assets divided by the standard deviation of ROA; in Column (2) it is the volatility of firm ROA; in Column (3) it is the volatility of the market to book ratio; in Column (4) it is the capital to assets ratio, computed as the ratio of equity to total assets; in Column (5) is the ratio of combined (Tier1 and Tier2) risk-adjusted capital ratio. Female Executive is a dummy equal to one if a firm's executive is a woman, and zero otherwise. RS index is a variable (ranging between zero and four) measuring the number of roadblocks erected by a state in a given year against the IBBEA provisions (see Table A1 for details on the construction of the index). Other explanatory variables (coefficients unreported) are the one-year lagged logarithm of firm sales, cash holdings ratio, tangibility and ROA. Details on the construction of these variables are reported in Table A2. Each regression further includes state and year fixed effects as well as the interaction between SIC and year dummies. Standard errors are clustered by states of headquarter. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

Dependent variable:	*z*-score	σ (ROA)	σ (market to book)	Capital to assets	Risk-adjusted capital ratio
	(1)	(2)	(3)	(4)	(5)
Female Executives	0.3364	-0.2793	-0.4614*	0.0085	0.0080
	(0.2621)	(0.2505)	(0.2358)	(0.0054)	(0.0069)
Female Executives ×RS index	-0.1616*	0.1402*	0.2675**	-0.0039*	-0.0055**
	(0.0895)	(0.0828)	(0.1126)	(0.0022)	(0.0027)
RS index	-0.0535	0.0827	0.0952	0.0035**	0.0024**
	(0.0741)	(0.0776)	(0.0729)	(0.0017)	(0.0011)
Year fixed effects	✓	✓	✓	✓	✓
State fixed effects	✓	✓	✓	✓	✓
SIC×year fixed effects	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Number of observations	636	638	635	664	560

Appendix B: Variables and RS index description

Table A1. Description of the RS index

This table illustrates changes in the RS index measuring the regulatory changes in the banking industry. RS index is a variable (ranging between zero and four) measuring the number of roadblocks erected by a state in a given year against the IBBEA provisions (see Table A1 for details on the construction of the index). Data source: Rice and Strahan (2010).

State	RS index	Effective date	Age restriction (years)	*De novo* interstate branching restriction	Single branch acquisition restriction	Statewide deposit cap on branch acquisition
Alabama	3	5/31/1997	5	Yes	Yes	30%
Alaska	2	01/01/1994	3	Yes	No	50%
Arizona	2	8/31/2001	5	Yes	No	30%
Arizona	3	09/01/1996	5	Yes	Yes	30%
Arkansas	4	06/01/1997	5	Yes	Yes	25%
California	3	9/28/1995	5	Yes	Yes	30%
Colorado	4	06/01/1997	5	Yes	Yes	25%
Connecticut	1	6/27/1995	5	No	No	30%
Delaware	3	9/29/1995	5	Yes	Yes	30%
DC	0	6/13/1996	No	No	No	30%
Florida	3	06/01/1997	3	Yes	Yes	30%
Georgia	3	05/10/2002	3	Yes	Yes	30%
Georgia	3	06/01/1997	5	Yes	Yes	30%
Hawaii	0	01/01/2001	No	No	No	30%
Hawaii	3	06/01/1997	5	Yes	Yes	30%
Idaho	3	9/29/1995	5	Yes	Yes	No
Illinois	0	8/20/2004	No	No	No	30%
Illinois	3	06/01/1997	5	Yes	Yes	30%
Indiana	1	07/01/1998	5	No	No	30%
Indiana	0	06/01/1997	No	No	No	30%
Iowa	4	04/04/1996	5	Yes	Yes	15%
Kansas	4	9/29/1995	5	Yes	Yes	15%
Kentucky	3	3/22/2004	No	Yes	Yes	15%
Kentucky	3	3/17/2000	No	Yes	Yes	15%
Kentucky	4	06/01/1997	5	Yes	Yes	15%
Louisiana	3	06/01/1997	5	Yes	Yes	30%
Maine	0	01/01/1997	No	No	No	30%
Maryland	0	9/29/1995	No	No	No	30%
Massachusetts	1	08/02/1996	3	No	No	30%
Michigan	0	11/29/1995	No	No	No	No
Minnesota	3	06/01/1997	5	Yes	Yes	30%
Mississippi	4	06/01/1997	5	Yes	Yes	25%

Missouri	4	9/29/1995	5	Yes	Yes	13%
Montana	4	10/01/2001	5	Yes	Yes	22%
Montana	4	9/29/1995				Increases 1% per year from 18% to 22%
Nebraska	4	5/31/1997	5	Yes	Yes	14%
Nevada	3	9/29/1995	5	Limited	Limited	30%
New Hampshire	0	01/01/2002	No	No	No	30%
New Hampshire	1	08/01/2000	5	No	No	30%
New Hampshire	4	06/01/1997	5	Yes	Yes	20%
New Jersey	1	4/17/1996	No	Yes	No	30%
New Mexico	3	06/01/1996	5	Yes	Yes	40%
New York	2	06/01/1997	5	Yes	No	30%
North Carolina	0	07/01/1995	No	No	No	30%
North Dakota	1	08/01/2003	No	No	No	25%
North Dakota	3	5/31/1997	No	Yes	Yes	25%
Ohio	0	5/21/1997	No	No	No	30%
Oklahoma	1	5/17/2000	No	No	No	20%
Oklahoma	4	5/31/1997	5	Yes	Yes	15%
Oregon	3	07/01/1997	3	Yes	Yes	30%
Pennsylvania	0	07/06/1995	No	No	No	30%
Rhode Island	0	6/20/1995	No	No	No	30%
South Carolina	3	07/01/1996	5	Yes	Yes	30%
South Dakota	3	03/09/1996	5	Yes	Yes	30%
Tennessee	1	3/17/2003	3	No	No	30%
Tennessee	1	07/01/2001	5	No	No	30%
Tennessee	2	05/01/1998	5	Yes	No	30%
Tennessee	3	06/01/1997	5	Yes	Yes	30%
Texas	2	09/01/1999	No	No	No	20%
Texas	4	8/28/1995	N/A	N/A	N/A	20%
Utah	1	4/30/2001	5	No	No	30%
Utah	2	06/01/1995	5	Yes	No	30%
Vermont	0	01/01/2001	No	No	No	30%
Vermont	2	5/30/1996	5	Yes	No	30%
Virginia	0	9/29/1995	No	No	No	30%
Washington	1	05/09/2005	5	No	No	30%
Washington	3	06/06/1996	5	Yes	Yes	30%
West Virginia	1	5/31/1997	No	No	No	25%
Wisconsin	3	05/01/1996	5	Yes	Yes	30%
Wyoming	3	5/31/1997	3	Yes	Yes	30%

Table A2. Variable description

Name	Description	Source
ROA	(*ebitda*/*at*), where *ebitda* are earnings before interest, taxes, depreciation and amortization, and *at* measures total assets. We drop 1% of the observations at each tail of the distribution to mitigate the effect of outliers.	Compustat
Market to book	(*prcc_f*×*cshtr_f*)/*ceq*, where *prcc_f* is the market price of common shares at the end of the fiscal year, *cshtr_f* is the number of common shares outstanding and *ceq* is the book value of equity. The variable is trimmed between 0 and 10.	Compustat
σ (ROA)	Standard deviation of (*ebitda*/*at*) computed over three years (1997-1999, 2000-2002 and 2003-2006).	Compustat
σ (market to book)	Standard deviation of (*prcc_f*×*cshtr_f*)/*ceq*, where *prcc_f* is the market price of common shares at the end of the fiscal year, *cshtr_f* is the number of common shares outstanding and *ceq* is the book value of equity computed over three years (1997-1999, 2000-2002 and 2003-2006).	Compustat
Capital to assets	*ceq*/*at*, where *ceq* is the book value of a firm's equity and *at* are total assets. Values outside the [0, 1] range are excluded.	Compustat
Risk-adjusted capital ratio	Ratio of Tier1 and Tier2 capital to total assets.	Compustat bank
z-score	(ROA×*capital to assets*)/σ(ROA), where ROA is computed as (*ebitda*/*at*) and capital to assets is computed as *ceq*/*at*, i.e., the ratio of equity to total assets. σ(ROA) is the standard deviation of (*ebitda*/*at*) computed over three years (1997-1999, 2000-2002 and 2003-2006).	Compustat
Ln (Sales)	ln(*sale*), where *sale* measures a firm's revenues.	Compustat
Tangibility	1-(*intan*/*at*), where *intan* are intangible assets and *at* are total assets. Values outside the [0, 1] range are excluded.	Compustat
Cash holdings	(*che*/*at*), where *che* are cash and marketable securities and *at* are total assets. Values outside the [0, 1] range are excluded.	Compustat
Debt to assets	(*dlc*+*dltt*)/*at*, where *dlc* is the size of financial debt due in one year, *dltt* is the size of long-term financial debt and *at* is the size of assets. Values outside the [0, 1] range are excluded.	Compustat
Year fixed effects	Full set of year dummies.	Compustat
State fixed effects	Full set of state-of-headquarter dummies, where headquarter corresponds to the last state of headquarter as reported in Compustat	Compustat

SIC fixed effects	Full set of 4-digit SIC dummies.	Compustat
RS index	Index ranging between zero and four measuring the number of roadblocks erected by a state in a given year against the IBBEA provisions (see Table A1 for a list of roadblocks introduced/removed by state and year).	Rice and Strahan (2010)
Female Executives	Dummy equal to one if the executive is female, and zero otherwise.	IRRC
Female directors	Dummy equal to one if the firm has at least one female director, and zero otherwise.	IRRC
Ln (Executives age)	Ln (*Executives age*), where *Executives age* is the age of a CEO measured in years.	IRRC

CHAPTER VI

CONCLUSIONS

Motivated by the introduction of new policies that aim to increase the presence of women in leadership positions, an emerging literature explores gender differences in economic behavior and their implications for organizations and the labor market. My dissertation provides novel findings on the role played by gender, and in particular gender pairings, using laboratory experiments and empirical data.

In the first two chapters, using a laboratory experiment, we investigate on the role played by gender interactions in an accountable setting. In particular, we study how the compensation scheme, and the gender pairing between a decision-maker and an audience that calls for a justification of the decision-maker's choices, influence her/his level of cognitive effort. The main findings show that both male's and female's behavior is influenced by the audience's gender and the compensation scheme. Females' effort provision is positively influenced when they are paired with an audience of the opposite gender and the when the compensation scheme includes incentives and payoff sharing; by contrast, when male are paired with an audience of the opposite gender and their compensation scheme includes either incentives or not, their effort provision significantly decrease.

In the last two chapters, we use empirical data to study the interactions between gender and firms' specific characteristics, and between gender and organization's external context. In particular, we study how the gender interactions in the top layers of a firm's hierarchy affect corporate performance; and how women and men executives respond to an exogenous increase in competitive pressures in the banking sector. Our main findings show that an increasing fraction of women on the board of directors significantly improves the profitability of firms led by female CEOs. Concerning the interactions between gender and external context, we find that when the competitive pressure increases, female-led banks experience lower performance than their counterparts; however, despite the lower profitability, female leadership mitigates the potential exacerbating effect on competition on risk-taking.