Experimental Economics (2023) 26:879–914
https://doi.org/10.1007/s10683-023-09796-9

ORIGINAL PAPER

Gender differences in alternating-offer bargaining: an experimental study

Iñigo Hernandez-Arenaz[1] · Nagore Iriberri[2]

Received: 2 January 2022 / Revised: 16 February 2023 / Accepted: 16 February 2023 /
Published online: 17 April 2023

Abstract

A laboratory study was carried out to analyze the relationship between ambiguity regarding the sharing norms in structured alternating-offer bargaining and gender differences in bargaining. Symmetric environments, where a 50:50 split emerges as the unique sensible norm, showed the lowest ambiguity and gender differences are absent. We increased ambiguity by introducing asymmetries into the bargaining environment by making one bargaining party get a higher share than the other (due to empowerment, entitlement or informational asymmetries), but without imposing new sharing norms. In these situations, men are less likely to reach an agreement, but, when they do, they obtain a larger share of the pie. As a result, men and women show similar overall earnings but earnings are lower when bargaining *with* men. We find suggestive evidence that gender differences diminish when we reduce ambiguity regarding the sharing norms by providing information about other participants' agreements in asymmetric environments.

Keywords Bargaining · Gender differences

JEL Classification codes J16 · D91 · C75

The authors would like to thank Antonio Cabrales, Javier Gardeazabal, Pedro Rey-Biel, anonymous referees, and participants of seminars at various universities for helpful comments. Iñigo Hernandez-Arenaz acknowledges the financial support provided by Vicerrectorado de Investigación de la UPV/EHU (PIF//13/015), Departamento de Educación, Política Lingüística y Cultura del Gobierno Vasco (IT869-13), Ministerio de Ciencia, Innovación y Universidades (PID2019-108343GA-I00), and Grant PID2021-127119NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe". Nagore Iriberri acknowledges the financial support from Grant PID2019-106146GB-I00, funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe", Departamento de Educación and Política Lingüística y Cultura del Gobierno Vasco (IT1697-22), and the Norwegian Research Council (TOPPFORSK 250506). The replication material for the study is available at https://osf.io/9wu2e/ (DOI https://doi.org/10.17605/OSF.IO/9WU2E).

✉ Nagore Iriberri
nagore.iriberri@ehu.eus

Extended author information available on the last page of the article

1 Introduction

The gender wage gap has long been a major subject for study in economics. Although it has shown a decreasing trend over time, its persistence in developed countries challenges classical explanations based on differences in human capital, preferences and statistical discrimination (Blau and Kahn, 2000, 2017). Gender differences in negotiation have been put forward as an alternative explanation for the gender wage gap. Starting wages are often the result of bilateral negotiation. Moreover, wages are also affected by negotiations that come later in one's career, e.g., for pay increases. If women are less likely to negotiate starting salaries and to ask for pay increases, and/or if women obtain worse deals when negotiating, this would clearly go some way towards explaining the gender wage gap (Azmat and Petrongolo, 2014; Card et al., 2016; Sin et al., 2020).

The stereotypical behavior in most real life bargaining settings is that men are better bargainers than women and so when differences are found the gender gap tends to be negative for women. In this paper we propose to switch the focus from *whether* there are gender differences to *when* they will be observed. Mazei et al. (2015) offer the most recent meta-analysis on gender differences in negotiation and their moderators, building on the previous work by Stuhlmacher and Walters (1999). Men were found to achieve better outcomes than women, but these gender differences were found to depend on the context. One important moderating factor is what psychologists labeled *structural ambiguity*. Building on Mischel (1977)'s notion of ambiguous (or weak) and unambiguous (or strong) situations, gender differences were mostly found in situations where people did not have a clear protocol or script for appropriate behavior. In these situations, people relied on more general behavioral schemata and available social norms, such as preconceived gender roles and stereotypes (Bowles et al., 2005; Major et al., 1984).[1]

This paper studies the relationship between the existing ambiguity regarding the sharing norms and the existence of gender differences in bargaining. Going back to the opening paragraph; do we expect the same gender differences in wage negotiations when workers know exactly the salary each worker is getting (full transparency) or when workers lack any guidance related to the existing wage distribution (full ambiguity)? We propose using a controlled environment such as the laboratory, to study when gender differences will be observed in structured alternating-offer bargaining environments. The design of the experiment was registered at the *AEA RCT registry*, under the reference

[1] Gender differences in the willingness to start a negotiation have also experienced a similar change of focus. Since the pioneering work by Babcock and Laschever (2009) "Women Don't Ask," Kugler et al. (2018) performed a meta-analysis of existing work in psychology and they concluded that there is ample evidence for the existence of gender differences in the likelihood of starting a negotiation. However, these differences are smaller when situational ambiguity regarding the appropriateness of negotiating is low rather than high, as well as when situational cues are more consistent with the female gender role than with the male gender role (see recent work in economics by Leibbrandt and List, 2015). See also Hernandez-Arenaz and Iriberri (2019) for a review.

AEARCTR-0002029.[2] A laboratory setting allows researchers to study gender differences in bargaining environments that vary in the existing ambiguity regarding the sharing norm, which is our main treatment variable. In addition, the laboratory offers the possibility of measuring individuals' self-assessment of their ability to perform a task and to bargain, as well as their risk and social preferences, which are hard, if not impossible, to control for when using observational data. Gender differences in those behavioral traits are known to be the mediating factor for gender differences (see Niederle et al. 2011; Gillen et al. 2019; Van Veldhuizen, 2022).

We hypothesize that, ceteris paribus, the higher the ambiguity regarding the existing sharing norm, the more likely and stronger will be the gender differences.

We use a symmetric bargaining setting as a benchmark, where bargaining parties show equal strength so that a 50:50 split of the pie is the only expected sharing norm. The benchmark shows a bargaining setting with the lowest ambiguity. We hypothesize that in this benchmark setting, participants will follow the norm so that no gender differences appear.

We then modify the symmetric environment to introduce three common sources of asymmetries existing in the real world: empowerment (only the proposer has a positive outside option), entitlement (the proposer is entitled to a greater share than the responder), and informational asymmetry (only the proposer knows the actual size of the pie).[3] We chose these particular asymmetries because they are present in many economic-relevant situations such as in salary negotiations. For example, whenever any of the bargaining roles has an outside option (an employer with multiple potential employees or employees with multiple job offers) empowerment will be in play. In situations with a feeling of ownership of the surpluses on which the participants are negotiating, such as in negotiations about performance based promotions, entitlement is in play. Finally, employees almost never know the exact size of the salary or promotion that is attainable, generating an informational asymmetry. The existence of these asymmetries not only makes one bargaining party stronger (the *proposer* in our setting) and the other weaker (the *responder* in our setting), but also increases the ambiguity regarding what one could expect as the bargaining outcome. To put it simply, in all three asymmetric environments, the proposer is expected to get more than the responder, but it is not clear how much more. We hypothesize that these environments would be the ones in which gender differences in bargaining are likely to flourish. Lastly, in a final treatment variation, we aim to maintain the asymmetry in the bargaining environment but reduce ambiguity by providing participants with the modal agreements of other participants in past experimental sessions. We hypothesize that, if anything, gender differences should decrease when ambiguity is reduced.

Our laboratory study consisted of three main tasks. Subjects first performed a real effort task, where each subject obtained a score for productivity which then

[2] The pre-plan analysis can be checked at https://www.socialscienceregistry.org/trials/2029/history/15499. Additional treatments with past agreements, see Fig. 1, were not part of the pre-plan analysis but were added after suggestions by referees.

[3] Kagel et al. (1996) do a similar manipulation on informational asymmetries.

determined the pie to be shared. In the second task, subjects were randomly paired and had 3 minutes to bargain over the pie via alternating-offer. The bargaining task consisted of 10 bargaining periods of 3-minutes each with a different paired participant each time. Finally, in the third task, we elicited a set of beliefs to measure their self-assessed ability in the task and in bargaining, as well as risk and social preferences.

The laboratory design relied on random pairing of individuals to form the pairs that will bargain over a pie, and on men and women being ex-ante equally likely to be allocated to either the strong or the weak bargaining position. This design allowed us to study two main important questions on gender differences when bargaining. Firstly, we studied gender differences in three interconnected bargaining outcomes (probability of reaching an agreement, earnings conditional on reaching an agreement and overall earnings) in the symmetric and the asymmetric bargaining environments, with and without information about past agreements. Secondly, we tested whether men and women react differently to the presence of asymmetries and to the presence of information about past agreements in asymmetric bargaining environments, i.e., whether gender is an effect modifying factor. To do this, we compared gender differences in each asymmetric environment with those in the symmetric environment, together with gender differences in the asymmetric environments with and without information about past agreements.

In the symmetric bargaining benchmark, as expected, we find that the 50:50 split is largely followed. Indeed 69.1% of successful negotiations end up with the pie being split exactly equally, showing to be the bargaining setting with the lowest ambiguity. As conjectured, we find no hard evidence for gender differences. When asymmetries are introduced, we find important gender differences in the stereotypically expected direction. Firstly, men show a lower likelihood of reaching an agreement, a result that in our opinion has not been emphasized enough in the literature given its implications for efficiency, especially in the responder's role. Secondly, when an agreement is reached, men show the ability to secure a higher share of the pie, especially in the proposer's role. These differences are consistent with most findings both in economics and psychology. As these two differences have the opposite effect on overall earnings, it turns out that men and women do not show significant differences in either role when focused on overall earnings. However, overall earnings are lower when bargaining *with* men. When comparing gender differences across the three different bargaining asymmetric environments, we find that gender differences are strongest in those with empowerment and informational asymmetries. These are the environments in which ambiguity is highest. Third and finally, when past agreements are provided in the asymmetric bargaining environments, the fact that men show a lower likelihood of agreement is no longer significant, attenuating in part the existing gender differences.

With regard to gender as effect modifying factor, we compare gender differences in each asymmetric bargaining environment with those in the symmetric bargaining environment. Despite not finding significant results, we find all coefficients going in the expected direction, increasing gender differences when increasing ambiguity. Furthermore, when comparing gender differences in asymmetric environments with and without past agreements, again, although in the expected direction, we do not find hard evidence that gender is an effect modifying factor for the provision of past agreements. Finally, when we measure ambiguity in a continuous way using the distance between a particular split and the mean or modal split and interact gender differences with these ambiguity measures, we again find suggestive evidence that the higher the ambiguity the bigger the gender gap, although not always significant.

We further show three important robustness tests. First, we replicate the main analysis including individual level controls and, as expected, we find that the existing gender differences get attenuated when controlling for differences in confidence, stereotypical beliefs, and risk preferences (consistent with Niederle et al., 2011; Gillen et al., 2019; Van Veldhuizen, 2022). Second, as experimental subjects were represented using gender avatars in the bargaining environment, experimenter demand effects may be at play. Having this concern in mind, we included a question at the end of the experiment (*What was the purpose of the experiment?*) to measure how many subjects indeed identified gender differences as the object of the study. In the robustness test, we show that this percentage is low (below 8% of the subjects) and we replicate the main analysis excluding those negotiations in which participated subjects who mentioned gender was the object of study, and the main results hold, although again the size and significance decreases, some of the results becoming now insignificant. However, notice that the number of observations also decreases significantly. We acknowledge this percentage may represent a lower bound and that therefore this offers an imperfect control for potential experimenter demand effects. Third and finally, we also observe an important deadline effect, as about a quarter of the negotiations were still going on in the last 10 seconds. We replicate the main analysis excluding those negotiations and the main results hold.

The rest of the paper is organized as follows. Section 2 places our paper and results in the context of existing work. Section 3 describes the procedures and design of the laboratory experiment, the data, the identification strategy, and the hypotheses. Section 4 describes the main results. Section 5 concludes.

2 Literature review and contribution

Gender differences in bargaining have been studied by economists. For example, gender differences in negotiation were found by studying male proposers' behavior in field experiments in which the gender of potential scripted buyers varied (Ayres, 1991; Ayres and Siegelman, 1995; Castillo et al., 2013).[4] To study gender differences in wage negotiation, Säve-Söderbergh (2019) and Roussille (2021) used wage bids and wage offers and find evidence for substantive gender gap in wages. Andersen et al. (2018) found that men obtain better deals than women among the members of a patriarchal society while the reverse was true for a matriarchal society.

[4] In particular, Ayres (1991) and Ayres and Siegelman (1995), both found that women obtain worse deals when buying a car while Castillo et al. (2013) found that women obtain better fares in taxi-rides. In a related literature, Andersen et al. (2021) and Goldsmith-Pinkham and Shue (2020), also study gender differences in the housing market. Both studies show that observed gender differences can be mostly explained by men and women either demanding different types of properties or by having different transaction timing and location. These differences are unlikely to play a role in a laboratory setting.

Biasi and Sarsons (2022) compared the wage gap in salaries among teachers in Wisconsin after a law change from collective bargaining to flexibility in bargaining for individual salaries. They found that this increased the gender gap in wages.

Economists have also studied gender differences in controlled settings such as laboratories, mostly using the ultimatum game, which represents a reduced-form bargaining setting, as it allows for a single offer (or demand) and the response to it. Using face-to-face ultimatum games, Eckel and Grossman (2001) found that women are more likely to accept offers from women (*solidarity*) and that men are more likely to accept offers from women (*chivalry*). In an ultimatum game where gender is commonly known Solnick (2001) found that women are more likely to accept offers from male proposers than from female proposers. Sutter et al. (2009) found much more competition and retaliation and, thus, lower efficiency when the bargaining partners had the same gender than when they had the opposite gender. Huang and Low (2022) showed that gender differences can reverse when negotiating in a battle-of-the-sexes type setting when participants can use verbal communication as opposed to no communication.

Closer to our alternating-offer bargaining setting, Dittrich et al. (2014) used a laboratory face-to-face alternating-offer wage-bargaining game where the firm was empowered, and found that starting salaries offered by men to women were lower than those offered by women to men, resulting in significant gender interaction effects on wage-bargaining outcomes. Using data from a TV-show in which bargaining parties showed major asymmetries in all three dimensions (empowerment, entitlement and information), Hernandez-Arenaz and Iriberri (2018) found that the pairing between a male proposer (strong) and a female responder (weak) was the only one that differed from the rest, yielding higher profits for the proposer. Contrary to our findings here, they found significant interaction effects. However, this may be because their settings have more than one type of asymmetry simultaneously. Rigdon (2012) found that women demand less than men in a demand-ultimatum-game in the laboratory, and more interestingly, that this gap diminishes when previous demands are provided, similar to our treatment to reduce ambiguity. More recently, when studying gender differences in the choice to negotiate, Exley et al. (2020) included a baseline treatment, where subjects were *forced* to negotiate in an unstructured setting with limited time. They found that men and women achieve similar earnings.

Note that apparently contradictory findings can be rationalized through our hypotheses and results. Those studies that find gender differences in bargaining are those that show greater ambiguity with respect to which sharing norm is adequate. However, Exley et al. (2020) used a setting that, while asymmetric, displayed a clear sharing norm that dictated how the pie should be divided, as bargaining parties knew exactly how much of the pie each bargaining party contributed. Consequently, in line with our hypotheses and results, these studies suggest that gender differences are likely to flourish only in those situations that show enough ambiguity.

Our paper makes three contributions over existing work. Firstly, and most importantly, it proposes a way to determine when gender differences in bargaining can be expected: when ambiguity, proxied by the variance in bargaining earnings, is highest. This is confirmed by our experimental results and is consistent with other findings in the literature, both in economics and psychology. Regarding the contribution over the studies in economics, it is the first study that offers different ways to systematically manipulate the existing ambiguity, which is our main treatment variable, when studying bargaining outcomes in connection to gender differences. Regarding the work in psychology, we offer a framework to think about what *structural ambiguity* means or materializes into, providing the comparison of symmetric (no ambiguity) versus asymmetric environments that lack a clear sharing rule (highest ambiguity), and propose a way in which ambiguity can be reduced (through the provision of past agreements). Secondly, it proposes an experimental framework for studying gender differences in a rich structured bargaining environment such as the alternating-offer bargaining, bringing the environment closer to reality and at the same time being observable to the researcher. In this regard, we find a significant gender difference that has not been stressed enough in our opinion, despite having important consequences for efficiency, that of men showing a lower likelihood of reaching an agreement. This is due to the studied bargaining settings in the laboratory, mostly ultimatum-like settings, showed a reduced-form of bargaining. Finally, this paper offers a rationale for the most recent studies using observational data that show effective policy recommendations on how to reduce the gender gap in wages: transparency. Hospido et al. (2019), Recalde and Vesterlund (2020) and Bennedsen et al. (2022) show that the more transparency with regard to when to apply for promotions and with regard to the salary increase involved in a promotion, the lower the gender gap. Transparency is clearly at the other extreme of ambiguity.

3 Experimental procedures and design

A laboratory experiment was run at the Bilbao Laboratory of Experimental Analysis (Bilbao Labean) at the University of the Basque Country and at the Experimental Economics Lab (LEE) at University University Jaume I, on a computer-based form using z-Tree experimental software (Fischbacher, 2007). Subjects were recruited through ORSEE (Greiner, 2015), with a total of 562 participants—278 (49.4%) men and 284 (50.6%) women—split into sixteen different sessions. Recruiting was carried out in such a way that the gender balance in each session was assured while subjects were unaware of this at the time of recruiting.

At the beginning of each session, subjects were provided with written general instructions, which informed them that the experiment consisted of 3 different tasks and that the detailed instructions would be displayed on their computer screens before the start of each task. All instructions, both written general instructions and detailed instructions regarding each of the tasks, were read aloud to ensure that the information was public knowledge. A translation of the instructions can be found in Online Appendix B. Each session lasted for about one and a half hours, including payment. Average earnings were 15.32 euro (s.d. 5.71) including a show-up fee of 3 euro, and total earnings ranged from 5 euro to 34.5 euro.

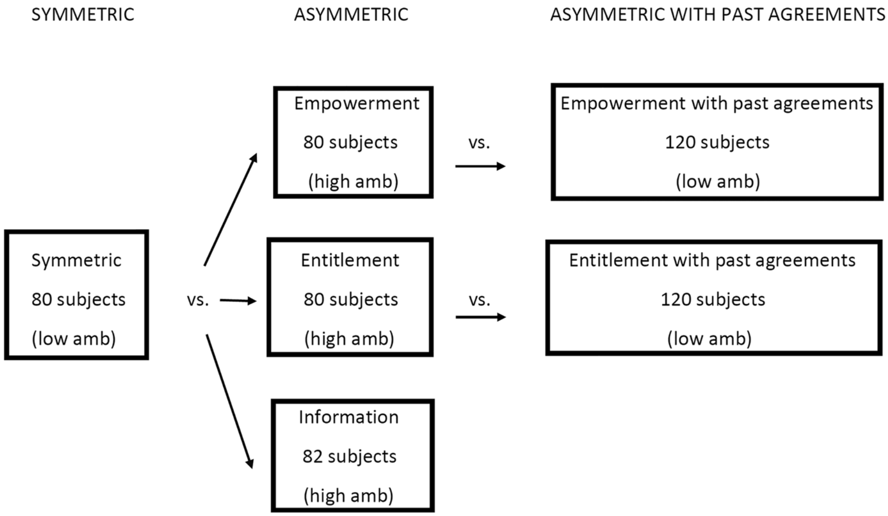

Fig. 1 Treatments: varying ambiguity in the sharing rules

3.1 Design: treatments and time-line of the experiment

3.1.1 Treatments

Figure 1 summarizes the experimental treatments that aim to change the ambiguity regarding the sharing norms. Firstly, these bargaining environments differ from one another in terms of the existence of symmetry and, among the asymmetric bargaining environments, in terms of the source of the induced bargaining asymmetry (through empowerment, entitlement and information). In the symmetric environment, we expected the ambiguity to be lowest, as the only sensible sharing rule is the 50:50 split. In the asymmetric environments, we expected the ambiguity to be highest, as the 50:50 rule is no longer sensible and there is no other sensible sharing rule. In all of these sessions we provided no information regarding what other participants in previous sessions agreed on. Secondly, we aimed to reduce ambiguity with regard to the available sharing rules for *Empowerment* and *Entitlement*, providing subjects with the modal split of the pie in previous sessions. We decided not to carry out additional treatment for the informational asymmetry as it is the lack of information that is the source of the asymmetry, such that providing past agreements may result in canceling out the asymmetry itself.

3.1.2 Time-line of the experiment

All sessions included three different tasks: a real effort task, an alternating-offer bargaining task, and a set of elicitation tasks. The real effort task and the elicitation tasks were identical in all sessions, but we varied the bargaining environment

Fig. 2 Example of a matrix shown to subjects during the real effort task

0	1	1	0	1	1
0	1	1	0	0	1
0	0	0	1	1	1
1	0	0	0	0	0
1	1	0	0	1	0
0	0	0	1	0	0

from one treatment to another, as described in Fig. 1. We now provide further details about each of the bargaining environments.

Real effort task: Subjects were presented with a matrix filled with "0"s and "1"s similar to that in Fig. 2 and asked to count the number of ones.[5] Once a number was entered for a matrix and the subject confirmed the input, a new matrix appeared on the screen. Subjects performed this task for 5 minutes and the performance measure was the total number of matrices for which the correct number of "1"s was provided.[6] This task was not directly incentivized but subjects were informed that their performance in this task was important for determining their earnings in the bargaining task.[7] Consistent with previous findings, this task proved to be gender neutral in performance, with regard to the number of matrices attempted, and the precision rate.[8]

Subjects' gender was elicited at the end of this task, just before taking on the bargaining task. In particular, they were presented with two avatars representing the silhouettes of a man and a woman and explicitly asked "Are you a man or a woman?". As can be seen in Fig. 3, these avatars were chosen to elicit subjects' gender in the most aseptic and neutral way possible, without giving any further cues such as facial expressions. These avatars were used to make bargainers' genders common knowledge, as illustrated by Fig. 4.

5 A similar task was used in Abeler et al. (2011) and Mengel (2015).

6 The z-Tree program was designed such that the maximum number of matrices that could be attempted was 60. This was explicitly stated in the instructions. Data show that this constraint was not binding, as the maximum number of attempted matrices was 45. The average number of attempted matrices was 24.02.

7 As will become clear in the explanation of the bargaining task, the relationship between performance and the pie to be bargained over in the bargaining task may induce competitive attitudes. To preclude any feeling of competition while subjects performed the real effort task, the instructions stated: "The number of correct answers that you provide will determine your productivity. The higher your productivity, the more money, on average, you will have for the next task".

8 Men (278 observations) on average provided the correct number of "1"s in 19.48 matrices (s.d. 4.41), while women (284 observations) in 19.32 (s.d. 4.32). Moreover, this gender neutrality in terms of performance is also present in effort (number of attempted matrices), and precision (number of correct over number of attempted matrices).

Fig. 3 Gender avatars

Bargaining task: symmetric. Based on their relative performances in the real effort task, subjects were assigned a score for *productivity*, which determined the pie to be bargained over. Specifically, the top third of performers were endowed with a productivity of €15 , the middle third with a productivity of €10, and the bottom third with a productivity of €5. Subjects were only given precise details about this protocol after they completed the real effort task, but no information was provided about the actual number of matrices they solved correctly or about their individual productivities.

Each subject was then randomly paired with another subject. One was assigned the role of Participant A (hereafter referred to as the *Proposer*) and the other that of Participant B (hereafter referred to as the *Responder*). The *Proposer* was the paired subject with the higher score in the real effort task, although this protocol was not revealed.[9] Within each pairing, the pie to be bargained over was randomly drawn from the productivity of the proposer and that of the responder with equal probabilities. This means that the pie could be of €5, €10, or €15. Only once the pie size was randomly determined, this information was made public, so the bargaining parties do not know whether they will be bargaining over the pie determined by the proposer's or the responder's productivity. Each pairing had 3 minutes to reach a deal on how to split the pie through an alternating-offer bargaining process. During the bargaining, proposers decided on offers to responders while responders decided on demands from proposers. In other words, the whole bargaining process took place in terms of the amount of money that the responder would get. Proposers started the negotiation making the first offer to the responders. During the bargaining, the information available to all subjects consisted of their own avatar and that of the opponent (their gender and that of their paired partner), the size of the pie to be shared, and the bargaining history of offers and demands. See Fig. 4 for an illustration. Importantly, subjects could not see their own productivity or their opponent's. If they reached a deal within the 3-minute limit, the agreed split was implemented. Otherwise they got 0.

9 Subjects were only told that they would be given a bargaining role. Roles were assigned in this way in order to facilitate comparison across different bargaining environments. In the event of ties, roles were randomly assigned.

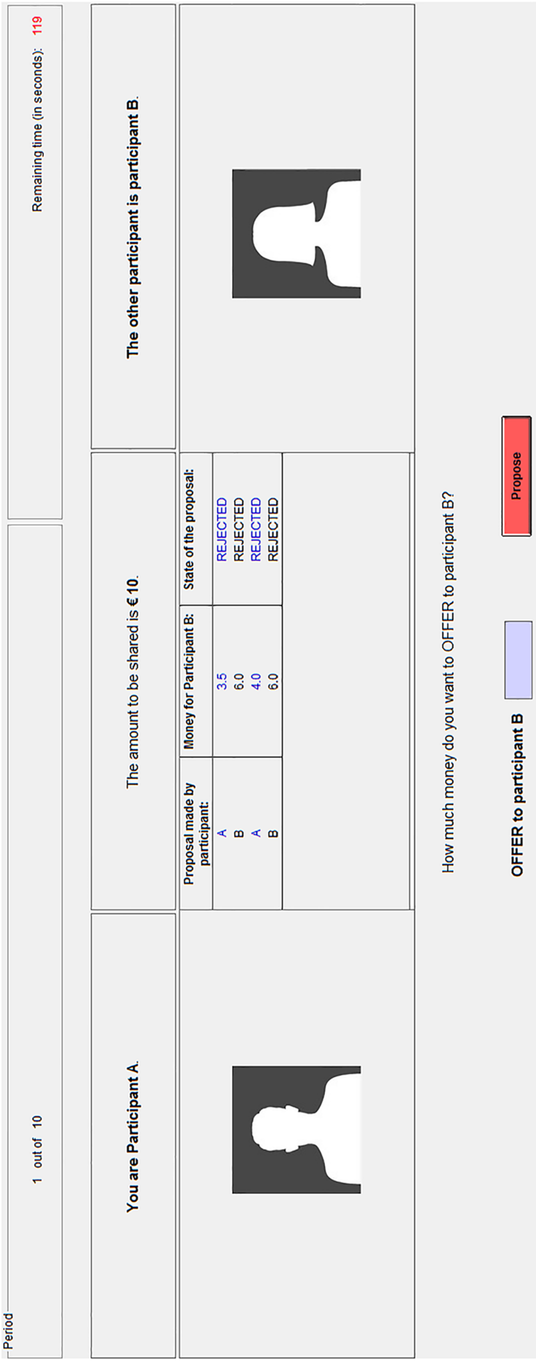

Fig. 4 Screen seen by Proposers during the bargaining task (Symmetric Environment)

The whole bargaining process was repeated for 10 periods in all treatments, with a different paired participant each time.[10] Importantly, from one period to the next the role in the bargaining pairing (proposer or responder) and the pie to be split could change. For payment, subjects were informed that the computer would take two periods randomly—one from periods 1–5 and another from periods 6–10—and the resulting outcomes would be implemented.

Bargaining task: empowerment. Everything was the same as for the *Symmetric* bargaining, except that there was an outside option for the proposer. In particular, if a deal was not reached within the 3-minute limit, the proposer had an outside option while the responder got 0. The outside option available to the proposer was a random amount drawn from a uniform distribution between 50% and 85% of the pie. Both parties knew about the outside option but neither knew its exact value when bargaining.

Bargaining task: empowerment with past agreements. Everything was the same as for *Empowerment* bargaining, except that we provided subjects with past agreements, i.e., the most frequent amount (mode) agreed for the responder in the sessions with empowerment. These amounts depended on the pie to be shared: 1, 1.5 and 5 euro, when the pies were 5, 10 and 15 euro, respectively. This information was presented to subjects during the negotiation just above the dialog box about the offer/demand.

Bargaining task: entitlement. Everything was the same as for *Symmetric* bargaining, except that subjects were able to see their own productivity and that of their partners. This was public knowledge. This bargaining environment thus informed subjects of whose productivity determined the size of the pie. This was intended to generate a feeling of entitlement.[11] In the event of a tie, there is no entitlement effect, meaning that we do not consider those bargaining pairings in the analysis in the rest of the paper (note the lower number of observations in the entitlement treatment).[12]

Bargaining task: entitlement with past agreements. Everything was the same as for *Entitlement* bargaining, except that we provided subjects with past agreements, i.e., the most frequent amount (mode) agreed for the responder in the sessions with

[10] In one of the *Symmetric* sessions there was a technical problem and the z-Tree program stopped at the second repetition. We ran the bargaining module again and everything worked fine the second time. Thus, we gathered data from 12 bargaining periods for the *Symmetric* environment, instead of 10 but, given that periods 1 and 3 and periods 2 and 4 involve exactly the same pairings, we only considered periods 1–2 and 5–12 when analyzing this bargaining environment.

[11] Note that, by design, the productivity of the proposer was at least as high as that of the responder, because the role of proposer was assigned to the participant with the higher productivity score. Consequently, we argue that, when the size of the pie is the proposer's productivity, the proposer feels a positive entitlement—the pie is high thanks to the proposer's productivity—while when the pie size is the responder's productivity the responder feels a negative entitlement—the pie is low because of the responder's productivity. In footnote 23, we comment on gender differences when entitlement effects are split into these two cases.

[12] For cases in which there was no entitlement per se (the proposer and the responder recorded the same productivity), which should be closest to the symmetric case, we did find some gender differences. However, we believe that this last effect was due to their experience in bargaining with entitlement, meaning that this behavior might be largely contaminated by their experience in negotiations in which there was entitlement.

entitlement. These amounts depended on the pie to be shared: 2.5, 5 and 7 euros, when the pies were 5, 10 and 15 euros, respectively. This information was presented to subjects during the negotiation just above the dialog box about the offer/demand.

Bargaining task: information. Everything was the same as for the *Symmetric* environment, except that only the proposer could see the actual size of the pie, while the responder only knew that it could be 5, 10 or 15 euro. This was public knowledge.

Elicitation tasks. After completing the 10 bargaining periods, subjects entered the third and last task of the experiment. We first asked the subjects explicitly: "What do you think the objective of this experiment is?." This answer was not incentivized and they were allowed to provide their answers in free format. One potential concern with the way we made subjects' genders common knowledge is that this feature of the design could yield some type of experimenter demand effect, which we address in the robustness checks (at the end of Sect. 4.2). Furthermore, in this task we elicited beliefs about self-assessed relative ability both in the real effort task and the bargaining task. As far as the real effort task is concerned, subjects were asked to reveal which quartile of the performance distribution they thought they were in and to state which gender they believed had performed better (or whether there were no gender differences). Similarly, for the bargaining task, subjects were asked to reveal which quartile of the distribution they thought they were in based on the relative surplus obtained during the 10 negotiations and to state which gender on average had obtained a greater share of the pie over the 10 periods (or whether there were no gender differences). Finally, we also elicited risk attitudes following the methodology in Eckel and Grossman (2002) and social preferences via the primary slider measure items described in Murphy et al. (2011) and implemented for z-Tree by Crosetto et al. (2012). All these measures were incentivized.[13] Table A1 in Online Appendix A shows the mean values for these control variables by gender. The main notable gender differences show up in risk preferences, where women appear to be more risk averse than men, and less confident in both their ability at the real effort task and in their bargaining ability. Figures A1 and A2 in Online Appendix A show subjects' perceptions about the gender nature of the task and bargaining by gender. Perceptions about the gender nature of the real effort task are split, with slightly more male subjects tending to believe it is a male task, while slightly more female subjects put more weight on the task being a female task. However, both male and female subjects perceive bargaining to be a male task.

[13] At the end of the experiment subjects also completed a non-incentivized questionnaire that asked for standard demographics and for the big five personality traits (Gosling et al., 2003). When we ran a principal component analysis on the self-reported answers provided by our subjects, the resulting 5 principal factors do not match the structure provided by Gosling et al. (2003). Therefore, we decided not to use personality traits measures as individual controls.

3.2 Data, hypotheses and identification strategy

3.2.1 Data

We gathered data on 2487 different negotiations from 562 different experimental subjects.[14] We focused on three important bargaining outcomes. The first outcome in a negotiation was whether the parties reached an agreement or not. The success rate measured the efficiency of bargaining: only when an agreement was reached could surplus be created. Another important outcome was earnings, measured as the share of the pie. This outcome, however, can be measured in two different ways: overall earnings, not conditional on reaching an agreement, and earnings conditional on reaching an agreement. For example, data from the field on labor markets usually involves the second one, as failed negotiations are rarely observed. However, from an efficiency point of view, the former variable is the most important, for example, when deciding whether to negotiate or in deciding on whom to delegate a negotiation. To sum up, we considered all three variables: probability of reaching an agreement, share of the pie conditional on reaching an agreement, and overall share of the pie or earnings.

3.2.2 Hypotheses

The experimental design consisted of a 2 (Male Proposer, Female Proposer)×2 (Male Responder, Female Responder)×6 (Symmetric, Empowerment, Empowerment with past feedback, Entitlement, Entitlement with past feedback, Information) factorial design. The first two factors allowed us to test for the existence of gender differences in each of the bargaining roles. Meanwhile, the third factor allowed us to check for the role of gender as an effect modifying factor between symmetric and asymmetric bargaining environments, and between environments with and without past agreements.

Given the experimental design and treatments, we started by testing two different sets of hypotheses. Firstly, we tested for the existence of gender differences in each of the six environments considered.

We hypothesized that gender differences would be non-existent in the symmetric bargaining environment, where the 50:50 norm is prevalent (Hypothesis 1). Our symmetric bargaining setting is closest to the one modeled in Ma and Manove (1993), where players do not know with certainty whether their offer will be the last one. The reason is that, if they wait for too long, they might not be able to submit the offer and get a response from the other player, while if they send their offer too early,

[14] We actually collected data on 2810 different negotiations, but 323 are from the *Entitlement* and *Entitlement with past agreements* treatments from pairings in which no entitlement was implemented and, therefore, we dropped these observations from our data set.

the opponent might send a counteroffer so that their offer is not the last one. In this framework, the expected division of the pie is unique and close to an even split.[15]

We hypothesized that asymmetric bargaining environments without past agreements may yield gender differences in all three environments: empowerment, entitlement, and information (Hypotheses 2). Note that, by making the proposer the stronger bargaining party, asymmetries break the 50:50 sharing norm but in a way that an alternative clear sharing norm is absent. This lack of clear sharing rule also allows for enough ambiguity and wiggle room for the bargaining parties to show their bargaining abilities. In particular, for the empowerment setting we decided not to provide the exact value of the outside option so as not to make that amount too salient.[16] In the entitlement setting, although it was clear the proposer was entitled to a higher share of the pie, because their productivity was higher, it was not clear how much their share of the pie should be, because the pie was determined randomly by the productivity of only one bargaining party. Finally, in the informational asymmetry, bargaining parties might expect the stronger party to try to take advantage of the informational asymmetry.

In asymmetric bargaining environments with past agreements, we intended to maintain the strength of the proposer by breaking with the 50:50 split, but in a way that a new sharing rule arises by providing bargaining parties with past agreements. Under this scenario, we hypothesized that gender differences should be less pronounced and somewhere half way between the symmetric environment and the empowerment and entitlement situations without past agreements (Hypotheses 3).

Secondly, given the fact that we also varied bargaining environments to change the existing ambiguity regarding the available sharing norms, we tested the null hypothesis of whether gender is an effect modifying factor when changing from a symmetric to an asymmetric bargaining environment (Hypothesis 4), and when changing from a bargaining situation without past agreements to a setting with information on past agreements (Hypothesis 5). With this in mind, we compared each of the asymmetric treatments with the *Symmetric* one, and asymmetric environments without past agreements with those with past agreements.

3.2.3 Identification strategy

In order to test hypotheses 1–3 (whether gender differences exist in different bargaining environments), we started with the following specification:

[15] In Ma and Manove (1993), the authors characterize a symmetric Markov-perfect equilibrium, unique at almost all nodes, in which players adopt strategic delay early in the game, make and reject offers later on, and reach agreements late in the game. In equilibrium, players miss the deadline with positive probability.

[16] In addition, we decided to ensure that the outside option would be at least 50% of the pie in order to properly implement a bargaining asymmetry through the introduction of an outside option. Note that in this case, the Nash bargaining solution (Nash Jr, 1950) and the *deal-me-out* solution (Binmore et al., 1989) return the same and, more importantly, agree on the effect of the outside option. By contrast, if the outside option was lower than 50%, these two solution concepts disagreed on whether the existence of an outside option had any effect.

$$Y_{ij} = \alpha + \beta_1 MaleProp_i + \beta_2 MaleResp_j + \gamma X_{ij} + \epsilon_{ij} \tag{1}$$

where $MaleProp_i$ ($MaleResp_j$) takes a value of 1 if the *Proposer i* (*Responder j*) is a man and 0 for a woman. To control for the characteristics in which the bargaining between *Proposer i* and *Responder j* took place, the term X_{ij} includes session, period, and pie fixed effects. Specification (1) enables us to test whether gender differences in bargaining can be detected, i.e., whether men and women in the role of *Proposer/Responder* obtain different outcomes from bargaining or whether bargaining with men is different from bargaining with women. In this specification, our coefficients of interest are β_1 and β_2.[17] The estimation results for these tests are shown in Tables 3, 4 and 5.

To test hypothesis 4 (whether introducing asymmetries are gender effect modifying factors), we compared gender differences in each asymmetric environment (without including the treatments with past information) with the symmetric environment by running the following regression:

$$\begin{aligned} Y_{ij} = \alpha &+ \beta_1 MaleProp_i + \beta_2 MaleResp_j + \beta_3 Asym_{ij} + \beta_4 Asym_{ij} * MaleProp_i \\ &+ \beta_5 Asym_{ij} * MaleResp_j + \gamma X_{ij} + \theta Asym_{ij} * X_{ij} + \epsilon_{ij} \end{aligned} \tag{2}$$

while to test hypothesis 5 (whether providing past information is a gender effect modifying factor), we compared gender differences in each asymmetric environment with the ones under the provision of past information:

$$\begin{aligned} Y_{ij} =&\alpha + \beta_1 MaleProp_i + \beta_2 MaleResp_j + \beta_3 PastAgree_{ij} + \beta_4 PastAgree_{ij} * MaleProp_i \\ &+ \beta_5 PastAgree_{ij} * MaleResp_j + \gamma X_{ij} + \theta PastAgree_{ij} * X_{ij} + \epsilon_{ij} \end{aligned} \tag{3}$$

where, as before, $MaleProp_i$ ($MaleResp_j$) takes a value of 1 if the *Proposer i* (*Responder j*) is a man and 0 for a woman and X_{ij} incorporates session, period, and pie fixed effects into the analysis to control for the environment in which the bargaining took place. In regression 2, the omitted treatment is the symmetric one, while Asym$_{ij}$ takes the value of 1 if the pair *ij* bargained in an asymmetric environment without past agreements. In regression 3, the omitted treatment is the asymmetric one without the past agreements, while PastAgree$_{ij}$ takes the value of 1 if the pair *ij* bargained in an environment in which past agreements were provided. In regressions 2 and 3, the coefficients of interest are β_4 and β_5 whose sign and significance show whether the introduction of asymmetries/past agreements indeed modify gender differences with respect to the symmetric/without past agreement environments. The estimation results for these tests are shown in Tables 6 and 7.

Given our interest in understanding gender differences in the three main interdependent outcomes of probability of reaching an agreement, earnings conditional on agreement and overall earnings, the estimation is carried out by using Cragg's

[17] Additionally, for each analysis we also tested for the existence of any gender interaction effect. We did not find strong enough evidence in favor of this hypothesis and, thus, results on gender interaction effects are not shown.

two-part model (Cragg, 1971).[18] Cragg's two-part model relies on the existence of a process that determines whether the outcome is positive or zero –i.e., whether the negotiation ended in agreement or not–, and on a different process that determines the participants' share of the pie conditional on reaching an agreement.

An interesting feature of two-part models is the interpretation of their coefficients, which provides us with a more comprehensive understanding of the impact of gender on bargaining. Firstly, the model allows us to compute the unconditional semi-elasticity ($S_z(y)$), i.e., the percentage by which the *overall* earnings, taking into account failed negotiations, of men differ from that of women in role $z \in$ Proposer, Responder. Secondly, it allow us to decompose this overall average effect of gender into two different components. On the one hand, we can isolate how the gender of the subject playing in role z impacts the probability of reaching an agreement ($S_z(P = 1)$). On the other hand, it allows us to examine how the gender of the person in role z impacts the earnings of the proposer and the responder, conditional on reaching an agreement ($S_z(y|y > 0)$).[19] See Table A3 in Online Appendix A.

This decomposition is crucial to understand gender differences in bargaining as having a significant result on overall earnings could be due to different facts: because there are gender differences on the probability of reaching a deal or because there are gender differences in earnings conditional to close a deal. Even more importantly, the absence of gender differences in overall earnings does not necessarily imply the absence of gender differences. It could be the case that one gender in a particular role makes agreement less likely but captures more of the pie when an agreement is reached (for example, because this gender negotiates more aggressively). Since these two gender differences may have opposite effects on overall earnings, it is possible to find a non significant coefficient for the overall earnings despite the existence of gender differences in both parts of the process.

A few final observations are necessary to clarify the subsequent analyses.

First, each of the regression returns 5 closely linked outcomes: probability of reaching a deal, proposer's earnings when a deal is reached, responder's earnings when a deal is reached, proposer's overall earnings and responder's overall earnings. When talking about earnings, it is necessary to differentiate between proposers' and responders' ones. For conditional earnings, this is the case because as we work with the semi-elasticities, the percentage change that the gender of bargainer in role z causes on proposer's and responder's earnings will be different when the average earnings of proposers and responders differ (which, as will be shown below,

[18] More specifically, analyses were performed using the model described in equations (7) and (9) in Cragg (1971). Using the lognormal model described by equations (7) and (11) in Cragg (1971) yielded similar results. Results were also robust with respect to using the type II Tobit model proposed by Heckman (1976, 1979). Although it has the advantage of allowing for dependency between the first and the second parts, the latter model has the important disadvantage that it requires an exclusion restriction (i.e., an instrument) to properly identify the model. The more standard Tobit model does not suit this setting because, by construction, the independent variables have the same effects in both parts of the model which, as will become clear later, is not desirable in our setting. Notice, however, that the Tobit model is nested within the Cragg's model.

[19] In particular, by working with the semi-elasticities, and similar to McDonald and Moffitt (1980), we can decompose the overall effect such that $S_z(y) = S_z(P = 1) + S_z(y|y > 0)$.

is the case in all asymmetric treatments). Notice however that, despite the magnitude being different, it will always be, by construction, of opposite sign for proposers and responders, as the situation represents a zero-sum game. That is, if one party gets more then the other must get less. For overall earnings, we need to differentiate between proposers' and responders' earnings due to the fact that, overall, the game is not a zero-sum game since if the bargaining fail, both parties get zero. This means that for overall earnings, observing a gender in a role affecting the proposers' overall earnings in certain direction does not imply that it affects the responders' ones in the opposite direction.

Second, in all specifications for bargaining outcomes, we use a two-way clustering at the subject level, that is, at the proposer and responder level simultaneously (Cameron et al., 2011; Thompson, 2011; Gu and Yoo, 2019), such that the number of clusters is the same as the number of different subjects playing the role of proposer and responder respectively.[20]

Finally, given the large number of hypothesis to be tested in the experiment, we also present the *p*-values corrected through the Romano-Wolf multiple hypothesis correction (Romano and Wolf, 2005a, b, 2016).[21] Consequently, in Sect. 4, we will only consider that a gender coefficient is significant after taking into account this correction which involves, broadly, those coefficients that are significant at least at the 5% significance level without the correction.

3.3 Assessing the experimental design

We started checking for the suitability and validity of our experimental design to test for gender differences in bargaining settings that differ in their ambiguity with regard to the existing sharing norms.

We first assessed whether the pairing protocol generated a balanced gender pairing distribution. While the pairings between subjects were done randomly, the role assigned to each party was not. Specifically, although not publicly revealed to subjects, within each pairing the party with the higher score in the real effort task was the one that was assigned the role of proposer (see footnote 9). However, given the gender neutrality of the real effort task, we would expect that all pairings should be evenly represented.

This is confirmed in Table 1, where it can be checked that, within each treatment, each different pairing accounts for close to 25%, the figure expected under full randomization. It can also be checked in Table 1 that within each treatment close to 50% of the pairings have a male proposer and 50% a male responder. In order words, men and women had ex-ante equal probabilities of being assigned the strong and weak

[20] Note that, for outcome variables, we have two non-nested clusters: proposers' and responders' clusters. The two-way clustering proposed in Cameron et al. (2011) allows us to account for the dependency of observations across both clusters by adding up the variance when clustering at the first cluster and when clustering at the second cluster and subtracting from this the variance when clustering at the intersection of both clusters.

[21] In particular, we have made used of the command `rwolf2` implemented by Clarke (2021) when obtaining the adjusted *p*-values. Notice that this approach is very demanding given that in each table we correct for twenty-four simultaneous hypotheses, which reduces considerably the power of the analysis.

Table 1 Distribution of gender pairings across and within each treatment

	Symmetric	Empowerment		Entitlement		Information
		No past agree.	Past Agree.	No past agree.	Past Agree.	
FF	24.50%	24.25%	25.17%	22.52%	26.02%	25.85%
MF	24.25%	25.50%	30.17%	20.61%	26.75%	28.05%
FM	26.75%	26.00%	22.83%	30.92%	21.45%	22.68%
MM	24.50%	24.25%	21.83%	25.95%	25.78%	23.41%
Male proposer	48.75%	49.75%	52.00%	46.56%	52.53%	51.46%
Male responder	51.25%	50.25%	44.67%	56.87%	47.23%	46.10%
Observations	400	400	600	262	415	410

bargaining roles. This allowed us to test for the existence of gender differences and gender interaction effects in bargaining.

Next, we checked whether the *Empowerment*, *Entitlement*, and *Information* treatments generated the ambiguity we aimed for, and whether the provision of past agreements in *Empowerment* and *Entitlement* reduced the ambiguity when compared to the sessions without past agreements. The distributions of responder's share of the pie when an agreement was reached across the four different bargaining environments, showed the clearest evidence for this (Fig. A3 in Online Appendix A). Firstly, while there was a clear prevalence of the 50:50 sharing rule in the *Symmetric* setting, followed in 69.1% of the successful negotiations, no such rule existed in the asymmetric ones. Secondly, in the absence of a clear sharing rule, the responder's pie shares showed much more variation in all three asymmetric bargaining settings. In a similar way, when comparing *Empowerment* and *Entitlement* with and without past agreements, we can see a reduction of the dispersion in the former environments, although this reduction is milder than the differences between the symmetric and the asymmetric bargaining environments.

To test these impressions more formally, we used two measures of dispersion to measure the existing ambiguity: the absolute value of the difference between each responder's share and the mean value of the responder's share (adjusted by treatment and pie) and the absolute value of the difference between each responder's share and the modal value of the responder's share (adjusted by treatment and pie). Table A1 in Online Appendix A shows the mean values of these two ambiguity measures by treatment and by pie. The ordering is clear. The symmetric bargaining environment shows the lowest ambiguity values, while the empowerment and informational asymmetric bargaining environments show the highest, followed by the entitlement. The provision of past agreements shows intermediate ambiguity values, higher than the symmetric but lower than those without the provision of past agreements.

Table 2 shows the average treatment effect on the ambiguity in a regression analysis. As intended by the design, the results in columns (1) and (2) of Table 2 show that all three asymmetric bargaining environments increased significantly the ambiguity in implemented sharing rules, such that the dispersion is increased. The magnitude of the increase in *Entitlement* seems to be more moderate, but as can be seen at the bottom of

Table 2 Average treatment effect on ambiguity

	Symmetric vs. Asymmetric		Emp. without vs. with past agreements		Ent. without vs. with past agreements	
	Dist_Mean	Dist_Mode	Dist_Mean	Dist_Mode	Dist_Mean	Dist_Mode
	(1)	(2)	(3)	(4)	(5)	(6)
Empowerment	0.0476***	0.0635***				
	(0.0148)	(0.0194				
Entitlement	0.0292	0.0452**				
	(0.0192)	(0.0213)				
Information	0.0403**	0.0582***				
	(0.0188)	(0.0222)				
Past agreements			– 0.0366***	– 0.0541***	– 0.0117	– 0.0287
			(0.0106)	(0.0154)	(0.0168)	(0.0204)
Observations	1263	1263	844	844	577	577
R-squared	0.088	0.112	0.114	0.118	0.051	0.056
H_0: Emp=Ent	0.2159	0.3773				
H_0: Emp=Inf	0.5955	0.8052				
H_0: Ent=Inf	0.5469	0.5720				

OLS for the mean effect of each treatment on *Dist_Mean* and *Dist_Mode* for successful agreements. *Dist_Mean* is the absolute difference between the responder's share and the mean value of the responder's share by treatment and pie. *Dist_Mode* is the absolute difference between the responder's share and the mode of the responder's share by treatment and pie. The omitted environment in columns (1) and (2) is *Symmetric*, while in columns (3) to (6) is the bargaining environment without past agreements. All regressions control for *Pie Size*, *Period* and *Session* fixed effects. Standard errors are clustered at the subject level using two-way clustering. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

the table, we cannot reject the fact that the magnitudes on the increased dispersion are comparable across the three asymmetric environments. In addition, columns (3) and (4) show that the provision of the modal amount agreed in *Empowerment* indeed reduces ambiguity when compared to *Empowerment* without past agreements. Finally, in a similar vein, columns (5) and (6) show that this reduction also occurred for *Entitlement*, although again this effect seems to be more moderate.

In summary, asymmetric bargaining environments led to more variation on splits of the pie, as intended. This allowed us to test whether gender is an effect modifying factor when moving from a symmetric to an asymmetric bargaining environment. In addition, the provision of past agreements in *Empowerment* and *Entitlement* did bring a reduction in ambiguity. Therefore, we would expect gender differences to be most important in the *Empowerment*, *Entitlement* and *Information* bargaining settings (without past agreements) compared to the symmetric environments. In addition, we would also expect that the provision of past agreements would decrease gender differences in the case of *Empowerment* and *Entitlement*. The treatment effects on ambiguity are strongest in *Empowerment* and most moderate in *Entitlement*, which will also be important when commenting on the results on gender differences.

4 Results

4.1 Estimation results 1: gender differences in bargaining

When do men and women obtain different results in alternating-offer bargaining? When does it matter whether one bargains *with* women or men? To find out, we started by testing for gender differences in bargaining environments that differ in ambiguity with respect to sharing norms.

Table 3 shows the aggregate results in the symmetric environment (panel A), in asymmetric environments (panel B), and in asymmetric environments with past agreements (panel C). The first column shows of how gender in each role affects the probability of reaching a deal, an outcome that is jointly determined by both proposers and responders, and therefore common to both. It represents the cooperative part of the bargaining and, thus, measures the effect of gender on the efficiency of the bargaining process. Columns (2) and (4) show the results, for the proposer and the responder respectively, for the share of the pie conditional on reaching an agreement. Finally, columns (3) and (5) show the results for overall earnings (taking failed negotiations into account) for the proposer and the responder respectively.

Notice that the coefficients in column 3 (column 5) are, by construction, the sum of the coefficients from columns 1 and 2 (columns 1 and 4) for the proposer (responder), i.e., $S_z(y) = S_z(P = 1) + S_z(y|y > 0)$. Also take into account that, as semi-elasticities are reported, coefficients should be interpreted as the percentage change in the bargaining outcome variable when there is a male proposer/responder compared to a female proposer/responder. Finally, note that two different channels can be found through which gender can affect earnings (columns 2–5). The first is *direct effects*, namely the impact of the gender of the bargainer on their own outcomes (i.e. the effect of being a male bargainer vs. a female bargainer). This is the case of the coefficients β_1*: Male Prop.* in regard to proposer's outcomes (columns 2 and 3) and of β_2*: Male Resp.* in regard to responder's outcomes (columns 4 and 5). The second channel is *indirect effects*, namely the impact of the gender of the bargainer on the other party's outcomes (i.e. the effect of bargaining *with* a male bargainer vs. a female bargainer) which corresponds to β_2*: Male Resp.* when looking at the proposer's outcomes (columns 2 and 3) and to β_1*: Male Prop.* in regard to responder's outcomes (columns 4 and 5). This distinction is important to separate the impact of gender in bargaining as it might be the case that, no evidence of gender affecting own outcomes is found but there is evidence of gender impacting other party's outcomes (or vice versa). To stress the existence of these two different channels, Table 3 shows the direct effects in bold.

Table 3 shows important patterns regarding the sign of the female coefficients. Firstly, men are found to be less likely to reach an agreement. Note that in every environment and bargaining role, the male coefficient is systematically negative for the probability of reaching an agreement (column 1), impairing both men's overall earnings and the overall earning of those bargaining *with* men. This negative effect is particularly strong when men are in the weaker bargaining position in asymmetric environments (β_2 in panel B). Secondly, the male proposers' coefficient in column (2) and male responders' coefficient in column (4) are always

Table 3 Gender differences: aggregate results

	$S_z(P = 1)$	Proposer's outcomes		Responder's outcomes	
		$S_z(y\|y > 0)$	$S_z(y)$	$S_z(y\|y > 0)$	$S_z(y)$
	(1)	(2)	(3)	(4)	(5)
Panel A: Symmetric					
β_1: Male Prop	– 0.0349	**0.00837**	**– 0.0265**	– 0.00832	– 0.0432
	(0.0436)	**(0.0287)**	**(0.0517)**	(0.0286)	(0.0527)
	[0.866]	**[0.966]**	**[0.933]**	[0.966]	[0.858]
β_2: Male Resp	– 0.0466	– 0.0439	– 0.0905	**0.0437**	**– 0.00290**
	(0.0470)	(0.0329)	(0.0633)	**(0.0323)**	**(0.0504)**
	[0.736]	[0.492]	[0.403]	**[0.492]**	**[0.966]**
Observations	400	400	400	400	400
Clusters	80	80	80	80	80
Panel B: Asymmetric					
β_1: Male Prop	– 0.0339	**0.0601**^{**,a}	**0.0262**	– 0.112^{**,a}	– 0.146^{***,a}
	(0.0268)	**(0.0235)**	**(0.0338)**	(0.0446)	(0.0544)
	[0.544]	**[0.008]**	**[0.881]**	[0.008]	[0.005]
β_2: Male Resp	– 0.0991^{***,a}	– 0.0405^{**,b}	– 0.140^{***,a}	**0.0755**^{**,b}	**– 0.0236**
	(0.0284)	(0.0180)	(0.0360)	**(0.0335)**	**(0.0402)**
	[0.001]	[0.024]	[0.001]	**[0.024]**	**[0.993]**
Observations	1072	1072	1072	1072	1072
# Clusters	242	242	242	242	242
Panel C: Asymmetric + Past agreements					
β_1: Male Prop	– 0.0124	**0.0214**	**0.00898**	– 0.0494	– 0.0617
	(0.0300)	**(0.0166)**	**(0.0317)**	(0.0386)	(0.0527)
	[0.937]	**[0.529]**	**[0.966]**	[0.529]	[0.612]
β_2: Male Resp	– 0.0522*	– 0.0420^{***,a}	– 0.0943^{**,a}	**0.0963**^{***,a}	**0.0441**
	(0.0317)	(0.0129)	(0.0371)	**(0.0300)**	**(0.0381)**
	[0.232]	[0.001]	[0.009]	**[0.001]**	**[0.625]**
Observations	1015	1015	1015	1015	1015
# Clusters	240	240	240	240	240

Cragg's hurdle model for the pie share captured by the proposer and by the responder. Semi-elasticities are reported. $S_z(P = 1)$ in column (1) refers to the effect of gender in the probability of reaching a deal. $S_z(y|y > 0)$ in columns (2) and (4) refers, for the proposer and the responder outcomes respectively, to the effect of gender in the share captured conditional on reaching a deal. $S_z(y)$ in columns (3) and (5) refers, for the proposer and the responder outcomes respectively, to the effect of gender in the overall share captured (including fail negotiations) such that $S_z(y) = S_z(P = 1) + S_z(y|y > 0)$. All regressions control for each bargaining environment, *Pie Size*, *Period*, and *Session* fixed effects. All fixed effects are interacted with each bargaining environment. Direct effects displayed in bold. Standard errors are clustered at subject level using two-way clustering. Romano-Wolf multiple hypothesis corrected p-values in brackets with 1000 bootstrap replication. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ for standard p-values. $^{a}p < 0.01$, $^{b}p < 0.05$, $^{c}p < 0.1$ for Romano-Wolf multiple hypothesis corrected p-values

positive, showing that men obtain a higher share of the pie conditional on reaching an agreement. In short, men are less likely to close a deal but aresssable to secure higher shares when reaching an agreement, thus direct effects of gender on overall earnings (β_1 in column 3 and β_2 in column 5) are not statistically different from zero. However, the indirect effect on overall earnings (β_2 in column 3 and β_3 in column 5) is always negative, suggesting that men do not earn more overall, but bargaining *with* men is worse than bargaining *with* women.

Turning the focus to *when* these gender differences are significant, Table 3 also shows evidence of the hypothesized effect of ambiguity on gender differences (hypotheses 1–3). In the symmetric environment (Panel A), which is the benchmark, and consistently with Hypothesis 1, we find no evidence for gender differences.

In line with Hypothesis 2, in asymmetric environments (Panel B), we find ample evidence of gender differences. Most importantly, having a male responder lowers the probability of reaching an agreement by 10%. A look at the proposer's side shows that when men close a deal, they get 6% more than women. This translates into an indirect effect such that when responders negotiate with men they obtain 11.2% less than when they deal with women. Since the effect of a male proposer on the probability of closing a deal is negative (although not significant), the direct effect of male proposers on overall earnings is attenuated while the indirect effect is enhanced. Thus, in an asymmetric environment and in terms of overall earnings, male proposers do not obtain significantly higher profits, but negotiating with a male proposer reduces total earnings by 14.6%. We find a similar effect when we turn our attention to responders. Male responders show a direct effect of 7.5% and an indirect effect of −4% but, as they reduce the probability of reaching an agreement by 10%, the direct effect on overall earnings is not significant, while the indirect effect is negative and significant with a size of −14%. In sum, men's negotiating strategies in asymmetric environments do not favor them on average, because although they obtain more when they are successful, they are less likely to reach deals. However, men's more aggressive bargaining behavior results, on average, in their bargaining counterpart obtaining lower overall profits.

Finally, panel C provides evidence in favor of Hypothesis 3. It can be seen that when information about past agreements is made available to bargainers in order to reduce ambiguity, having a male responder does not significantly lower the probability of reaching an agreement. Moreover, all the effects on the proposer's side (direct and indirect) vanish, but on the responder's side male responders are still observed to obtain more when reaching a deal and, although their lower probability of reaching an agreement is no longer significant, this translates into lower overall earnings for proposers when bargaining with male responders.

In Table A4 in Online Appendix A, we show results complementary to those in Tables 3 but with added individual level controls (self-confidence, risk attitudes and social preferences).[22] As one would expect, the main results hold, but they become weaker in terms of both the size and significance of gender coefficients, particularly

[22] In this analysis including individual level controls we have also used alternative specification for risk preferences, with dummies instead of a continuous variable. The results remain the same.

in panel B, as mediating factors such as confidence and risk aversion are now controlled for. In other words, controlling for individual characteristics, where women are found to be more risk averse and less confident in their ability to perform the task and to bargain, as shown in Table 9 in Appendix A, attenuates the estimated gender differences in bargaining. Thus, as shown in the context of gender differences in competition (Niederle et al., 2011; Gillen et al., 2019; Van Veldhuizen, 2022), gender differences in bargaining seem to be primarily driven by other well known gender differences other than pure ability to perform in negotiations, such as risk tolerance and self-confidence. Comparing the results for Tables 3 and A4 in Online Appendix A, it is further shown that these dimensions matter more when ambiguity is greater, as it is precisely in this situation where results change more when controlling for individual level traits.

To sum up, when we find gender differences, they go in the stereotypically expected direction: men prove to be tougher bargainers, bringing a higher likelihood of failure, but they obtain better deals than women when an agreement is reached. The higher likelihood of negotiation failure is especially strong in highly asymmetric bargaining environments, making it undesirable to bargain *with* men. In addition, the ability to secure a higher share of the pie when an agreement is reached is canceled out in most cases by the always (significant or not) higher probability of failure, meaning that men and women end up with similar overall earnings. When the three types of bargaining environments (symmetric, the three asymmetric environments and the asymmetric environment with past agreements) are compared, we find evidence in the hypothesized direction: gender differences are strongest in asymmetric environments (hypothesis 2), followed by asymmetric environments with past agreements (hypothesis 3), and finally symmetric environments, where ambiguity is lowest (hypothesis 1).

4.2 Estimation results 2: disaggregated results

In the aggregate analysis, in Table 3, the estimated figures reflect an average gender difference across all three different asymmetric bargaining environments (in panel B) and in the two different bargaining environments with past agreements (in panel C). We now turn to the treatment-by-treatment analysis, ending with a test of whether gender is an effect modifying factor when ambiguity manipulations are compared across different environments.

Table 4 shows the gender coefficients in the probability of reaching an agreement (Panel A) and in proposer's and responder's earnings conditional on reaching an agreement (Panel B1 and B2, respectively) for each of the 6 bargaining environments in each column.[23]

[23] We also broke down negotiations with entitlement further into those with positive entitlement (the pie to be shared is that of the proposer), negative entitlement (the pie to be shared is that of the responder) and cases in which there was no entitlement (the proposer and the responder had the same productivity). The only major difference between environments with positive and negative entitlements was that, in the latter, male responders were less likely to reach an agreement.

The first point to be considered is the probability of reaching an agreement (Panel A of Table 4). The first noticeable fact is that, except for the treatment of empowerment with information on past agreements (column 5), having a male bargainer always decreases the probability of reaching a deal. However, this effect is only significant in the treatments of empowerment and information (columns 2 and 4), i.e. when ambiguity is strongest. In particular, having a male responder facing an empowered party decreases the probability of reaching a deal by 12.5%. When the proposer holds more information, the drop is 10.3% with respect to female responders. Thus, having male bargainers hurts the overall efficiency of the bargaining process, especially when men occupy weak positions in highly asymmetric bargaining environments and there are no salient sharing norms.

What happens when bargainers reach a deal? Panel B1 of Table 4 shows that, despite its coefficient being always positive, male proposers are only able to capture significantly more than female proposers –8.2 % more– when they can exploit an advantage in environments with empowerment without past agreements (column 2). This direct effect implies that, as shown in Panel B2, responders bargaining *with* empowered male proposers get 24.5% less than when they bargain with female proposers. However, when the asymmetry in empowerment is maintained but information about the modal agreement is provided, the effects of gender (both direct and indirect) decrease to the point where they are no longer significant (column 5).

Regarding the effects of gender in the role of responder (β_2's), we find that male responders obtain a 13% greater share than their female counterparts when they bargain with a more informed proposer, as shown in Panel B2, column (4). This translates into an indirect effect in Panel B1 by which proposers obtain 8.6% less when bargaining with a male responder. Under a situation with entitlement and information about past agreements (column 6), male responders were found to be able to secure 6.8% more of the pie (Panel B2), which means that proposers obtain 5.9% less pie when bargaining with men (Panel B2).[24]

Overall, Table 4 shows that when men bargain the probability of reaching a deal is decreased, especially when they are placed in the weak position at the bargaining table and when there are no clear sharing norms helping to reduce ambiguity but, at the same time, they are able to secure a greater share of the pie when they reach a deal. The next question that arises naturally is whether this strategy pays off.

This question is examined in Table 5, which shows the effects of gender on the overall earnings of the proposer (Panel A) and of the responder (Panel B). Remember that coefficients in Table 5 can be computed as just the sum of the coefficients in Panel A and Panel B from Table 4. Table 5 shows that although both direct effects (β_1 in Panel A and β_2 in Panel B) are typically positive, they are moderate in size and not significant regardless of the particular environment examined. This is because although men are typically able to secure a bigger share of the pie for themselves than women, this effect is canceled out by the loss of efficiency derived from men lowering the probability of reaching an agreement. However, a look at the indirect effects (β_2 in Panel A and β_1 in Panel B) shows that the point estimates are

[24] Notice, however, that the point estimate is still smaller when information is provided about past deals than in the entitlement without past agreements.

Table 4 Gender differences: disaggregated results, probability of agreement and conditional earnings $S_z(P = 1)$ and $S_z(y|y > 0)$

	Symm.	Asymmetric			Asymmetric +Past agreements	
		Emp.	Ent.	Info.	Emp.	Ent.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: probability of reaching a deal ($S_z(P = 1)$)						
β_1:Male Prop	− 0.0349	− 0.0565	− 0.0385	− 0.00847	0.0117	− 0.0511
	(0.0436)	(0.0487)	(0.0463)	(0.0383)	(0.0413)	(0.0431)
	[0.926]	[0.805]	[0.926]	[0.997]	[0.997]	[0.805]
β_2:Male Resp	− 0.0466	− 0.125**,c	− 0.0531	− 0.103***,a	− 0.0740	− 0.0166
	(0.0470)	(0.0578)	(0.0537)	(0.0336)	(0.0477)	(0.0368)
	[0.885]	[0.090]	[0.885]	[0.009]	[0.471]	[0.995]
Observations	400	400	262	410	600	415
Clusters	80	80	80	82	120	120
Panel B: conditional earnings ($S_z(y\|y > 0)$)						
Panel B1: proposer						
β_1:Male Prop	0.00837	0.0818**,c	0.0712*	0.0235	0.0308	9.92e− 05
	(0.0287)	(0.0339)	(0.0406)	(0.0430)	(0.0202)	(0.0255)
	[0.997]	[0.053]	[0.299]	[0.985]	[0.500]	[0.997]
β_2:Male Resp	− 0.0439	0.00793	− 0.0606	− 0.0864***,b	− 0.0312**	− 0.0590**,b
	(0.0329)	(0.0244)	(0.0417)	(0.0307)	(0.0149)	(0.0240)
	[0.671]	[0.997]	[0.562]	[0.019]	[0.119]	[0.048]
Observations	400	400	262	410	600	415
Clusters	80	80	80	82	120	120
Panel B2: responder						
β_1:Male Prop	− 0.00832	− 0.245**,c	− 0.0880*	− 0.0354	− 0.108	− 0.000116
	(0.0286)	(0.104)	(0.0503)	(0.0642)	(0.0720)	(0.0298)
	[0.997]	[0.053]	[0.299]	[0.985]	[0.500]	[0.997]
β_2:Male Resp	0.0437	− 0.0234	0.0748	0.130***,b	0.108**	0.0689**,b
	(0.0323)	(0.0724)	(0.0508)	(0.0444)	(0.0514)	(0.0291)
	[0.671]	[0.997]	[0.562]	[0.019]	[0.119]	[0.048]
Observations	400	400	262	410	600	415
Clusters	80	80	80	82	120	120

The dependent variables are: The probability of reaching a deal ($S_z(P = 1)$, Panel A) and earnings conditional to reaching a deal ($S_z(y|y > 0)$) for the proposer (Panel B1) and the Responder (Panel B2). Semi-elasticities are reported such that $S_z(y) = S_z(P = 1) + S_z(y|y > 0)$. All regressions control for *Pie Size*, *Period*, and *Session* fixed effects. Standard errors clustered at subject level using two-way clustering in parentheses. Romano-Wolf multiple hypothesis corrected p-values in brackets with 1000 bootstrap replication. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$ for standard p-values.

[a]$p < 0.01$,[b]$p < 0.05$, [c]$p < 0.1$ for Romano-Wolf multiple hypothesis corrected p-values

Table 5 Gender differences: disaggregated results, overall earnings $S_z(y)$

	Symm.	Asymmetric			Asymmetric +Past agreements	
		Emp.	Ent.	Info.	Emp.	Ent.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: proposer's overall earnings						
β_1:Male Prop	– 0.0265	0.0253	0.0327	0.0150	0.0425	– 0.0510
	(0.0517)	(0.0534)	(0.0604)	(0.0593)	(0.0419)	(0.0464)
	[0.998]	[0.998]	[0.998]	[0.998]	[0.945]	[0.928]
β_2:Male Resp	– 0.0905	– 0.117*	– 0.114	– 0.189***,^{a}	– 0.105*	– 0.0755*
	(0.0633)	(0.0676)	(0.0700)	(0.0473)	(0.0550)	(0.0439)
	[0.643]	[0.363]	[0.479]	[0.002]	[0.254]	[0.370]
Observations	400	400	262	410	600	415
Clusters	80	80	80	82	120	120
Panel B: responder's overall earnings						
β_1:Male Prop	– 0.0432	– 0.301**,^{b}	– 0.126*	– 0.0439	– 0.0961	– 0.0513
	(0.0527)	(0.123)	(0.0698)	(0.0730)	(0.0904)	(0.0562)
	[0.983]	[0.049]	[0.304]	[0.998]	[0.937]	[0.970]
β_2:Male Resp	– 0.00290	– 0.149*	0.0218	0.0267	0.0342	0.0523
	(0.0504)	(0.0818)	(0.0716)	(0.0521)	(0.0556)	(0.0468)
	[0.998]	[0.304]	[0.998]	[0.998]	[0.998]	[0.925]
Observations	400	400	262	410	600	415
Clusters	80	80	80	82	120	120

The dependent variable is overall earnings, which is the sum of the probability of reaching an agreement and earnings conditional on reaching an agreement. Cragg's truncated hurdle model for the overall earnings captured by the proposer (Panel A) and by the responder (Panel B) in each bargaining environment. Semi-elasticities are reported. All regressions control for *Pie Size*, *Period*, and *Session* fixed effects. Standard errors clustered at subject level using two-way clustering in parentheses. Romano-Wolf multiple hypothesis corrected p-values in brackets with 1000 bootstrap replication. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$ for standard p-values. $^a p < 0.01$,$^b p < 0.05$,$^c p < 0.1$ for Romano-Wolf multiple hypothesis corrected p-values

substantially larger than those for the direct effects and are always negative in sign, implying that bargaining *with* men is detrimental for both proposers and responders. When considering significant results, we only find that male empowered proposers decrease responders' earnings by 30.1% (β_1 in column 2 of Panel B) and that male responders decrease proposer's overall earnings by 18.9% when there are asymmetries of information (β_2 in column 3 of Panel A).

When are gender differences strongest? Empowerment and informational asymmetries without past agreements prove to be bargaining environments where gender differences are bigger. These are precisely the environments in which ambiguity is strongest.

Three robustness checks are worth noting.

First, as with the aggregate results, when individual background characteristics such as self-confidence and risk attitudes are controlled for, all gender differences become weaker in terms of both size and significance, as risk aversion and confidence are mediating factors (see Tables A5 and A6 in Online Appendix A).

Second, we find that 7.65% of participants mention gender as an objective to be studied by the experiment. We acknowledge that this percentage may represent a lower bound, as there could be subjects who thought the experiment was about gender but did not say so. Despite this limitation, we test how robust the main findings are to potential experimenter demand effects by excluding from the sample those negotiations in which participant mentioned gender as an aim of the experiment (see Tables A8 and A7 in Online Appendix A). The main results hold, although again the size and significance fall and some of the results become insignificant. However, note that the number of observations also decreases significantly. Therefore, it is not clear whether the weaker gender differences are due to experimenter demand effects or to the lower number of observations. We acknowledge that this robustness test provides an imperfect control for potential experimenter demand effects.

Third, we find an important deadline effect. As shown by Fig. A4 in the Online Appendix A, about a quarter of the successful negotiations were reached within the last 10 seconds of the 3-minute time limit. This is consistent with previous findings in bargaining, both with field data (e.g., Cramton and Tracy, 1992) and in the lab (e.g., Roth et al. 1988; Sterbenz and Phillips, 2001; Gneezy et al., 2003; Gächter and Riedl, 2005). We therefore identify and refer here to proposals (regardless of whether they are offers or demands) made within the last 10 seconds as *ultimatums*. Accordingly, Table A9 in Online Appendix A shows that in *ultimatum* agreements, the final split of the pie is affected by whether the last proposal is a demand or an offer, but this is not the case for non-*ultimatum* agreements. In light of this so called *deadline effect*, two questions arise. Firstly, given that bargaining parties might self-select into an ultimatum bargaining environment, it is advisable to test for gender differences in the likelihood of ending up in an *ultimatum* type of bargaining process. Table A10 in Online Appendix A looks at gender differences in the propensity to close a deal via an *ultimatum* in each of the six treatments. The results suggest that a bargaining pairing involving men is more likely in general to close a deal of this type. Second, we test whether results remain the same when these negotiations closed in the last 10 seconds are excluded (see Tables A11 and 12 in Online Appendix A) and find that the main findings are robust to the existence of deadline effects.[25]

Next, we examine whether gender differences change when asymmetries are introduced (by comparing column 1 with columns 2–4 from Tables 4 and 5) and when the asymmetries are maintained but the bargaining wiggle room is reduced

[25] Two alternatives to fixed time limit are random stopping time (e.g., Dittrich et al., 2014) and shrinking pie in real time (e.g., in Embrey et al., 2014). As men and women are known to differ in their risk preferences, these alternative methodologies might yield major gender differences in bargaining due to their different risk aversion levels, so we decided to stick to the fixed time limit.

by providing past information (by comparing columns 2 and 3 with 5 and 6). These tests are presented in Table 6—for the probability of reaching a deal and conditional earnings—and Table 7—for overall earnings. In both tables, columns 1–3 test for the effect of introducing asymmetries, and columns 4 and 5 for the effects of reducing wiggle room in asymmetric environments. No statistically significant results are found, but there are some noteworthy patterns in line with our initial hypotheses. Consistent with Hypothesis 4, columns 1–3 in Table 6 show that introducing asymmetries in a high ambiguity context makes the presence of men decrease the likelihood of reaching a deal but increases the share of the pie that they get when a deal is closed, both as proposers and responders, and reduces what bargainers get when bargaining with men. Hence, compared to symmetric environments, introducing asymmetries with no clear sharing norm tends to make men more aggressive in their bargaining strategy, which pays off when they close a deal. However, these are only suggestive findings as they are not statistically significant.

But, is this effect due to the asymmetry itself or to the fact that asymmetries also blur the existing sharing norm? Columns 4 and 5 of Tables 6 and 7 answer this question by comparing the coefficients of the empowerment and entitlement treatments with those when past information about the modal deal reached in previous experimental sessions is provided. In virtually all cases the comparison shows that the coefficients from columns 1 and 4 and those from columns 2 and 5 have contrary signs. Despite there being no significant interactions, this observation suggests that providing a potential sharing norm undoes the gender differences found in the asymmetric treatment and reinforces the idea that it is not so much the existence of an asymmetry that results in gender differences but rather the lack of a clear idea about what deal is acceptable. This is in line with Hypothesis 5. However, again, these are only suggestive findings as they are not statistically significant.

We perform a final test to shed some additional light on how ambiguity affects the existence and size of gender differences. As shown by Table 2, each of the six treatments generates a bargaining environment with different degrees of ambiguity. Therefore, we can linearize the relationship by using the degree of ambiguity in each of the treatments as our variable of interest and interact it with the *Male Prop* and *Male Resp* dummies. The idea is to test whether increasing ambiguity (abstracting from the treatment itself), makes gender differences greater. The results in Table 8 show that accounting only for the degree of ambiguity generated by each treatment returns results that are consistent with those described above. The results show a significant coefficient for gender as an effect modifying factor if we use the distance to the mean as a measure of ambiguity, but a non-significant effect modifying factor (although in the right direction) if we use the distance to the mode as a measure of ambiguity. This high consistency in the direction of the effect with the hypotheses put forward in the paper further suggests a link between the degree of ambiguity in a bargaining environment and the likelihood of finding gender differences.

Table 6 Gender as modifying factor: probability of agreement and conditional earnings $S_z(P = 1)$ and $S_z(y|y > 0)$

	High ambiguity Vs. Symmetric				Past info Vs. No info	
	Emp.	Ent.	Info.		Emp.	Ent.
	(1)	(2)	(3)		(4)	(5)
Panel A: Probability of reaching a deal ($S_z(P = 1)$)						
β_4: Male Prop#Asym	− 0.0192	− 0.00789	0.0257	β_4: Male Prop#Past	0.0692	− 0.00281
	(0.0656)	(0.0625)	(0.0574)		(0.0641)	(0.0618)
	[0.997]	[0.997]	[0.996]		[0.810]	[0.997]
β_5: Male Resp#Asym	− 0.0741	− 0.0123	− 0.0587	β_5: Male Resp#Past	0.0545	0.0459
	(0.0673)	(0.0703)	(0.0521)		(0.0711)	(0.0705)
	[0.810]	[0.997]	[0.810]		[0.969]	[0.982]
Observations	800	662	810		1000	677
Clusters	160	160	162		200	200
Panel B: Conditional earnings ($S_z(y|y > 0)$)						
Panel B1: Proposer						
β_4: Male Prop#Asym	0.0953**	0.0678	0.0178	β_4: Male Prop#Past	− 0.0489	− 0.0722
	(0.0456)	(0.0478)	(0.0492)		(0.0386)	(0.0467)
	[0.103]	[0.558]	[0.996]		[0.696]	[0.467]
β_5: Male Resp#Asym	0.0465	− 0.0223	− 0.0530	β_5: Male Resp#Pas	− 0.0395	0.00311
	(0.0392)	(0.0499)	(0.0424)		(0.0270)	(0.0456)
	[0.784]	[0.996]	[0.728]		[0.535]	[0.997]
Observations	800	662	810		1,000	677
Clusters	160	160	162		200	200
Panel B2: Responder						
β_4: Male Prop#Asym	− 0.177**	− 0.0735	− 0.0223	β_4: Male Prop#Past	0.150	0.0862
	(0.0839)	(0.0520)	(0.0608)		(0.125)	(0.0556)
	[0.103]	[0.558]	[0.996]		[0.696]	[0.467]
β_5: Male Resp#Asym	− 0.0800	0.0242	0.0667	β_5: Male Resp#Past	0.130	− 0.00373
	(0.0691)	(0.0542)	(0.0520)		(0.0869)	(0.0544)
	[0.784]	[0.996]	[0.728]		[0.535]	[0.997]
Observations	800	662	810		1000	677
Clusters	160	160	162		200	200

The dependent variables are: The probability of reaching a deal ($S_z(P = 1)$, Panel A) and earnings conditional to reaching a deal ($S_z(y|y > 0)$) for the proposer (Panel B1) and the Responder (Panel B2). Semi-elasticities are reported such that $S_z(y) = S_z(P = 1) + S_z(y|y > 0)$. Semi-elasticities are reported. All regressions control for *Pie Size*, *Period*, and *Session* fixed effects. All fixed effects are interacted with each bargaining environment. Standard errors clustered at subject level using two-way clustering in parentheses. Romano-Wolf multiple hypothesis corrected p-values in brackets with 1000 bootstrap replication. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$ for standard p-values. $^{a}p < 0.01$,$^{b}p < 0.05$,$^{c}p < 0.1$ for Romano-Wolf multiple hypothesis corrected p-values

Table 7 Gender as modifying factor: overall earnings $S_z(y)$

	Asymmetric VS Symmetric				Past info VS No info	
	Emp.	Ent.	Info.		Emp.	Ent.
	(1)	(2)	(3)		(4)	(5)
Panel A: Proposer's overall earnings						
β_4: Male Prop#Asym	0.0760	0.0599	0.0435	β_4: Male Prop#Past	0.0203	– 0.0750
	(0.0733)	(0.0775)	(0.0770)		(0.0668)	(0.0734)
	[0.896]	[0.982]	[0.998]		[1.000]	[0.902]
β_5: Male Resp#Asym	– 0.0276	– 0.0346	– 0.112	β_5: Male Resp#Past	0.0150	0.0490
	(0.0856)	(0.0925)	(0.0732)		(0.0833)	(0.0859)
	[1.000]	[1.000]	[0.494]		[1.000]	[0.998]
Observations	800	662	810		1000	677
Clusters	160	160	162		200	200
Panel B: Responder's overall earnings						
β_4: Male Prop#Asym	– 0.196*	– 0.0814	0.00344	β_4: Male Prop#Past	0.220	0.0834
	(0.115)	(0.0826)	(0.0821)		(0.153)	(0.0878)
	[0.329]	[0.908]	[1.000]		[0.573]	[0.926]
β_5: Male Resp#Asym	– 0.154*	0.0119	0.00803	β_5: Male Resp#Past	0.184*	0.0422
	(0.0841)	(0.0816)	(0.0660)		(0.0941)	(0.0868)
	[0.256]	[1.000]	[1.000]		[0.178]	[1.000]
Observations	800	662	810		1000	677
Clusters	160	160	162		200	200

The dependent variable is overall earnings, which is the sum of the probability of reaching an agreement and earnings conditional on reaching an agreement. Cragg's truncated hurdle model for the overall earnings captured by the proposer (Panel A) and by the responder (Panel B) in each bargaining environment. Cragg's truncated hurdle model for gender as modifying factor of asymmetry (Columns (1)–(3)) and ambiguity (Columns (4) and (5)) on the overall earnings captured by the proposer (Panel A) and by the responder (Panel B). Semi-elasticities are reported. All regressions control for *Pie Size*, *Period*, and *Session* fixed effects. All fixed effects are interacted with each bargaining environment. Standard errors clustered at subject level using two-way clustering in parentheses. Romano-Wolf multiple hypothesis corrected p-values in brackets with 1000 bootstrap replication. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$ for standard p-values. $^{a}p < 0.01$,$^{b}p < 0.05$, $^{c}p < 0.1$ for Romano-Wolf multiple hypothesis corrected p-values

5 Discussion

In this paper, we have addressed the question of when gender differences in bargaining will be most likely to manifest. We proposed an experimental design that varied bargaining environments with the goal of changing ambiguity regarding the sharing norm, to test whether increasing the existing ambiguity will make more likely to observe gender differences.

We find that gender differences are absent in symmetric settings, where a 50:50 split is the norm and hence ambiguity is the lowest, but they become significant when asymmetries between bargaining roles were introduced. Furthermore, all the detected gender differences are in the stereotypically expected direction, namely that men make reaching an agreement harder, and, if anything, they obtain a higher share of the pie when an

Table 8 Gender differences: ambiguity as independent variable

	$S_j(P = 1)$	Proposer's outcomes		Responder's outcomes	
		$S_z(y\|y > 0)$	$S_z(y)$	$S_z(y\|y > 0)$	$S_z(y)$
	(1)	(2)	(3)	(4)	(5)
Panel A: Ambiguity measured as the distance to the mean					
Male Prop	– 0.0387	– 0.0518	– 0.0905	0.0949	0.0562
	(0.0698)	(0.0414)	(0.0804)	(0.0749)	(0.103)
Male Resp	0.0805	– 0.0378	0.0427	0.0654	0.146*
	(0.0729)	(0.0347)	(0.0872)	(0.0622)	(0.0854)
Ambiguity	4.174	3.432	7.606	– 6.458	– 2.284
	(4.245)	(2.542)	(4.880)	(4.761)	(6.470)
Male Prop#Amb	0.175	**1.150***	**1.325**	– 2.122*	– 1.947
	(0.899)	**(0.594)**	**(1.066)**	(1.086)	(1.424)
Male Resp#Amb	– 1.997**	– 0.0534	– 2.051*	**0.151**	**– 1.846***
	(0.940)	(0.446)	(1.129)	**(0.804)**	**(1.088)**
Observations	2,487	2,487	2,487	2,487	2,487
# Clusters	562	562	562	562	562
Panel B: Ambiguity measured as the distance to the mode					
Male Prop	– 0.0418	0.0181	– 0.0237	– 0.0340	– 0.0758
	(0.0447)	(0.0262)	(0.0537)	(0.0471)	(0.0623)
Male Resp	– 0.00878	– 0.0338*	– 0.0426	0.0618*	0.0530
	(0.0479)	(0.0204)	(0.0569)	(0.0367)	(0.0521)
Ambiguity	3.200	3.572	6.772	– 6.659	– 3.460
	(4.112)	(2.481)	(4.731)	(4.648)	(6.305)
Male Prop#Amb	0.208	**0.223**	**0.430**	– 0.406	– 0.198
	(0.516)	**(0.327)**	**(0.648)**	(0.588)	(0.728)
Male Resp#Amb	– 0.782	– 0.0948	– 0.876	**0.179**	**– 0.603**
	(0.563)	(0.211)	(0.665)	**(0.378)**	**(0.561)**
Observations	2,487	2,487	2,487	2,487	2,487
# Clusters	562	562	562	562	562

Cragg's hurdle model for the pie share captured by the proposer and by the responder using as our treatment variable the degree of ambiguity measured as the dispersion from the mean deal (Panel A) and as the dispersion from the modal deal (Panel B). Semi-elasticities are reported. $S_z(P = 1)$ in column (1) refers to the effect of gender in the probability of reaching a deal. $S_z(y|y > 0)$ in columns (2) and (4) refers, for the proposer and the responder outcomes respectively, to the effect of gender in the share captured conditional on reaching a deal. $S_z(y)$ in columns (3) and (5) refers, for the proposer and the responder outcomes respectively, to the effect of gender in the overall share captured (including fail negotiations) such that $S_z(y) = S_z(P = 1) + S_z(y|y > 0)$. All regressions control for each bargaining environment, *Pie Size*, *Period*, and *Session* fixed effects. All fixed effects are interacted with each bargaining environment. Direct effects displayed in bold. Standard errors are clustered at subject level using two-way clustering. Direct effects displayed in bold. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

agreement is reached. Although the latter result resonates with the stereotypically expected behavior, the former result, i.e., men having a higher likelihood of failure (also found by Eckel and Grossman 2001; Sutter et al., 2009, Andersen et al., 2018), has not been stressed enough in our opinion, despite having important consequences for efficiency. This finding is also important because it shows that when looking at *overall earnings*, men and women may obtain the same results from bargaining for themselves (*direct effect*), but men tend to decrease other people's earnings by reducing the efficiency of the bargaining (*indirect effect*). Also, even in asymmetric environments, when past agreements were provided to decrease existing ambiguity, the effect of men having a lower likelihood of an agreement prove to be no longer significant. Furthermore, the detected gender differences are strongest in asymmetric environments were ambiguity is highest: empowerment and informational bargaining environments without past agreements. However, when tested if we can conclude that gender is an effect modifying factor with respect to ambiguity we only find suggestive evidence for this.

What did we learn about real-life negotiations? We argue that most, if not all, bargaining situations in economically relevant situations, such as in wage negotiations in labor markets, are not only highly asymmetric but they also have enough wiggle room, without clear guidance with respect to which split to settle on, to allow men and women to get different outcomes. The counterexample would be a firm in which salaries for each of the positions are fully transparent and there is no room for negotiation (Hospido et al., 2019; Recalde and Vesterlund, 2020; Bennedsen et al., 2022), where one would not expect bargaining to play any role in setting wages.

Future research should focus on studying *how much* wiggle room is needed to be able to detect these gender differences. Note that our study shows that ambiguity is a *necessary* condition for gender differences to flourish, although even in asymmetric environments it is possible to reduce ambiguity by providing bargaining parties with past agreements, which may serve as reference points.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10683-023-09796-9.

Acknowledgements The authors would like to thank Antonio Cabrales, Javier Gardeazabal, Pedro Rey-Biel, anonymous referees, and participants of seminars at various universities for helpful comments. Iñigo Hernandez- Arenaz acknowledges the financial support provided by Vicerrectorado de Investigación de la UPV/EHU (PIF//13/015), Departamento de Educación, Política Lingüística y Cultura del Gobierno Vasco (IT869-13), Ministerio de Ciencia, Innovación y Universidades (PID2019-108343GA-I00), and Grant PID2021-127119NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe". Nagore Iriberri acknowledges the financial support from Grant PID2019-106146GB-I00, funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe", Departamento de Educación and Política Lingüística y Cultura del Gobierno Vasco (IT1697-22), and the Norwegian Research Council (TOPPFORSK 250506). The replication material for the study is available at https://osf.io/9wu2e/ (DOI https://doi.org/10.17605/OSF.IO/9WU2E).

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

References

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference points and effort provision. *The American Economic Review, 101*(2), 470–492.

Andersen, S., Ertac, S., Gneezy, U., List, J. A., & Maximiano, S. (2018). On the cultural basis of gender differences in negotiation. *Experimental Economics, 21*(4), 757–778.

Andersen, S., Marx, J., Nielsen, K. M., & Vesterlund, L. (2021). Gender differences in negotiation: Evidence from real estate transactions. *The Economic Journal, 131*(638), 2304–2332.

Ayres, I. (1991). Fair driving: Gender and race discrimination in retail car negotiations. *Harvard Law Review*, (pp. 817–872).

Ayres, I. & Siegelman, P. (1995). Race and gender discrimination in bargaining for a new car. *The American Economic Review*, (pp. 304–321).

Azmat, G., & Petrongolo, B. (2014). Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics, 30*, 32–40.

Babcock, L., & Laschever, S. (2009). *Women don't ask: Negotiation and the gender divide*. Princeton University Press.

Bennedsen, M., Simintzi, E., Tsoutsoura, M., & Wolfenzon, D. (2022). Do firms respond to gender pay gap transparency? *The Journal of Finance, 77*(4), 2051–2091.

Biasi, B., & Sarsons, H. (2022). Flexible wages, bargaining, and the gender gap. *The Quarterly Journal of Economics, 137*(1), 215–266.

Binmore, K., Shared, A., & Sutton, J. (1989). An outside option experiment. *The Quarterly Journal of Economics, 104*(4), 753–770.

Blau, F. D., & Kahn, L. M. (2000). Gender differences in pay. *The Journal of Economic Perspectives, 14*(4), 75–99.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature, 55*(3), 789–865.

Bowles, H. R., Babcock, L., & McGinn, K. L. (2005). Constraints and triggers: Situational mechanics of gender in negotiation. *Journal of Personality and Social Psychology, 89*(6), 951.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics, 29*(2), 238–249.

Card, D., Cardoso, A. R., & Kline, P. (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics, 131*(2), 633–686.

Castillo, M., Petrie, R., Torero, M., & Vesterlund, L. (2013). Gender differences in bargaining outcomes: A field experiment on discrimination. *Journal of Public Economics, 99*, 35–48.

Clarke, D. (2021). *RWOLF2: Stata module to calculate Romano-Wolf stepdown p-values for multiple hypothesis testing*. Boston College Department of Economics: Statistical Software Components.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica: Journal of the Econometric Society*, (pp. 829–844).

Cramton, P. C., & Tracy, J. S. (1992). Strikes and holdouts in wage bargaining: Theory and data. *The American Economic Review*, (pp. 100–121).

Crosetto, P., Weisel, O., & Winter, F. (2012). A flexible z-tree implementation of the social value orientation slider measure (Murphy et al. 2011)–manual–. *Jena Economic Research Papers*, 2012, 062.

Dittrich, M., Knabe, A., & Leipold, K. (2014). Gender differences in experimental wage negotiations. *Economic Inquiry, 52*(2), 862–873.

Eckel, C. C., & Grossman, P. J. (2001). Chivalry and solidarity in ultimatum games. *Economic Inquiry, 39*(2), 171.

Eckel, C. C., & Grossman, P. J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior, 23*(4), 281–295.

Embrey, M., Fréchette, G. R., & Lehrer, S. F. (2014). Bargaining and reputation: An experiment on bargaining in the presence of behavioural types. *The Review of Economic Studies, 82*(2), 608–631.
Exley, C. L., Niederle, M., & Vesterlund, L. (2020). Knowing when to ask: The cost of leaning in. *Journal of Political Economy, 128*(3), 816–854.
Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*(2), 171–178.
Gächter, S., & Riedl, A. (2005). Moral property rights in bargaining with infeasible claims. *Management Science, 51*(2), 249–263.
Gillen, B., Snowberg, E., & Yariv, L. (2019). Experimenting with measurement error: Techniques with applications to the caltech cohort study. *Journal of Political Economy, 127*(4), 1826–1863.
Gneezy, U., Niederle, M., Rustichini, A., et al. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics, 118*(3), 1049–1074.
Goldsmith-Pinkham, P., & Shue, K. (2020). *The gender gap in housing returns*. National Bureau of Economic Research: Technical report.
Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504–528.
Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association, 1*(1), 114–125.
Gu, A., & Yoo, H. I. (2019). vcemway: A one-stop solution for robust inference with multiway clustering. *The Stata Journal, 19*(4), 900–912.
Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In *Annals of economic and social measurement, volume 5, number 4* (pp. 475–492). NBER.
Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, (pp. 153–161).
Hernandez-Arenaz, I., & Iriberri, N. (2018). Women ask for less (only from men): Evidence from bargaining in the field. *Journal of Economic Behavior and Organization, 152*, 192–214.
Hernandez-Arenaz, I., & Iriberri, N. (2019). A review of gender differences in negotiation. *Oxford Research Encyclopedia of Economics and Finance*.
Hospido, L., Laeven, L., & Lamo, A. (2019). The gender promotion gap: Evidence from central banking. *Review of Economics and Statistics*, (pp. 1–45).
Huang, J., & Low, C. (2022). The myth of the male negotiator: Gender's effect on negotiation strategies and outcomes. *Journal of Economic Behavior & Organization, 202*, 517–532.
Kagel, J. H., Kim, C., & Moser, D. (1996). Fairness in ultimatum games with asymmetric information and asymmetric payoffs. *Games and Economic Behavior, 13*(1), 100–110.
Kugler, K. G., Reif, J. A., Kaschner, T., & Brodbeck, F. C. (2018). Gender differences in the initiation of negotiations: A meta-analysis. *Psychological Bulletin, 144*(2), 198.
Leibbrandt, A., & List, J. A. (2015). Do women avoid salary negotiations? Evidence from a large-scale natural field experiment. *Management Science, 61*(9), 2016–2024.
Ma, C. A., & Manove, M. (1993). Bargaining with deadlines and imperfect player control. *Econometrica: Journal of the Econometric Society*, (pp. 1313–1339).
Major, B., McFarlin, D. B., & Gagnon, D. (1984). Overworked and underpaid: On the nature of gender differences in personal entitlement. *Journal of Personality and Social Psychology, 47*(6), 1399.
Mazei, J., Hüffmeier, J., Freund, P. A., Stuhlmacher, A. F., Bilke, L., & Hertel, G. (2015). A meta-analysis on gender differences in negotiation outcomes and their moderators. *Psychological Bulletin, 141*(1), 85.
McDonald, J. F., & Moffitt, R. A. (1980). The uses of tobit analysis. *The Review of Economics and Statistics*, (pp. 318–321).
Mengel, F. (2015). Gender differences in networking.
Mischel, W. (1977). The interaction of person and situation. *Personality at the crossroads: Current issues in interactional psychology, 333*, 352.
Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. (2011). Measuring social value orientation. *Judgment and Decision Making, 6*(8), 771–781.
Nash Jr, J. F. (1950). The bargaining problem. *Econometrica: Journal of the Econometric Society*, (pp. 155–162).
Niederle, M., & Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics, 3*(1), 601–630.

Recalde, M., & Vesterlund, L. (2020). *Gender differences in negotiation and policy for improvement*. National Bureau of Economic Research: Technical report.
Rigdon, M. (2012). An experimental investigation of gender differences in wage negotiations.
Romano, J. P., & Wolf, M. (2005). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association, 100*(469), 94–108.
Romano, J. P., & Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica, 73*(4), 1237–1282.
Romano, J. P., & Wolf, M. (2016). Efficient computation of adjusted *p*-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters, 113*, 38–40.
Roth, A. E., Murnighan, J. K., & Schoumaker, F. (1988). The deadline effect in bargaining: Some experimental evidence. *The American Economic Review, 78*(4), 806–823.
Roussille, N. (2021). *The central role of the ask gap in gender pay inequality*. Berkeley: University of California.
Säve-Söderbergh, J. (2019). Gender gaps in salary negotiations: Salary requests and starting salaries in the field. *Journal of Economic Behavior & Organization, 161*, 35–51.
Sin, I., Stillman, S., & Fabling, R. (2020). What drives the gender wage gap? Examining the roles of sorting, productivity differences, bargaining and discrimination. *Review of Economics and Statistics*, (pp. 1–44).
Solnick, S. J. (2001). Gender differences in the ultimatum game. *Economic Inquiry, 39*(2), 189.
Sterbenz, F. P., & Phillips, O. R. (2001). Bargaining experiments with deadlines and random delays. *Economic Inquiry, 39*(4), 616–626.
Stuhlmacher, A. F., & Walters, A. E. (1999). Gender differences in negotiation outcome: A meta-analysis. *Personnel Psychology, 52*(3), 653–677.
Sutter, M., Bosman, R., Kocher, M. G., & van Winden, F. (2009). Gender pairing and bargaining—Beware the same sex! *Experimental Economics, 12*(3), 318–331.
Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics, 99*(1), 1–10.
Van Veldhuizen, R. (2022). Gender differences in tournament choices: Risk preferences, overconfidence, or competitiveness? *Journal of the European Economic Association, 20*(4), 1595–1618.
Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data* (pp. 521–524). MIT press.

Authors and Affiliations

Iñigo Hernandez-Arenaz[1] · **Nagore Iriberri**[2]

Iñigo Hernandez-Arenaz
iharenaz@gmail.com

1 Department of Economics, Universidad Pública de Navarra & INARBE, Campus Arrosadia, 31006 Pamplona/Iruña, Spain

2 Department of Economic Analysis, University of the Basque Country, UPV/EHU and IKERBASQUE, Basque Foundation for Science, Av. Lehendakari Aguirre 83, 48015 Bilbao, Spain