Does Information About the Parenthood Penalty Influence Support for Gender Equality?

Pre-Analysis Plan

Principal investigator: Matteo Gamalerio

16th of January 2026

1 Experimental design

We will conduct an incentivised online information provision experiment on a nationally representative sample from the general population in Spain. The sample consists of approximately 2,200 Spanish citizens recruited in partnership with the market research company Netquest. The provider complies with the General Data Protection Regulation (GDPR). The sample will be representative of the adult population of Spain, using quotas based on gender, age, and geographical regions (*Comunidades Autónomas*). Around two-thirds of the participants will be recontacted for a follow-up survey one or two weeks later, as part of the second wave. The survey will be conducted in Spanish, although this pre-analysis plan is presented in English for consistency. A full Spanish version of the questionnaire will be included in the online appendix of the paper.

All panel members will receive an invitation to participate via email and may freely choose whether or not to take part. The aim of this experiment is to understand whether providing information on the so-called "motherhood penalty" can increase support for gender-equality policies.

1.1 Main survey

Before entering the survey, panelists will be informed that the study is being conducted by researchers from the Universitat de Barcelona and will be required to provide explicit and informed consent. They will also be informed of their right to anonymity and their ability to withdraw from the survey at any time without any consequences (except for forfeiting the chance to receive the potential lottery award). An email address will be provided in case respondents wish to contact the research team with any questions or requests for clarification.

Demographics After providing explicit and informed consent, we ask respondents a battery of socio-demographic questions, including age, gender, highest level of education, and marital status. We restrict the sample to Spanish adults between 18 and 64 years old. We also elicit whether they currently have children or would like to have them in the future. Additionally, we ask about their current employment status and wage, following the questions from CIS (2025). We also collect information on political inclination and affiliation, as well as religion. Netquest will complete this set of demographic characteristics with geographical information on their municipality of residence, postal code, province, region and Nielsen macroregion. These data are updated approximately every six months. This information will enable linkage with electoral data, allowing us to identify whether the respondent lives in a municipality governed by a female or a male mayor. All data will be fully anonymised, making it impossible to reconstruct the identity or place of residence of participants. Crucially, we gather these demographic variables at the beginning of the experiment to assess whether there is self-selection in the survey completion.

Prior beliefs The participants are informed that the focus of this block of questions will be on wages and working conditions for men and women in Spain. Particularly, they are asked to think about men and women in Spain with similar characteristics and comparable jobs.
We elicit prior beliefs using both qualitative and quantitative approaches. First, we ask respondents whether they believe the gender wage gap exists and to explain what they think is the main reason behind it. Then, we invite them to estimate the salary of a woman with the same characteristics as a man who, applying the estimates by de Quinto et al. (2021), earns: (1) 100 € in the first scenario with no children, (2) 100 € one year after the birth of the first child, and (3) 95 € when the child is ten years old. For half of the respondents, the beliefs' elicitation is incentivised by an extra payment after the survey, in case of a correct answer (within a threshold of ± 2) in scenario (3). This is question 20 in the survey. This ensures that panelists are really putting effort in the task. It also reduces the opportunities for participants to engage in political motivated reasoning, which makes them to reply to questions on conditions (economic ones in Mian et al. (2023) and Prior et al. (2015)) as opinion questions. Regarding the scenario (3), we ask participants to what extent they are confident in the belief they have reported on the motherhood penalty. We control for the time spent in reading the questions and performing the task, as a proxy for commitment and attention. Prior beliefs are captured by means of sliders. By default, Netquest proposes sliders in which the starting position is on the very left side. This is the standard for respondents and to avoid confusion, we have kept the same setting (not considering a random positioning).

Treatment The respondents will be randomly assigned in equal proportion to either a *treatment* or a *control* group. These two experimental conditions differ in the amount of information received on the actual size of the motherhood penalty after the birth of the first child.

- *Treatment* group: Participants are first provided with a definition from the literature on the child penalty. They are then shown a colorful histogram indicating that, before having their first child, a woman and a man with the same characteristics exhibit no difference in earnings. Salaries are normalized to 100€. One year after the birth of the first child, a wage gap of 11€ per 100€ earned by a male worker emerges. When the child is 10 years old, the gap increases to up to 28 €. At this point, the salaries of both men and women have declined relative to the pre-birth level of 100€; however, while the man's wage declines only to 95€, the woman's wage falls to approximately 69€. Finally, the video presents two potential mechanisms explaining this gender wage gap: an increase in part-time work and a reduction in the number of days worked by women. Participants are informed that this evidence comes from a study conducted by researchers at the Bank of Spain.
- *Control* group: no additional information.

Given the nature of the design, individuals will belong to one of four groups:

$$\text{Groups} = \begin{cases} \text{Treatment with incentive} \\ \text{Treatment without incentive} \\ \text{Control with incentive} \\ \text{Control without incentive,} \end{cases}$$

where the incentive is derived from scenario (3) described in the prior beliefs subsection. Half of the respondents are randomly assigned to view a statement indicating that their belief elicitation is incentivized by an additional payment after the survey in the case of a correct answer. However, the main analysis will be focused on being part of either the treatment or the control group.
Some details of the design are worth noting. First of all, the information will be provided by means of a short and colourful video (less than 2 minutes), explaining the concept of the child penalty. The content of the video will be based on de Quinto et al. (2021). It will present the results from this study quantifying the gender wage gap attributable to the parenthood penalty, by

comparing the earnings trajectories of men and women with similar characteristics before having children, one year after the birth of the first child, and ten years later. The video will also explain the main mechanisms behind the penalty, such as the increase in part-time female employment and the reduction in working days. To enhance engagement and ease the understanding of the treatment, the video will include colorful clear histograms and in digits, animations, and visual aids (Haaland and Roth, 2023). Second, we provide the information treatment as an *indirect* feedback to participants' guesses at the beliefs elicitation stage. This characteristic of the design aims to reduce the experimenter demand effect (Haaland et al., 2023). Finally, the control group does not receive any additional relevant information on the topic of child penalty. Thus, we will be able to precisely capture the impact of the information on the participants' changes in beliefs and support for gender equality policies. (Haaland and Roth, 2020; Alesina et al., 2023)

Posterior beliefs and treatment comprehension We evaluate posterior beliefs about the magnitude of the motherhood penalty to determine whether the information treatment shifts these beliefs in the intended direction. Participants are then asked to estimate how much a woman earns when the child is fifteen years old. This question serves both to capture posterior beliefs and to verify comprehension of the treatment. An additional question on part-time employment and annual working days is included to further assess understanding.
The elicitation of these beliefs is not incentivized, as panelists might otherwise hedge their responses to maximize rewards. By comparing how posterior beliefs differ across groups, we can assess how participants update their views in response to the information provided.

Gender equality policies Our main outcomes variables are related to the perceptions regarding the introduction (or the possibility) of some gender equality policies. Respondents will be asked to what extent they agree or disagree with various government interventions, including the recent extension of paternity leave to 17 weeks, the non-transferability of parental leave, tax programs supporting families, fiscal incentives, and gender quotas in the labour market. To mitigate social desirability bias, each policy will be presented alongside both arguments in favour and against its adoption.
We also plan to construct an index capturing the overall support for gender equality policies.

Mechanisms Mechanisms will be explored through agreement with statements regarding the fairness of government interventions, perceptions of discrimination against men, and views on gender roles. The latter focuses on traditional norms assigning women the primary responsibility for childcare and domestic tasks over career advancement. We will also ask participants about their intention to support political parties that propose to expand or reduce gender equality policies. We will also investigate heterogeneous treatment effects depending on the characteristics of participants.

Behavioral outcomes We elicit whether the partcipants' beliefs about the existence of child penalty may affect their demand for information. We will present respondents with the opportunity to request additional information on the gender wage gap through a report published by the EsadeEcPol Center for Economic Policy (Hupkau and Contreras, 2025). Interested participants will receive access to the report at the end of the survey. Netquest gives panellists the option to directly download the content.We will track participants' interest in receiving additional information and use this as a behavioral outcome. The report includes findings from de Quinto et al. (2021), thereby also giving control group participants access to key information after completing the survey.
In addition, respondents will be informed that, upon entering the survey, they are automatically eligible for a lottery to win 600 points equivalent to a monetary value of 100 €. Participants will be asked how they would allocate a potential award between personal use and a donation to research promoting gender equality. The donation will be made after the draw, which is conducted by the survey company.

Conclusion In the final section, respondents will be invited to share personal experiences related to discrimination after taking parental leave. If they do not have children, they will be asked whether they are aware of unfair treatment experienced by relatives, friends, or colleagues. We will analyze responses to these open-ended questions using text analysis methods. We will obtain a measure of perceived political bias of the survey. At this stage of the survey, respondents interested in receiving the report will have the opportunity of doing it.

1.2 Second wave

Approximately two-thirds of the respondents will be followed up in a second, obfuscated survey conducted one or two weeks later. To reduce the likelihood of participants linking the follow-up to the initial survey, the first wave will be presented as a study conducted by researchers from the Universitat de Barcelona, while the second will appear to be carried out by researchers from Università Cattolica del Sacro Cuore (Italy). This approach is feasible, as the research team includes members from both institutions. The contact email used for the second wave will differ from that of the first and will feature an Italian domain. The follow-up survey will not adopt a different format, as this is standard for Netquest, but it will use revised wording for similar questions. The questionnaire is designed to be brief and will take approximately five minutes to complete.
We elicit respondents' beliefs about the importance of several societal issues, including immigration, overall public spending, and redistribution, apart from gender inequality. These questions are included to obscure the main purpose of the study and to test whether the information treatment has persistently altered respondents' beliefs.

Socio-demographics Following Haaland and Roth (2020), we aim to make the follow-up survey appear as an independent study. To this end, the questionnaire begins with a brief set of introductory questions on age, income, gender, and highest level of education. Although this information was already collected in the first wave, these questions are included to reinforce the perceived independence of the second survey.

Gender equality and other policies We replicate selected questions from the main survey related to gender equality policies. To minimise the risk that respondents recognise the connection between the two surveys, we rephrase the wording of these questions. In addition, we include items on other salient policy issues—such as immigration and government expenditures and taxes—to further diversify the content and mask the true focus of the survey.

Persistence of the treatment Using different wording, we re-ask our main quantitative outcome measure related to the treatment. Specifically, we ask respondents to evaluate the progression of a woman's salary compared to a man's salary ten or more years after the birth of the first child, assuming similar characteristics and occupations. Additionally, we assess second-order beliefs by asking respondents what they think the majority of people believe about this same topic.

1.3 Biases

We included a feedback entry field at the end of the survey to test whether respondents perceived the survey as politically inclined.

Bias in answer selection To detect the presence of extreme response bias, we will build an extreme response sum-score (ERS) to serve as a control in the analysis (Johnson et al., 2005). This sum is calculated by transforming each item into a binary variable (1 for an extreme response, 0 for a middle response) and then adding them up to create a single score that represents the individual's tendency to give extreme answers. We will re-run the analysis on a subsample of respondents who

have not provided only extreme responses in the robustness. In addition, to reduce response-order bias, we avoid long response lists and employ seemingly open-ended questions.

Experimenter Demand Effect Although experimenter demand effects are typically moderate (de Quidt et al., 2018), we adopt several measures to minimize this concern. First, the elicitation of prior beliefs is incentivized, while other behavioral measures carry monetary implications for participants, thereby reducing the risk of demand effects. By randomizing monetary incentives for prior beliefs among half of the participants, we can test for differences in accuracy between the two groups. Furthermore, the experiment is designed to preserve respondents' anonymity, making them less prone to experimenter demand effects. As suggested by Haaland et al. (2023), we will also conduct an obfuscated follow-up survey with the same respondents as in the initial experiment, in which dependent variables are elicited to estimate treatment effects, without respondents knowing that the two surveys are related. Lastly, we will test for bounding effect (de Quidt et al., 2018).

Social desirability bias Respondents are assured of complete anonymity in both the survey landing page and the consent form. We also consistently remind participants that we are only interested in their views, and that all responses are strictly confidential and anonymous. The block of questions related to support for gender equality policies presents arguments both in favor of and against these policies, thereby reducing the likelihood that respondents answer in a socially desirable way.

2 Setting, Sample size and restrictions

We will run our experiment through the certified market research company Netquest, which is widely used to conduct experiments. We will only recruit participants who currently live in Spain and have a Spanish nationality. Moreover, panelists must be between 18 and 64 years old. We will apply quotas depending on age group, gender and region of residence. Upon completion of the survey, Netquest pays participants a reward in korus (points), depending on the length of the survey. We apply several rules for inclusion in the survey. In particular, we exclude participants that:

- do not satisfy the quota requirement;
- do not pass the validation check;
- exit and then re-enter the survey as a new subject (as these individuals might see multiple treatments);

In total, we recruit 2,200 Spanish respondents. We will test whether the treatment and the control group will be balanced on pre-determined characteristics and variables. Additionally, half of the respondents will be eligible for a bonus if their prior belief in scenario (3) (question 20, as specified in the prior beliefs block) does not differ from the true value (67) by more than ± 2.
Lastly, upon entering the survey, participants are told to be eligible for a lottery of 600 korus (equivalent to 100€). We will have 6 winners, three from each group. Before knowing the result, participants may decide an amount that they would like to donate to a research project on gender equality. This information is **binding**.

3 Hypotheses

The information collected through the two survey waves will allow us to test the following set of hypotheses. First, we assess participants' knowledge of the child penalty and then examine the determinants of their prior knowledge about it:

1. Female respondents are more likely to be aware of the parenthood penalty;
2. Higher educational attainment is associated with greater awareness
3. Parental status affects awareness, with participants with and without children differing in their perceptions
4. Having a working mother is associated with greater awareness of the parenthood penalty
5. Younger and older individuals differ in their awareness levels
6. Socioeconomic status/income: The role of income is not straightforward and may operate in different directions. For example, higher-income individuals, who enjoy greater economic security and have better access to private solutions (e.g., childcare), may be less affected by the child penalty and therefore less aware of it than lower-income individuals. On the other hand, wealthier individuals may also be more educated and thus more aware of the issue.
7. Individuals with more extreme political inclinations—particularly conservative views—are less likely to be aware of the issue
8. Exposure to female leaders—such as a female mayor in the participant's municipality—is associated with greater awareness of the parenthood penalty

With regard to the experimental treatment, we hypothesize that receiving information on the existence and magnitude of the child penalty will:

9. Correct misperceptions related to the parenthood penalty
10. Increase support for gender equality laws and more egalitarian gender attitudes

We aim to identify the mechanisms underlying the observed effects, focusing on the role of gender norms, potential political backlash, and behavioural responses. With respect to the mechanisms, we expect that receiving the information treatment will:

11. increase the agreement that intervention to reduce inequalities between men and women in the workplace is fair;
12. decrease the agreement that gender equality policies may harm men;
13. decrease traditionalist views, such as that the role of the woman is to take care of the home and family, while increasing the belief that the current wage gap is unfair.
14. decrease agreement with the fact that mothers should sacrifice career, meaning respondents are less likely to attribute the wage gap to women's "choices".
15. increase the likelihood to vote for a party that promotes greater representation of women in leadership positions;
16. decrease the likelihood of voting for a party that promotes a model in which women take on a more central role in home and family care.
17. have ambiguous effects on the likelihood of voting for a party that proposes to remove the requirement for fathers to take the first six weeks of paternity leave immediately after birth, instead allowing families to freely choose when to use that time during the child's first year.

We hypothesize that:

18. Participants who initially underestimated the gender gap will be more likely to:
    (a) update their beliefs

(b) respond to the treatment by expressing greater support for gender equality policies

(c) respond to the treatment by expressing stronger disagreement with gender-stereotypical statements or traditional gender norms

19. Individuals who were already strongly supportive or opposed to gender equality may:

(a) exhibit limited change, with the most pronounced effects expected among those who initially held neutral views.

Through heterogeneity analysis, we expect the treatment effects to vary by demographic and attitudinal characteristics. For instance, we expect individuals with stronger confidence in their prior beliefs to be less responsive to new information.
Finally, we will examine the persistence of the treatment effects through the second, obfuscated survey wave. Specifically, we will test whether participants can accurately recall the magnitude of the parenthood penalty one or two weeks after receiving the information and whether their support for gender equality policies remains stable over time.

4 Analysis

4.1 Main analysis

We will match the dataset with electoral data and characteristics of local politicians based on the municipality of residence. Additionally, we will recover socio-demographic characteristics of the cities.

4.1.1 Baseline balance

We will test for baseline balance across treatment and control groups for the following variables:

- Gender - quota
- Age group - quota
- Geographical areas - quota
- Marital status
- Presence and number of children
- Desire of having children
- Log income (income is the midpoint of the interval selected)
- Employment status (dummies for unemployed, part-time employed and employed full-time)
- Education (dummy for person with at least bachelor degree)
- Political orientation (dummies for being left-wing, right-wing or center)
- Voting behaviour (dummies for voting for a left-wing, right-wing, center or another party, blank vote, no votar, no decision)
- Religion (dummies for being strongly religious, somewhat religious, no religious)
- Living arrangements as a minor (dummies for all the specifications)
- Employment situation of parents as a minor (dummies for all the specifications)

We will regress each of these variables on the treatment indicator to look for imbalances. Additionally, we will test for imbalances in prior beliefs across treatment and control groups. We expect no differences between treated and control units in terms of observable characteristics due to the random assignment. We will account for multiple hypothesis testing by regressing the treatment indicator on all of the variables, and we will conduct a joint F-test, to test if the coefficients are jointly different from zero. We will run specifications both without and with control variables. The control variables will be the pre-treatment variables used in the balance test. We include these controls to increase the precision of the estimates and to verify that the results are robust to the inclusion of variables that are unlikely to be unbalanced.

Lastly, we will obtain information on the general Spanish population to implement a mean comparison t-test on these variables and measure selection into the survey.

4.1.2 Differential Attrition

We will test whether attrition is related to the treatment by estimating the following equation:

$$A_i = \delta_0 + \delta_1 * Treatment_i + \eta' * \bar{X}_i + \epsilon_i, \tag{1}$$

where A_i indicates if a participant did finish our survey, $Treatment_i$ is a binary variable for being in the treatment group, and where $\bar{X}_i$ is a vector of pre-determined characteristics. We will use the same set of pre-determined characteristics as for the baseline balance test. If the coefficient δ_1 is not statistically significant, we will conduct the statistical analysis after dropping the units affected by attrition. If the coefficient δ_1 on the treatment indicator is statistically significant (using standard significance levels of 1, 5, and 10 percent), we will, as a robustness check, apply Lee bounds in the analysis. This approach will allow us to bound our estimates.

4.1.3 Attention check

Before the prior beliefs bloc, we present respondents with a simple attention check: they are asked to select the second option ("Brown") as their favorite color, even if it is not true. We will discard all the respondents that do not pass this validation. The survey company will run a second attention check before the mechanisms section. Netquest attention checks typically involve simple mathematical sums or questions about the current year.

4.1.4 Prior beliefs

We measure prior beliefs in terms of gaps. In the paper, we will plot the distribution of each of these variables, including a vertical line for the true value. We will look for under or over-estimating within boundaries, creating several measures (of ± 2, 5, 10, 20) and creating binary variables for being above or below the interval. We will also consider these variables:

- `Existence of gap` = 1 if they believe there is a gender wage gap
- `Gaps in beliefs _ no children` = Prior beliefs _ no children - 100 (the value they should have guessed);
- `Gaps in beliefs _ one year after` = Prior beliefs _ one year after - 89 (the value they should have guessed);
- `Gaps in beliefs _ when the child is ten years old` = Prior beliefs _ when the child is ten years old - 67 (the value they should have guessed);
- Confidence: for each of these variables, we assess their confidence in answering;
- Incentived = dummy for whether beliefs for `Prior beliefs _ when the child is ten years old` have been incentived and t-test to compare the differences between the groups

Additionally, we ask participants about what they think the main reason for gender wage gap is. We will analyse the short text adopting standard techniques of text analysis, including tokenization, stop-word removal, stemming or lemmatization. We will present results under means of word clouds, network graphs and topic visualization, among others. The goal is to analyse the sentiment related to the topic, including polarity, subjectivity and detection of emotions.

4.1.5 Treatment

We will create a dummy variable - `Treatment` - for whether the individual belonged to the treatment group or not. Respondents are required to visualize a short video and paying attention to it. They are instructed to continue with the questions by clicking "NEXT", once they have finished. It is not possible to proceed to the next section if they have not finished the video. At the end of the video, respondents may re-watch the information another time. Thus, we will measure the time spent on the video section and generate an indicator for having re-watched the video.

4.2 Outcomes

The estimation strategy will be mainly based on OLS regressions.

$$Y_i = \alpha + \beta * Treatment_i + \gamma' * \bar{X}_i + \epsilon_i \tag{2}$$

Where,

- Y_i is any of the outcome variables related to primary or secondary outcomes:
  - Primary outcomes: Posterior beliefs, Support for gender equality policies, Persistency;
  - Secondary outcomes: Behavioural outcomes, Mechanisms, Heterogeneity
- $Treatment_i = 1$ if the respondent is in the treatment group. We will investigate in the heterogeneity analsys the interaction between treatment and incentive;
- $\bar{X}_i$ is a battery of control of personal characteristics of the respondent and municipality characteristics. Since we are randomly allocating respondents to either the treatment or the control group, we expect characteristics to be balanced across the two groups. Thus, we only include controls to increase precisions.
  - Personal characteristics include: all the variables defined in the baseline part (refer to subsection 4.1.1), confidence in prior beliefs;
  - Municipal characteristics include: socio-demographic and geographical characteristics, electoral data (gender of the mayor and share of female councillors).
  - We will replicate the results by including either personal or municipal characteristics, and then present a specification that includes all of them. We will also report a specification without control variables.
- We will add province of residence FE and day of the interview FE to increase precision.
- Standard errors are robust to heteroskedasticity.

4.3 Main outcomes

4.3.1 Descriptive analysis

We will present descriptive statistics based on the pre-treatment characteristics.

4.3.2 Posterior beliefs

- Independent t-test to compare each of the Posterior beliefs variables in the Treatment and Control groups;
- Two-sided t-test to check whether `Posterior beliefs _ no children _ days worked` differ from 100 (the value respondents have to guess);
- Two-sided t-test to check whether `Posterior beliefs _ one year after _ days worked` differ from 85 (the value respondents have to guess);
- Two-sided t-test to check whether `Posterior beliefs _ when the child is ten years old _ days worked` differ from 74 (the value respondents have to guess);
- Check whether `Posterior beliefs _ when the child is ten years old _ part-time`: dummy for right answer ("Women have a higher probability");
- Two-sided t-test to check whether `Posterior beliefs _ when the child is 15 years old after _ earnings` differ from 60 (the value respondents have to guess);
- For each of the Posterior beliefs variables, we will construct several dummies related to: 1) correct answer within a range (we will iterate over several ranges, going from 0 to $\pm$ 10 from the correct value), 2) overestimate of the measure, 3) underestimate of the measure,
- We estimate linear regressions of posterior beliefs on a treatment dummy. We first present a specification without controls and then iteratively include a set of control variables capturing personal and municipal characteristics. Finally, we will test for heterogeneous treatment effects.

4.3.3 Support for gender equality policies

We construct several outcome variables to capture support for gender-equal family and labor-market policies. For each relevant survey question, we create the following measures: (1) an ordinal variable ranging from 1 to 5; (2) a dummy variable equal to 1 if the respondent selected option 1 or 2; (3) a dummy for the indifferent response; and (4) a dummy equal to 1 if the respondent selected option 4 or 5. In addition, 5) each item is rescaled to a 0-100 interval, where higher values indicate stronger support for gender equality. Finally, we construct a composite index by averaging across all items. We may even create a dummy for support of gender-equal family and labor-market policies including the indifferent answer.

Support for equal pay for work of equal value Do you think the government should make this legislation stricter, keep it as it is, or make it less strict?

- Ordinal variable ranging from 1 = "Much less strict" to 5 = "Much stricter".
- = 1 if the answer is 1) Much less strict; 2) Somewhat less strict;
- = 1 if the answer is 3) Keep as it is;
- = 1 if the answer is 4) Somewhat stricter; 5) Much stricter;
- **Equal-Pay Legislation Strictness (Index 1)**: Rescaling

$$\text{EqualPay}_i = \frac{\text{response} - 1}{4} \times 100.$$

Support for paternity leave Do you think paternity leave should be shorter, remain as it is, or even be extended?

- Ordinal variable ranging from 1 = "Much shorter" to 5 = "Much longer".
- = 1 if the answer is 1) Much shorter; 2) Somewhat shorter;
- = 1 if the answer is 3) Keep as it is;
- = 1 if the answer is 4) Somewhat longer; 5) Much longer;
- **Length of Paternity Leave (Index 2)**: Rescaling

$$\text{PaternityLeave}_i = \frac{\text{response} - 1}{4} \times 100.$$

Transferability of Parental Leave In a possible debate on this law, which of the following options do you prefer?

- Dummies for: 1) `NoTransfer`: Keep the current rule, non-transferable leave; 2) `FtoM`: Allow the father to transfer part or all of his leave to the mother; 3) `MtoF`: Allow the mother to transfer part or all of her leave to the father; 4) `BothWays`: Allow transfer in both directions, as decided by the family
- **Transferability of Parental Leave (Index 3)** We assign scores as:

$$\text{Transferability}_i = \begin{cases} 100 & \text{if leave is non-transferable (current rule)} \\ 75 & \text{if transfer allowed in both directions} \\ 50 & \text{if transfer allowed from mother to father} \\ 25 & \text{if transfer allowed from father to mother.} \end{cases}$$

Support for public spending Would you be willing for the government to increase spending on family policies, even if that means paying more taxes or reducing other budget allocations?

- Ordinal variable ranging from 1 = "Definitely in favor" to 5 = "Definitely against".
- = 1 if the answer is 1) Yes, definitely in favor; 2) Yes, somewhat in favor;
- = 1 if the answer is 3) Indifferent;
- = 1 if the answer is 4) No, somewhat against; 5) No, definitely against;
- **Public Spending on Family Policy (Index 4)**: Rescaling:

$$\text{FamilySpending}_i = \frac{5 - \text{response}}{4} \times 100.$$

Personal income tax deduction for mothers Do you think the Spanish government should reduce, maintain, or increase these kinds of incentives?

- Ordinal variable ranging from 1 = "Reduce significantly" to 5 = "Increase significantly".
- = 1 if the answer is 1) Reduce them significantly; 2) Reduce them slightly;
- = 1 if the answer is 3) Keep them as they are;
- = 1 if the answer is 4) Increase them slightly; 5) Increase them significantly;
- **Tax Incentives for Working Mothers (Index 5)**: Rescaling:

$$\text{TaxIncentive}_i = \frac{\text{response} - 1}{4} \times 100.$$

Gender quotas in the workplace and in politics To what extent do you agree or disagree with implementing mandatory gender quotas in these areas?

- Ordinal variable ranging from 1 = "Strongly disagree" to 5 = "Strongly agree".
- = 1 if the answer is 1) Strongly disagree; 2) Disagree;
- = 1 if the answer is 3) Neither agree nor disagree;
- = 1 if the answer is 4) Agree; 5) Strongly agree;
- **Mandatory Gender Quotas (Index 6)**: Rescaling:

$$\text{Quotas}_i = \frac{\text{response} - 1}{4} \times 100.$$

Pay transparency To what extent do you agree or disagree with this policy? Responses range from 1 = "Strongly disagree" to 5 = "Strongly agree".

- Ordinal variable ranging from 1 = "Strongly disagree" to 5 = "Strongly agree".
- = 1 if the answer is 1) Strongly disagree; 2) Disagree;
- = 1 if the answer is 3) Neither agree nor disagree;
- = 1 if the answer is 4) Agree; 5) Strongly agree;
- **Pay Transparency Directive (Index 7)** Rescaling:

$$\text{Transparency}_i = \frac{\text{response} - 1}{4} \times 100.$$

Composite Indices Finally, we define several composite indices:
1) **Gender-Equality Policy Support Index (GEPSI)** as the average of the seven rescaled scores above:

$$\text{GEPSI}_i = \frac{1}{7} \sum_{j=1}^{7} \text{score}_{ij},$$

Following Settele (2022), we standardize qualitative outcome measures based on the means and standard deviations in the control group. To account for multiple hypothesis testing, we construct 2) one summary index over each family of outcomes and test for the presence of an overall treatment effect on these indices. We will apply family-wise error rate (FWER) control. For the individual outcomes within these broad families, we present sharpened q-values based on false discovery rate (FDR) control. Both methods follow Anderson (2008).

4.4 Secondary outcomes

4.4.1 Robustness checks

- Sample excluding respondents who think the survey was biased;
- Sample excluding extreme patterns (namely first and last option to the questions);
- Sample excluding those who spent too little or too much time on questions (bottom 5 and 5% of the survey time distribution);

4.4.2 Mechanisms

In addition to our main outcomes, we measure potential mechanisms through which the treatment may affect attitudes, namely beliefs about gender roles and equality policies, in two blocs of questions.

Mechanisms: Gender Norms and Attitudes First of all, we ask to what extent they agree or disagree with certain statements. In line with the main outcome variables, we adopt the same methodology as before.

- **Government Intervention**: It is fair for the government to intervene with public policies to reduce inequalities between men and women in the workplace.
  * Ordinal variable ranging from 1 = "Strongly disagree" to 5 = "Strongly agree".
  * = 1 if the answer is 1) Strongly disagree; 2) Disagree;
  * = 1 if the answer is 3) Neither agree nor disagree;
  * = 1 if the answer is 4) Agree; 5) Strongly agree;
  * Rescaling:
$$\text{GovIntervene}_i = \frac{\text{resp} - 1}{4} \times 100.$$
- **Masculinity/Perceived Policy Fairness**: Public policies designed to improve women's situation in the labor market end up harming men, who lose opportunities or are treated unfairly
  * Ordinal variable ranging from 1 = "Strongly disagree" to 5 = "Strongly agree".
  * = 1 if the answer is 1) Strongly disagree; 2) Disagree;
  * = 1 if the answer is 3) Neither agree nor disagree;
  * = 1 if the answer is 4) Agree; 5) Strongly agree;
  * Rescaling: Since higher agreement reflects a less egalitarian stance, we reverse code:
$$\text{PolicyFairness}_i = \frac{5 - \text{resp}}{4} \times 100.$$
- `Family dynamics/Traditional Gender Roles`: The man's main role is to earn money and the woman's is to take care of the home and family
  * Ordinal variable ranging from 1 = "Strongly disagree" to 5 = "Strongly agree".
  * = 1 if the answer is 1) Strongly disagree; 2) Disagree;
  * = 1 if the answer is 3) Neither agree nor disagree;
  * = 1 if the answer is 4) Agree; 5) Strongly agree;
  * Rescaling: Reverse coded as:
$$\text{GenderRoles}_i = \frac{5 - \text{resp}}{4} \times 100.$$
- `Childcare/Mother's Career Sacrifice`: When a woman has children, she should be willing to partially or completely sacrifice her professional career to care for them.
  * Ordinal variable ranging from 1 = "Strongly disagree" to 5 = "Strongly agree".
  * = 1 if the answer is 1) Strongly disagree; 2) Disagree;
  * = 1 if the answer is 3) Neither agree nor disagree;
  * = 1 if the answer is 4) Agree; 5) Strongly agree;
  * Reverse coded as:
$$\text{MotherSacrifice}_i = \frac{5 - \text{resp}}{4} \times 100.$$

- For robustness, we define a **Gender Norms Egalitarianism Index (GNEI)** as the average of the four items above:

$$\text{GNEI}_i = \frac{1}{4}\big(\text{GovIntervene}_i + \text{PolicyFairness}_i + \text{GenderRoles}_i + \text{MotherSacrifice}_i\big).$$

  We report both individual items and the composite index.

Mechanism: Electoral Preferences on Gender Equality Platforms To capture the electoral implications of attitudes toward gender equality policies, respondents were asked about their likelihood of voting for a political party proposing different platforms. All items are measured on a 1–5 scale ranging from "No chance" (1) to "Certain that I would vote for them" (5).

- `Pro-Equality Party Platform`: Party proposes measures such as expanding paternity leave, increasing public support for working mothers, and introducing gender quotas in leadership positions.
  * = 1 if respondent has selected 1) No chance; 2) Low likelihood;
  * = 1 if respondent has selected 3) Neither low nor high likelihood;
  * = 1 if respondent has selected 4) High likelihood; 5) Certain that I would vote for them;
  * Rescaling:
$$\text{VoteProEquality}_i = \frac{\text{resp} - 1}{4} \times 100.$$
- `Anti-Equality Party Platform`: Party proposes limiting paternity leave, eliminating public support specifically for working mothers, and promoting a family model where women assume a more central role in home and family care.
  * = 1 if respondent has selected 1) No chance; 2) Low likelihood;
  * = 1 if respondent has selected 3) Neither low nor high likelihood;
  * = 1 if respondent has selected 4) High likelihood; 5) Certain that I would vote for them;
  * Rescaling: Since higher agreement here indicates support for a less egalitarian platform, we reverse code:
$$\text{VoteAntiEquality}_i = \frac{5 - \text{resp}}{4} \times 100.$$
- `Flexibility-Oriented Party Platform`: Party proposes removing the mandatory six weeks of immediate paternity leave, allowing families to choose freely when to use that time during the child's first year.

  This question does not map directly onto a pro- or anti-equality direction but instead captures attitudes toward flexibility. We therefore scale as:

$$\text{VoteFlexibility}_i = \frac{\text{resp} - 1}{4} \times 100.$$

4.4.3 Behavioural outcome: Demand for information

- t-test to compare demand for information in the Control and Treatment groups
- Linear regression of demand for report on a dummy for the treatment, including also a battery of control variables.
- Linear regression of the download of the reports (proxied by number of clicks depending on the group) on a dummy for the treatment, including also a battery of control variables.

4.4.4 Behavioural outcome: Lottery

- t-test to compare amount donated to an academic research project on gender equality in the Control and Treatment groups
- Linear regression of the amount donated on a dummy for the treatment, including also a battery of control variables.

4.4.5 Heterogeneous Treatment Effects

We will present heterogeneity analysis. In this setting, the equation will be of the form:

$$Y_i = \alpha + \beta * Treatment_i + \delta * Z_i + \theta * (Treatment_i * Z_i) + \gamma' * \bar{X}_i + \epsilon_i \quad (3)$$

Where Z_i is one of the pre-determined characteristics:

- Gender;
- Education level ;
- Parental status (current or intended);
- Parental employment background;
- Age group;
- Ideological orientation;
- Exposure to female leading figures, such as the mayor of their municipality.
- Confidence in prior beliefs;
- Exposure to the incentive statement

4.4.6 Personal experience

Lastly, we ask respondents a battery of questions related to their experience with gender equality issues, either in the company in which they work or in the company they own, if applicable.

Additionally, we collect text responses depending on whether having children has affected their employment situation or they have witnessed an episode of impact of parenting of people close to them (partner, family members, friends, coworkers) on their employment. We will analyse these answers with text analysis methods.

4.5 Second wave

4.5.1 Baseline balance

We will test for baseline balance for the following variables:

- Gender
- Age group
- Log income (income is the midpoint of the interval selected)
- Employment status (dummies for unemployed, part-time employed and employed full-time)
- Education (dummy for person with at least bachelor degree)

The survey company is not able to ensure quotas for the second wave. We will regress each of these variables on the treatment indicator to look for imbalances. We will account for multiple hypothesis testing by regressing the treatment indicator on all of the variables, and we will conduct a joint F-test, to test if the coefficients are jointly different from zero.
We test to see if there are any differences with respect to the baseline characteristics. In case there are differences, we will talk with the survey company to get a comprehensive view of the individual but we will consider characteristics from the first round.

4.6 Obfuscation

We ask respondents a battery of questions related to other topics:

- Main problem currently facing Spain → we will create a dummy if the panellist has selected `inequalities, gender, class, poverty`;
- Level of immigration
- Overall public spending;
- General level of taxes;
- Redistribution of wealth

We will use answers to these questions as placebo outcomes. Additionally, we will randomize the order of these questions with the questions of the following block.

4.7 Main outcome variables

In the second wave of the survey (*W2*), respondents were asked additional questions on beliefs and preferences regarding gender equality policies and parental leave. We will analyse these variables by 1) considering the rank of values, 2) creating binary variables and 3) rescaling values to a 0-100 interval, where higher values indicate more egalitarian beliefs or greater support for gender equality measures.

Government Intervention Do you think the government should or should not intervene with public policies to reduce gender inequalities in the workplace?

- = 1 if the respondent selects "Yes, the government should intervene with public policies to reduce inequalities", while the rest is 0;
- = 1 if the respondent selects "Yes, the government should intervene with public policies to reduce inequalities", while = 0 if the respondent selects "No" and missing if he or she decides for "I do not know";
- Scaling:

$$\text{GovIntervene}_{i,\text{W2}} = \begin{cases} 100 & \text{Yes, government should intervene} \\ 0 & \text{No, government should not intervene} \\ \text{missing} & \text{No opinion} \end{cases}$$

Perceived Negative Effects of Equality Policies Do you think these policies end up harming men, for example by causing them to lose job opportunities or be treated unequally?

- = 1 if the respondent selects "No, I don't think these policies harm men" and 0 otherwise;
- = 1 if the respondent selects "No, I don't think these policies harm men", = 0 if the panelist selects "Yes" and missing otherwise;

- Rescaling Reversed coding so higher values reflect more egalitarian perception:

$$\text{PolicyFairness}_{i,\text{W2}} = \begin{cases} 100 & \text{No, policies do not harm men} \\ 0 & \text{Yes, policies harm men} \\ \text{missing} & \text{No opinion} \end{cases}$$

Parental Leave Flexibility To what extent do you agree or disagree with this proposal?

- Ordinal variable ranging from 1 = "Strongly disagree" to 5 = "Strongly agree".
- = 1 if the answer is 1) Strongly disagree; 2) Disagree;
- = 1 if the answer is 3) Neither agree nor disagree;
- = 1 if the answer is 4) Agree; 5) Strongly agree;
- Rescaling:

$$\text{ParentalFlexibility}_{i,\text{W2}} = \frac{\text{resp} - 1}{4} \times 100$$

Spanish vs Italian Leave Model Preference Which of the two models do you think is better for promoting family well-being?

-

$$\text{LeaveModelPreference}_{i,\text{W2}} = \begin{cases} 100 & \text{Spanish model (more egalitarian)} \\ 0 & \text{Italian model} \\ \text{missing} & \text{None / Don't know} \end{cases}$$

Support for Gender Quotas To what extent do you agree or disagree with establishing mandatory gender quotas in these areas?

- Ordinal variable ranging from 1 = "Strongly disagree" to 5 = "Strongly agree".
- = 1 if the answer is 1) Strongly disagree; 2) Disagree;
- = 1 if the answer is 3) Neither agree nor disagree;
- = 1 if the answer is 4) Agree; 5) Strongly agree;
- Rescaling:

$$\text{GenderQuotaSupport}_{i,\text{W2}} = \frac{\text{resp} - 1}{4} \times 100$$

4.8 Persistence

We measure respondents' beliefs about the long-term effects of having children on the careers of men and women. Two types of responses are collected: personal opinion and social belief.

Personal Belief Do you believe that, in the long term (10 years or more) after the birth of the first child, the impact on salary is the same for men and women with similar characteristics and jobs?

- =1 if respondent has selected "No, the negative impact is greater for women"
- Higher values indicate the respondent believes women suffer a larger long-term career penalty.

$$\text{PerceivedMotherhoodPenalty}_{i,\text{W2}} = \begin{cases} 0 & \text{Yes, impact is similar} \\ 100 & \text{No, impact greater for women} \\ -100 & \text{No, impact greater for men} \\ \text{missing} & \text{I don't know / Prefer not to answer} \end{cases}$$

Perceived Social Belief Out of every 100 people, how many do you think would choose each of the following options?

- p_{equal} = % who say impact is similar
- p_{women} = % who say women are more negatively affected
- p_{men} = % who say men are more negatively affected
- p_{DK} = % who don't know / prefer not to answer

Respondents estimate, out of 100 people, how many would think that long-term career effects of having children are the same for men and women, greater for women, or greater for men. We summarize these responses with the *Social Motherhood Penalty Index*:

$$\text{SocialMotherhoodPenalty}_{i,\text{W2}} = p_{\text{women}} - p_{\text{men}},$$

where p_{women} and p_{men} are the percentages assigned to "women more negatively affected" and "men more negatively affected," respectively. Values range from -100 to 100, with higher values indicating that the respondent believes society perceives a larger negative impact on women.

4.8.1 Additional analysis

In the first part of the study, respondents were asked about their interest in downloading a report. We are now interested in testing whether actually downloading the report affected outcomes. Specifically, we expect treatment persistence to be higher among those who downloaded the report compared with 1) respondents who were not interested in receiving it and 2) those who expressed interest but ultimately did not download it.

References

A. Alesina, A. Miano, and S. Stantcheva. Immigration and Redistribution. *The Review of Economic Studies*, 90(1):1–39, 2023. URL https://doi.org/10.1093/restud/rdac011.

M. L. Anderson. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103:1481–1495, 2008.

CIS. Barómetro. *Centro de Investigaciones Sociológicas*, 2025.

J. de Quidt, J. Haushofer, and C. Roth. Measuring and bounding experimenter demand. *American Economic Review*, 108(11):3266–3302, 2018.

A. de Quinto, L. Hospido, and C. Sanz. The child penalty: Evidence from spain. *SERIEs: Journal of the Spanish Economic Association*, 12(4):585–606, 2021.

I. Haaland and C. Roth. Labor market concerns and support for immigration. *Journal of Public Economics*, 191, 2020.

I. Haaland and C. Roth. Beliefs about Racial Discrimination and Support for Pro-Black Policies. *The Review of Economics and Statistics*, 105(1):40–53, 2023. doi: 10.1162/rest_a_01036.

I. Haaland, C. Roth, and J. Wohlfart. Designing Information Provision Experiments. *Journal of Economic Literature*, 61(1):3–40, March 2023. doi: 10.1257/jel.20211658. URL https://www.aeaweb.org/articles?id=10.1257/jel.20211658.

C. Hupkau and J. Contreras. La brecha salarial de género en España: ¿qué factores contribuyen a su persistencia? *Es!deEcPol Policy Brief*, 50, 2025.

T. Johnson, P. Kulesa, Y. I. Cho, and S. Shavitt. The relation between culture and response styles: Evidence from 19 countries. *Journal of Cross-cultural psychology*, 36(2):264–277, 2005.

A. Mian, A. Sufi, and N. Khoshkhou. Partisan bias, economic expectations, and household spending. *The Review of Economics and Statistics*, 105(3):493–510, 2023.

M. Prior, G. Sood, and K. Khanna. You cannot be serious: The impact of accuracy incentives on partisan bias in reports of economic perceptions. *Quarterly Journal of Political Science*, 10(4): 489–518, 2015.

S. Settele. How do beliefs about the gender wage gap affect the demand for public policy? *American Economic Journal: Economic Policy*, 14:475–508, 2022. doi: https://doi.org/10.1257/pol.20200559.