ROCKWOOL Foundation Berlin
Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES 056/26

Misunderstandings of the Child Penalty Graph

Olle Folke, Johanna Rickne

www.rfberlin.com FEBRUARY 2026

Misunderstandings of the Child Penalty Graph

Authors

Olle Folke, Johanna Rickne

Reference

JEL Codes: J01

Keywords: Child penalty, gender inequality, graph design

Recommended Citation: Olle Folke, Johanna Rickne (2026): Misunderstandings of the Child Penalty Graph. RFBerlin Discussion Paper No. 056/26

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com

Misunderstandings of the Child Penalty Graph

Olle Folke and Johanna Rickne[1]

Abstract

We study misunderstandings of a widely used research graph that visualizes how parenthood affects gender inequality in the labor market. The child penalty graph typically displays overlapping trend lines for women and men before parenthood and a sharp divergence thereafter. A large survey experiment asks participants about the gender gap in earnings before and after random assignment across three alternative graph designs. Approximately one-third of respondents misinterpret the distance between the trend lines in the *child penalty graph* as the gender gap in earnings. After viewing the graph, nearly 50% of respondents (and 60% of those with a PhD) report an earnings gap of 0–5% before parenthood, far below the true gap of about 20%. We examine two consequences of this misunderstanding and assess whether alternative graph designs improve comprehension.

[1] Pre-registration at https://osf.io/8gjdy/overview. Folke: Uppsala University, Gamla torget 6, 751 20 Uppsala, Sweden (olle.folke@statsvet.su.se); Rickne: Swedish Institute for Social Research, Stockholm University, S-106 91 Stockholm, Sweden (Johanna.rickne@sofi.su.se). The authors thank Yoko Okuyama, Pamela Campa, Miguel Pereira, Petter Lundborg, Eleonora Alabrese, Ylva Moberg, and seminar participants at the Stockholm School of Economics/SITE, and ENS Paris-Saclay for helpful comments. The authors thank the Knut and Alice Wallenberg Foundation and the Swedish Research Council for financial support.

Introduction

Research on parenthood and economic gender inequality has a long history in economics.[2] Two papers published in the late 2010s revitalized this literature and provided one of the most influential research findings in the last decade of labor economics. These papers by Kleven (2019a, 2019b) use event study methods to demonstrate that parenthood generates a *child penalty* in earnings for women, but not for men.[3] A striking graphical visualization helped popularize their methodology across academic disciplines.[4] This *child penalty graph* typically shows the same pattern for different countries and labor market variables as the original rendition for earnings among Danish parents (see Figure 1A). Two trend lines for women and men overlap close to zero in the years before parenthood, followed by a sharp drop in women's trend line when the first child is born.

Outside the walls of academia, the *child penalty graph* continues to influence the conversation about gender inequality in the labor market through extensive media coverage (examples in Appendix Figures A1–A2). The researchers maintain a website called the Child Penalty Atlas, which allows users to click on a world map to obtain country-specific child penalty graphs (Figure 1B shows such a graph for Denmark; see also Kleven et al. 2024).

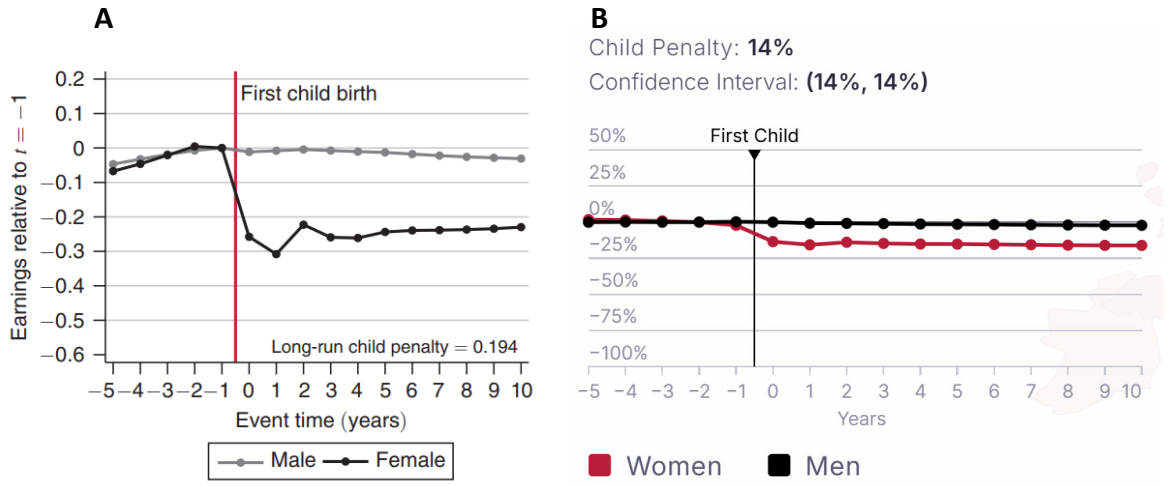


Figure 1. Example graphs from Kleven et al. (2019a) and the Child Penalty Atlas website.

Notes: Figure 1A reproduces Panel A in Figure 1 from Kleven et al. (2019a); and Figure 1B is a screenshot of the "Child penalty in employment" results for Denmark from childpenaltyatlas.org, accessed on December 12, 2025.

[2] Particularly prominent contributions include Reid (1934), Becker (1985), Goldin (1990), Korenman and Neumark (1992), Folbre (2001), and Bertrand et al. (2010).
[3] Earlier use of event study methods to study impacts of parenthood included Korenman and Neumark (1992) and Waldfogel (1997), with methodological refinements provided by, e.g., Anderson et al. (2003).
[4] Kleven et al. (2019a, 2019b) had together accumulated 3,958 Google citations by December 2025, and a search for "child penalty" on Google scholar that month returned 2,860 results. While the research literature using their child penalty graph is too vast to review, Appendix Table A1 lists studies in the Swedish context.

We use survey experiments to study misunderstandings of the *child penalty graph*. Our main tests examine whether people mistakenly interpret the gap between the trend lines for women and men as the gender gap in the outcome variable. We also test if they incorrectly read the graph's child penalty statistic as the size of that gender gap.

The misunderstanding might arise from typical presentations of the *child penalty graph*. Presenting the graph as showing how parenthood impacts gender inequality might make viewers interpret the trend line pattern as information about inequality levels. They might interpret the overlapping trend lines before parenthood as evidence that women and men have roughly equal outcomes, and the vertical distance after parenthood as the inequality level after childbirth. This differs from the correct interpretation, which is more complex. The regression analysis underpinning the graph creates a temporal benchmark at the year immediately before parenthood. The trend lines show estimated differences in the outcome relative to this year, recalculated after the estimation as percentages of what the outcome would have been if parenthood had not occurred.[5]

Our pre-registered survey experiment studies how Swedish respondents in a non-commercial research panel (N=12,130) interpret the *child penalty graph* for labor market earnings. We ask respondents to state the gender gap in earnings among Swedish adults "who are not yet parents" and "whose children are about 10 years old." These questions appear twice, before and after respondents view a research graph. We randomly assign respondents across three graphs: a *child penalty graph* like those on the website for the Child Penalty Atlas, and two alternative designs.

We test for the misunderstanding by analyzing whether viewing the *child penalty graph* makes respondents change their answers about the gender gap in earnings in ways that correspond to the predicted misunderstanding. We ascertain that changes stem from the graph's specific features by analyzing whether they *fail to appear* for respondents assigned to an alternative graph design that displays the same research result without these features. This

[5] Details in Appendix C. A more exact interpretation can be derived from the structure of the regression equation. The analysis uses panel data for first-time parents and regresses the labor market variable on dummies for event year (excluding t = -1), age, and calendar year. Consider the identifying variation when analyzing women's earnings. Removing variation by age and calendar year leaves variation mainly between women who became parents at different ages. Two 30-year-olds in the same calendar year can be at different event times if one had a child at 25 and the other at 31. The estimated coefficient for t = 5 compares earnings for women who have been parents for 5 years (those who had children at 25) with those who had their first child the next year (those who became parents at 31). Because coefficients mainly capture cross-sectional variation in the age when women have their first child, endogeneity concerns arise if women time their decision to become parents based on their earnings profile (see, e.g., Bensnes et al. 2026).

gender gap graph makes the potentially misunderstood information explicit in two ways. It combines one panel with event-study estimates for the gender gap in earnings with a second panel for women's and men's average earnings in the event window (Figure 2B).

The third graph design in the experiment evaluates whether a simple redesign of the original graph can mitigate the potential misunderstanding. This *child penalty graph with raw means* combines one panel showing the standard *child penalty graph* and a second panel showing average earnings by gender (Figure 2C).[6]

The results provide strong evidence of the misunderstanding. Seeing the *child penalty graph* makes one in three respondents change their answer about the gender gap in earnings before parenthood to a small (or zero) gap. A similar fraction of respondents change their answer about the size of the earnings gap after parenthood to reflect either the distance between the trend lines in the graph or the superimposed child penalty statistic. Neither of these changes occur for respondents presented with the *gender gap graph*. The addition of raw means to the child penalty graph mitigates our specific misunderstanding of interest, but causes other problematic misinterpretations.

We examine two potential consequences of the misunderstanding. Misinterpreting the graph as showing how parenthood turns initial gender parity into a large disadvantage for women might make people consider parenthood to be a more important explanation for gender inequality in the labor market. The results provide some evidence of this consequence. In contrast, we find little evidence that misunderstanding the graph makes people more likely to engage with it on social media.

Our study has three main weaknesses. The measurement method for detecting the misunderstanding might overestimate its prevalence. Asking questions about the size of the gender gap in earnings immediately after showing a research graph might push respondents to mistakenly extract this information from the figure. We sought to prevent this by adding a statement to these questions saying that "if you don't think the graphs provided useful information, you may enter the same answer as before" (i.e., the same answer about the gender gap in earnings). We nevertheless acknowledge that our method likely triggered information extraction that would not have happened otherwise.

[6] Showing raw means can help viewers understand event-study graphs (e.g., Miller 2023), and users on childpenaltyatlas.org can display a world map colored to indicate cross-country variation in the size of gender gaps.

We run a supplementary survey to evaluate the sensitivity of our main result to two alternative measurement methods (not pre-registered). Social media posts recruited a medium-sized sample (N=779) with 41% academic economists. Results from these data demonstrate similar levels of the misunderstanding as the main analysis when (1) directly asking whether we "learn" the misunderstood information from the graph (i.e., if we learn that the gender gap in earnings is 0—5% before parenthood), and (2) using free-text questions to allow respondents to interpret the graph in their own words.

A second weakness may derive from the specific *child penalty graph* in the main survey. Our chosen design matches the typical design in government reports (examples in Appendix Figure A3). This design likely underestimates the prevalence of the misunderstanding relative to more simplified reproductions in the media (examples in Figures A1–A2). Such reproductions typically lack a label on the y-axis or indications of the temporal benchmark. A striking example comes from the Nobel Foundation's illustration of 2023 Laureate Claudia Goldin's research (see Appendix Figure A1). This illustration shows the iconic pattern from the *child penalty graph*, but strengthens the impression of equal outcomes before parenthood by labeling the y-axis with a single dollar sign and adding cartoon figures of a man and woman walking hand in hand on the overlapping trend lines.

An opposite situation of over-estimating the prevalence of the misunderstanding might apply to our chosen design relative to more technical designs in research papers. Results from our supplementary survey speak to this concern. We find similar levels of the misunderstanding among respondents who were randomly assigned to view the graph from the survey or a more technical design from the original research article (i.e., Figure 1A from Kleven et al. 2019a).

A third weakness concerns potential trade-offs in graph design. Our study might be uncovering misunderstandings about a relatively marginal phenomenon (gender inequality levels) while ignoring correct communication about something more important (the size of the child penalty). This argument is particularly problematic for gender inequality before parenthood. Childless individuals are not a marginal group but account for 20—30% of the working-age population in our Swedish microdata.[7] These data show an (age- and year-adjusted) earnings gap of approximately 20% before parenthood, far above the 0–5% belief held by roughly 50% of our survey respondents after seeing the *child penalty graph* (and 60%

[7] In our microdata, childless individuals under age 50 account for 34% of the working-age population, those under 40 account for 29%, and those under 30 for 21%.

of respondents with a PhD).[8] It is also questionable if the *child penalty graph* efficiently communicates information about the child penalty. Responses to our free-text questions in the supplementary survey show that at most 40% of academic economists could correctly interpret the child penalty statistic, and 15–20% of all respondents.

Our results have real-world implications. They document that a common graph in applied economics creates misperceptions about an important economic phenomenon. Economic gender inequality hurts individuals and undercuts aggregate economic output (e.g., Goldin and Rouse 2000, Aizer 2010, Hsieh et al. 2019, Gagnon et al. 2025). Our results indicate that misperceptions created by the *child penalty graph* may hinder efforts to address this problem. Recent evidence from survey experiments shows that people who perceive smaller economic gender gaps are less likely to support policy action in this area (Settele 2022, Casarico et al. 2024).

We do not recommend a specific alternative graph design. Adding raw means to the *child penalty graph* mitigates the specific misunderstanding that we study, but creates other misinterpretations related to inequality levels and changes with parenthood. The *gender gap graph* conveys information about inequality levels more accurately than the *child penalty graph*, but applies a similar event-study methodology. Unless this methodology can deliver credible causal evidence about the impact of parenthood (e.g., Lundborg et al. 2017, 2024, Adams-Prassl et al. 2024, Bensnes et al. 2026), "no graph" might ultimately be the best recommendation.

Main survey experiment

Survey data and questions. We administered a survey with nine questions in Swedish (Appendix Tables B1–B2) to the Swedish Citizen Panel, a non-commercial online panel managed by the University of Gothenburg.[9] A first round of 16,000 invited participants was selected by stratified random sampling based on gender and age from the full panel (~70,000 persons). A second round of 1,779 invitations went to any panelist who was not invited in the

[8] Large gender inequality before parenthood is not an artifact of our Swedish data. Costas Dias et al. (2020) calculate a 10% wage gap before parenthood in the UK, and Cortés et al. (2023) establish a 10% wage gap immediately after labor market entry among US college graduates.

[9] For more information, see https://www.gu.se/en/som-institute/the-swedish-citizen-panel. Approximately 75% of the registered panelists joined voluntarily after seeing advertisements, for example when taking Electoral Compass tests linked to major Swedish newspapers such as *Aftonbladet*. The remaining 25% joined after a direct invitation from the University of Gothenburg, periodically sent out to randomly stratified samples of the Swedish adult population selected by Statistics Sweden for the purpose of improving social representativeness.

first round and who had a PhD or current job related to research communication (journalists, public relations specialists, public information officers, and communications officers). Participation was voluntary and we received 12,130 responses (a 68.2% response rate).

We described the survey topic as "parenthood and gender inequality" and the purpose as "investigating how people interpret the content of graphs presenting research results on this topic." The first two questions defined the gender gap in annual earnings[10] and asked about its size before and after parenthood without providing monetary incentives.[11] These questions asked respondents to "[t]hink of adults in Sweden who [have not yet had children/ who have become parents and whose child is around 10 years old]. If you were to guess, by what percentage do you think women's annual earnings are lower than men's?" Respondents entered their answer in a text box in a second sentence stating that "Among people [who are not yet parents/who are parents and whose oldest child is about 10 years old], women's average annual earnings are [text box] percent lower than those of men."

The survey introduced the graph by stating that "You will now see a graph showing research results on the earnings gap between women and men. The purpose of this research is to examine how having children affects economic gender equality. More precisely, the research examines whether women incur what is referred to as a 'child penalty,' that is, a reduction in earnings compared to men upon entering parenthood." We worded this text to mirror introductory explanations in the original research. The abstract for Kleven et al.'s (2019a) abstract explains that the paper studies "how children affect gender inequality" and finds as that "the arrival of children creates a long-run gender gap in earnings of around 20 percent." The methods subpage of the Child Penalty Atlas website describes the child penalty as parenthood's "long-term effect on opportunities of women and men" and the analysis in the site's graphs as "look[ing] at the earnings or employment of women and men around the birth of their first child and see[ing] what happens to women's employment relative to men."

We informed respondents that "[a]fter viewing the graphs, we will again ask questions about the gender gap in earnings." The two questions directly after the graph again asked about the gender gap in earnings, using the same wordings but adding the prompt discussed in the

[10] "[n]ow and then, there is talk of a gap in earnings between women and men, with women typically earning less from work or entrepreneurship over the course of a year."
[11] Settele 2022 finds highly similar belief distributions about the gender wage gap for incentivized and non-incentivized questions.

paper's introduction (i.e., that respondents should give the same answer as before if the graph did not think the graph provided the information).

Follow-up questions after the graph asked if respondents found it helpful for answering the questions about the earnings gap, and how certain they were of their answers. In the last two questions, they rated the importance of three explanations for gender inequality in the labor market, as well as how likely they would be to "like," share, or cite-post the graph on social media. The survey ended with a debriefing about gender inequality in earnings.

Research graphs. We create three graphs using Swedish administrative data on all first births between 2002 and 2012. Following Kleven et al. (2019b), we only include parents in the analysis sample if we can observe both of their earnings for five consecutive years before parenthood, and 10 consecutive years afterwards. We summarize the methods below and provide details in Appendix C.

We create the *child penalty graph* (Figure 2A) using replication files from Kleven et al. (2019a). The line colors and y-axis labels match the graph design on childpenaltyatlas.org (Figure 1A).[12] We add the title "earnings effect of becoming a parent, child penalty" and use quotation marks around "child penalty" to indicate the usage of an English term.[13]

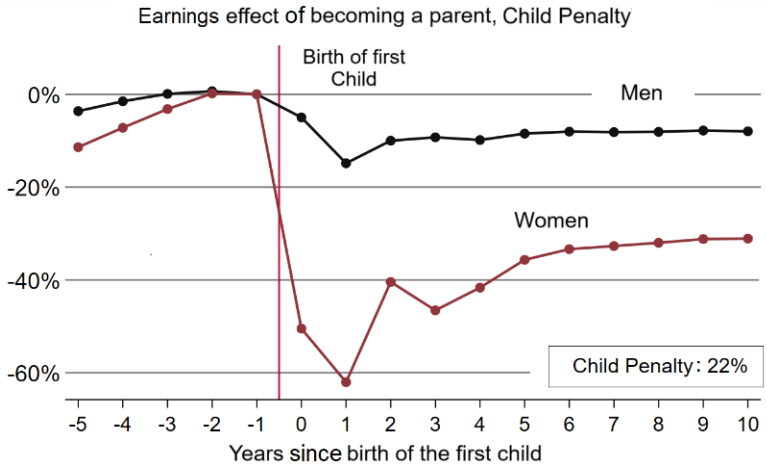


Figure 2A. Child penalty graph in the main survey experiment.

Notes: The graph was created using Swedish administrative data for annual earnings and the replication files from Kleven et al. (2019a). Details in Appendix C.

Our *child penalty graph* displays the standard pattern of overlapping trend lines around zero before parenthood and a subsequent drop in the women's trend line. These lines capture the covariate-adjusted earnings of each binary gender relative to the year immediately before

[12] We sought replication files for the slightly different method in the Child Penalty Atlas (Kleven et al. 2024), but had not received a response from the authors or located the replication files for this paper as of July 2025.
[13] Texts in Swedish generally use the English term because Swedish translations like *barnstraff* sound awkward.

parenthood, normalized as a percentage of that gender's counterfactual earnings if they did not have any children (see Appendix Figure C1). The graph exhibits the typical pattern, but with slightly different pre-trends by gender. This is likely because Sweden's parental leave program gives women substantial economic incentives to hold paid employment before childbirth.

The *gender gap graph* (Figure 2B) has two panels. The left panel displays women's and men's average annual earnings in each event year, including people with zero earnings. The right panel presents regression estimates for the gender gap in annual earnings. We obtain these by pooling women and men and restating the original regression as an interaction model. The graph plots the point estimates on the interaction terms between the event-year dummies and a dummy for female sex at birth. Keeping all event-year dummies in the regression removes the temporal benchmark in the year immediately before parenthood. We normalize each point estimate for women's relative earnings as a percentage of men's counterfactual earnings if they did not have any children (details in Appendix C).

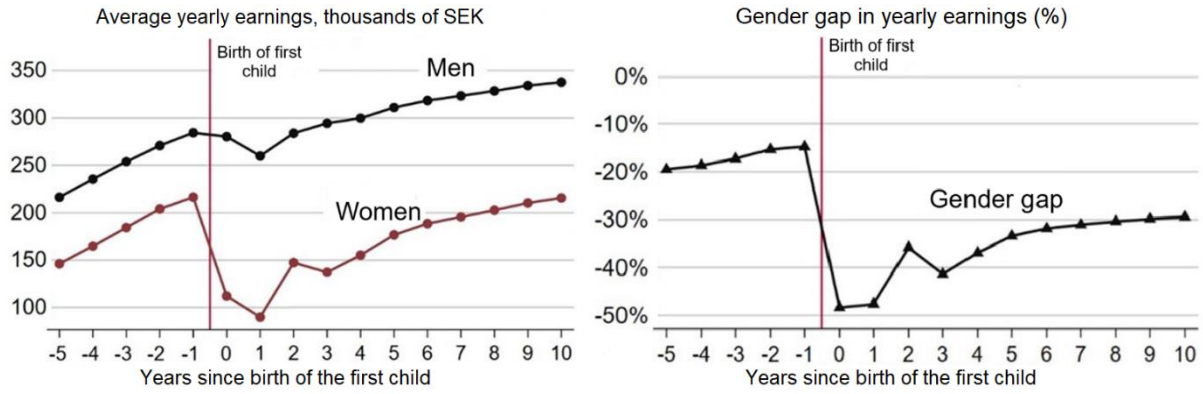


Figure 2B. Gender gap graph in the main survey experiment.

Notes: Graphs were created using Swedish administrative data for annual earnings. Appendix C provides details about the event-study estimates presented in the right graph. 250,000 SEK ≈ 30,000 USD.

The *gender gap graph* shows the same research finding as the *child penalty graph*, but replaces the gender-specific trend lines and child penalty statistic with elements that make the potentially misunderstood information explicit. Viewers may assess the *adjusted* gender gap in earnings from the normalized event-study estimates and calculate the *raw* gap from the plot of average earnings by gender. We hold the finding constant by duplicating Kleven et al.'s (2019a) original sample restrictions, control variables, and normalization procedures. However, we follow conventional methods and calculate the gender gap in earning by dividing the earnings difference between men and women by men's earnings. This differs from the division by *women's earnings* in Kleven et al.'s (2019a) calculation of the child penalty statistic.

Figure 2C combines the child penalty graph with the subplot of average earnings by gender into a *child penalty graph with raw means*.

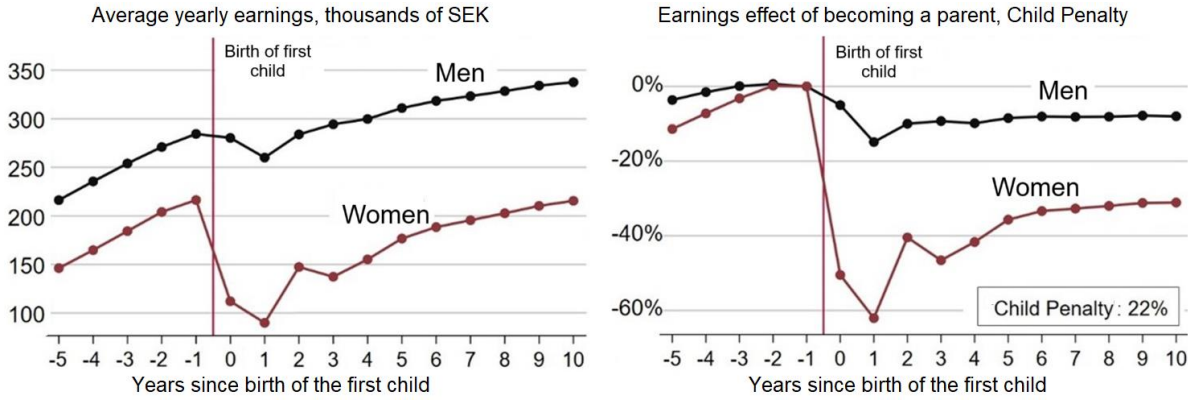


Figure 2C. Child penalty graph with raw means in the main survey experiment.

Notes: Graphs were created using Swedish administrative data for annual earnings and the replication files from Kleven et al. (2019a). Details in Appendix C.

Summary statistics

We clean the data by removing 297 respondents (2.4%) with missing responses for more than 80% of the survey questions and 121 respondents (1%) with short response times for the questions about the gender gap in earnings.[14] Respondents could skip any question in the survey, and all questions have less than 7% missing data (Table B3). The Citizen Panel's background data for individual traits also has little missing data: less than 1% for all variables in Table 1 except current economic activity (2.3%) and occupation (23.2%) (full list of percentages in Appendix Table B3).

Balance tests for observable traits and prior beliefs about gender inequality validate the randomization across the three graphs. We use binary OLS regressions to test for differences in means for the 25 traits in Table 1 (results presented in columns 4 and 5) and dummies for quartiles of the distributions of prior beliefs about the gender gap in earnings before and after parenthood (results in Appendix Table B4). The tests reveal small differences in means that rarely reach statistical significance at the 5% level. Only 4 of the 33 tests reach that level of significance when we compare respondents who saw Figures 2A and 2B, and 2 for the comparison of Figures 2A and 2C.

[14] These 121 respondents include 93 (0.7%) with a combined response time below 15 seconds for the first two questions about the gender gap in earnings, and 28 (0.2%) with a combined response time below 20 seconds for the second set of questions about that gap.

Table 1. Summary statistics and balance tests for individual traits.

	Child penalty graph	Gender gap graph	Child penalty graph with raw means	Diff. (2)-(1)	Diff. (3)-(1)
	(1)	(2)	(3)	(4)	(4)
Woman	0.47	0.46	0.46	-0.01	-0.01
Age < 40	0.12	0.11	0.12	-0.01	0.00
Age 40–59	0.50	0.49	0.48	0.00	-0.02
Age > 59	0.38	0.39	0.40	0.01	**0.02**
Tertiary education	0.82	0.83	0.81	0.01	-0.01
Current economic activity					
Working	0.69	0.69	0.69	-0.01	0.00
Not working	0.04	0.04	0.03	0.00	-0.01
Student	0.02	0.02	0.02	0.00	0.00
Retired	0.25	0.26	0.26	0.01	0.01
Relationship status					
Single without prior marriage	0.14	0.12	0.13	**-0.02**	-0.01
Divorced or widowed	0.10	0.11	0.11	0.00	0.01
In a relationship	0.76	0.77	0.76	0.01	0.00
Monthly pre-tax income, SEK					
< 36,999 (~4,000 USD)	0.35	0.33	0.33	-0.01	-0.02
37,000–44,999	0.17	0.17	0.19	0.00	0.02
45,000–64,999	0.32	0.30	0.31	-0.01	-0.01
> 65,000 (~7,000 USD)	0.17	0.19	0.18	**0.02**	0.01
Current occupation (SSYK 2012)					
Managers	0.11	0.12	0.12	0.00	0.01
Occupations requiring advanced higher education	0.65	0.64	0.62	-0.01	**-0.03**
Occupations requiring higher education qualifications or equivalent	0.12	0.12	0.12	0.01	0.00
Admin. and customer service clerks	0.03	0.03	0.04	0.00	0.01
Service, care and shop sales workers	0.05	0.04	0.05	0.00	0.01
Agricultural/horticultural/forestry/fishery workers	0.00	0.01	0.01	0.00	0.00
Building and manufacturing workers	0.02	0.02	0.02	0.00	0.01
Mechanical manufacturing and transport workers	0.02	0.02	0.02	0.00	0.00
Elementary occupations	0.01	0.01	0.01	0.00	0.00
Armed forces occupations	0.01	0.01	0.00	0.00	0.00
Non-missing data for answers given after the graph					
The gender gap in earnings before parenthood	0.98	0.97	0.97	-0.01	**-0.01**
The gender gap in earnings after parenthood	0.97	0.99	0.95	**-0.01**	**-0.02**
Importance of parenthood for gender inequality	0.96	0.95	0.95	-0.01	**-0.01**
Observations	3,912	3,969	3,831	7,743	7,881

Notes: Columns 1–3 show the shares of respondents by individual traits across treatment arms. Columns 4 and 5 report estimated coefficients from binary ordinary least squares (OLS) regressions with the individual trait as the dependent variable and a dummy for the gender gap graph (column 4) or the child penalty graph with raw means (column 5) as the independent variable. The *child penalty graph* is the reference condition for both comparisons. Data is restricted to non-missing observations, and the proportions of missing data are <1% for all traits except current economic activity (2%) and occupation (23%); see Appendix Table B3. Bold font indicates statistical significance at the 5% level. SSYK 2012 is Sweden's official occupation nomenclature and closely resembles ISCO-88.

The summary statistics in Table 1 show gender-balanced sample with an over-representation of older individuals and—by design—an over-representation of people with tertiary education (80%). Two-thirds of the respondents are employed, one-fourth are retired, and less than 5% are not working or students. Most are in a romantic relationship (75%) and about one in ten are single/never married (12%) or divorced/widowed (11%). Nearly half of the respondents have an above-median monthly income of over 45,000 SEK (the median was ~37,000 SEK in 2024). The sample has relatively large proportions of managers and occupations requiring advanced education, and small proportions of working-class occupations.

Response rates for the post-treatment questions about the gender gap in earnings only differ by 1–2 percentage points across the treatment arms (bottom rows of Table 1). These gaps are too small to drive our main treatment effects, which typically exceed 20 percentage points. Applying Lee's (2009) correction model shows that the largest change to the main results (in both relative and absolute terms) from applying the correction model is to reduce the estimated treatment effect from 8.2 to 6.9 percentage points.

Figure 3 displays answers about the gender gap in earnings before respondents saw the research graph. The vertical lines denote actual gap sizes from the microdata. Averaging over 5 years before parenthood gives an average covariate-adjusted earnings gap of 17%[15] and a raw gap of 29%, and 10 years after parenthood, these gap sizes are 27% and 36%.

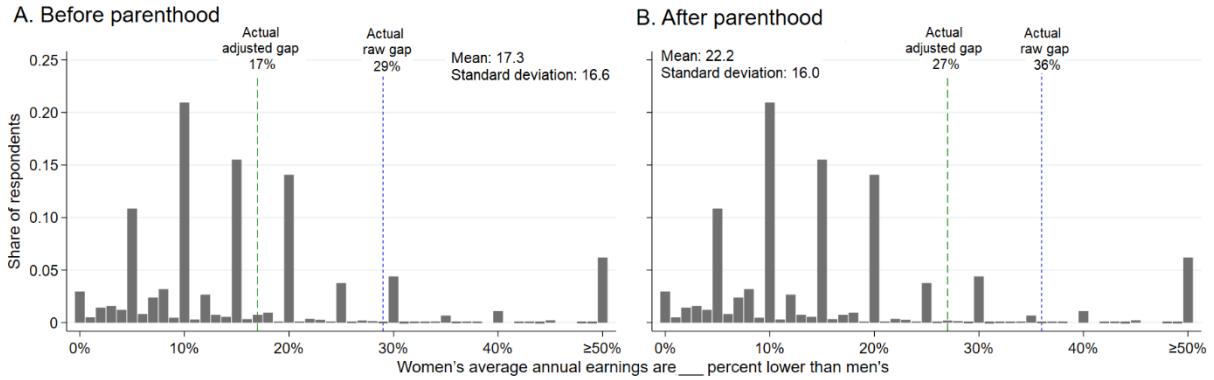


Figure 3. Prior beliefs about the gender gap in earnings.

Notes: The gender gap range is truncated at 50%. Vertical lines indicate actual gap sizes calculated in the microdata sample. The covariate-adjustment for age and year uses point estimates from the pooled event-study regression for the gender gap graph (Figure 2B) and the normalization to a percentage from Kleven et al. (2019a); details in Appendix C. N=11,712.

[15] We calculate the covariate-adjusted gap using the point estimates from the interaction model used to produce the right-side plot in the gender gap graph, i.e., taking the point estimates for the gender gap conditional on controls and normalizing it by dividing the monetary amount by men's counterfactual earnings (without children). If we instead normalize against *men's* average earnings, we obtain the same gap size before parenthood (17%) and a slightly larger size (32.0%) in year 10.

Respondents' beliefs about the gender gap in earnings varies across the traits in Table 1 (full results in Appendix Table B5). We observe larger perceived gender gaps among women, older respondents, and those with lower education or income. Ratings of the importance of parenthood for gender inequality vary less across individual traits, except for somewhat higher ratings among women and respondents with tertiary education. The 35% of respondents who self-reported previously hearing about child penalty research perceived slightly smaller gender gaps in earnings, and rated parenthood as slightly more important for gender inequality (both in bivariate regressions and conditional on the traits in Table 1, see Appendix Table B6).

Methods

We employ descriptive analysis and regressions to test for misunderstandings of the child penalty graph. The former uses histograms to compare responses about the gender gap in earnings before and after seeing the research graph. The latter estimate percentage-point differences in responses within and across individuals.

Our within-person analysis estimates changes in beliefs before vs. after seeing the graph. We reshape the data to have two observations per respondent (one pre- and one post-treatment) and regress a dummy for specific beliefs of interest (such as a 0–5% earnings gap before parenthood) on a dummy for the post-treatment observation, with standard errors clustered at the respondent level. The estimated coefficient quantifies the percentage-point change in beliefs generated by seeing the graph.

Comparing responses across graphs links changes in specific beliefs to the *child penalty graph*'s design elements by testing whether the changes disappear once we show the same research result with a graph that excludes the potentially misunderstood components. Our *gender gap graph* does not include gender-specific trend lines or a child penalty statistic, but instead makes the levels of gender inequality before and after parenthood explicit.

We formally compare the impact of seeing the *child penalty graph* and the *gender gap graph* by using OLS to estimate

$$DV_i = \alpha + \beta Child\_penalty\_graph_i + \boldsymbol{X}_i + \boldsymbol{Prior_i} + \epsilon_i \qquad (1)$$

where DV_i is a dummy for either the belief of interest (static post-treatment) or for switching to this belief after seeing the graph (a pre–post shift). The dummy for the $Child\_penalty\_graph_i$ takes a value of 1 for respondents who saw that graph and 0 for the *gender gap graph*. A positive

estimate of β implies a larger probability for the child penalty graph, conditional on controls for individual traits and prior beliefs about gender inequality. We include two vectors of controls, $\boldsymbol{X_i}$ (containing all traits in Table 1) and $\boldsymbol{Prior_i}$ (containing dummies for ventiles of prior beliefs)[16] and select among them using Belloni et al.'s (2014) post-double LASSO approach. We code missing values as zeros and add dummies for missing status to avoid losing observations; given the small amount of missing data, this pre-specified choice does not affect the results.

We test whether adding raw means to the child penalty graph mitigates the misunderstanding by repeating the analysis for Figures 2A and 2C. An analysis of approximately accurate beliefs in the supplementary material also uses the same sequence of methods. We analyze consequences of the misunderstanding by estimating Equation (1) and splitting the sample into respondents who are more vs. less likely to have misunderstood the *child penalty graph* (the Results section explains this analysis in more detail).

Misunderstanding the level of gender inequality *before* parenthood

We test whether respondents perceive the overlapping trend lines close to zero before parenthood in the child penalty graph as evidence of small or zero gender inequality. Because our graph has a small vertical distance between the trend lines in years 3–5 before parenthood (Figure 2A), we analyze the likelihood of stating a gender gap of 5% or less.

Comparing respondents' answers before and after being presented with the *child penalty graph* reveals a clear shift toward the 0–5% range (see Figure 4). The fraction of respondents stating such small or zero gender gaps in earnings increases by 29 percentage points from 0.18 to 0.47 (a 150% increase, $p < 0.01$; regression output in Table D1). Nearly one-third (32%) of the respondents initially state a larger gap but switch to the 0–5% range after seeing the graph. A particularly striking increase occurs at the value of 0. The share stating a gap of 0–1% increases by 19 percentage points from 0.03 to nearly one in four respondents (0.22, p-value < 0.01).

[16] Appendix Table B7 lists cutoff values and proportions of observations for these ventile brackets.

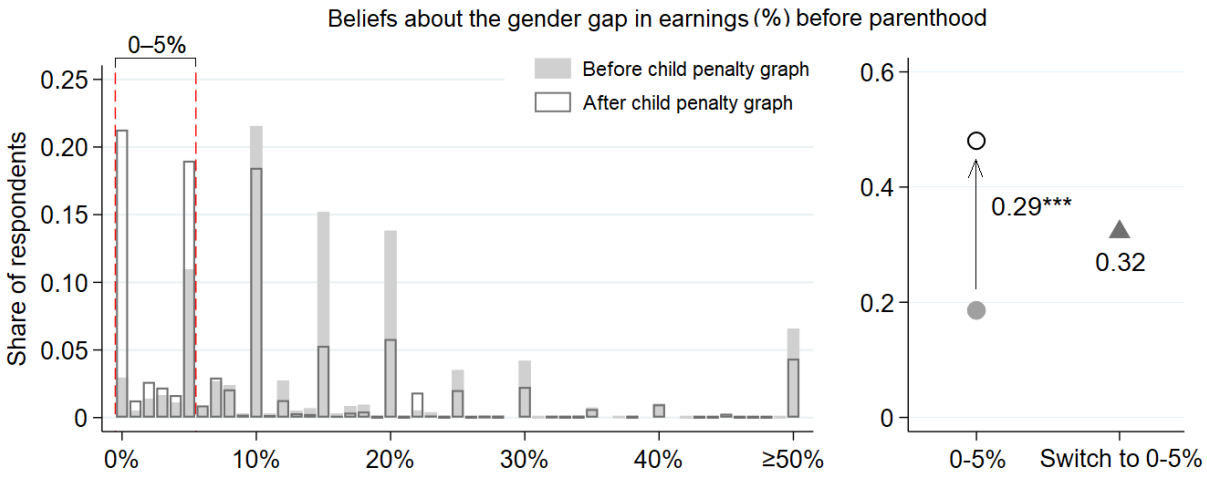


Figure 4. Shift in beliefs about gender inequality before parenthood.

Notes: Data for respondents who saw the child penalty graph in the survey experiment (N=3,786), excluding the 3% (N=126) who did not answer all questions about the gender gap in earnings. The gender gap range is truncated at 50%. Regression outputs for the increase in the share of beliefs in the 0–5% range appear in Appendix Table D1, Row 1, Columns 1 and 2. *** p<0.01, ** p<0.05, * p<0.1.

The *gender gap graph* did not generate any shift in beliefs toward the 0–5% range. Respondents who were assigned to receive this graph instead change their answers toward larger earnings gaps before parenthood, resulting in a 15-percentage-point drop for the 0–5% range (Figure 5). Estimating Equation (1) shows a 45-percentage-point difference in proportions relative to the *child penalty graph* (as shown by the estimate on the horizontal line in Figure 5, regression output in Appendix Table D2). These results link the observed shift for the *child penalty graph* to the pattern of overlapping gender-specific trend lines before parenthood.

Subsample analysis by education level and job title contradicts our expectation of larger misunderstandings among individuals who are "less familiar with event-study methods" (results in Appendix Table D3). We observe a 37-percentage-point shift to the 0–5% range among respondents with a PhD, compared to 29 percentage points in the full sample. After viewing the *child penalty graph*, 60% of PhD holders stated a gender gap of 5% or less. Smaller shifts occurred for respondents without tertiary education (20 percentage points from 0.13 to 0.33; N=690) or who have jobs related to research communication (25 percentage points from 0.16 to 0.41; N=248). PhD holders may be more confident in extracting information from statistics and figures, making them more likely to extract the wrong information from the *child penalty graph*.

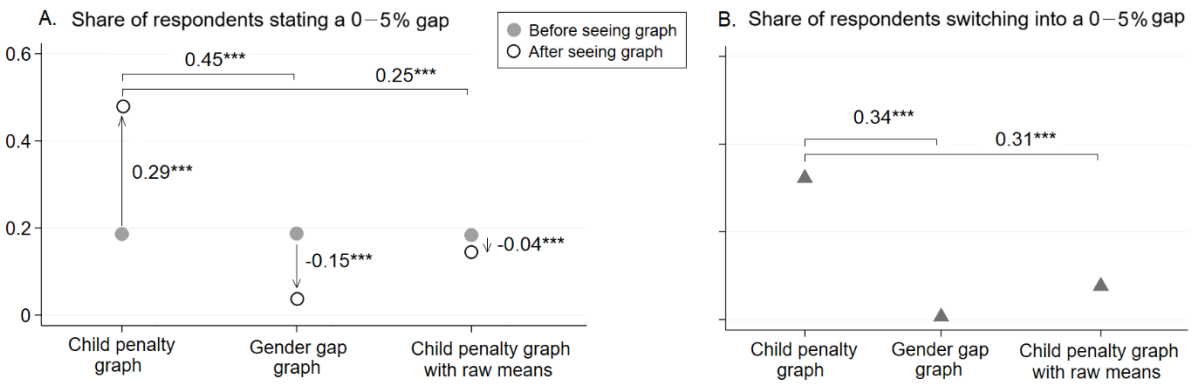


Figure 5. Beliefs about a 0–5% gender gap in earnings before parenthood.

Notes: Share of respondents stating a gender gap in earnings of 0–5% before and after seeing each graph (left) and the share that changed their belief into this range (right). Point estimates on the arrows from using OLS to regress a dummy for the 0–5% belief on a dummy for the post-treatment observation with standard errors clustered at the respondent level (detailed results in Table D1). Estimates on horizontal lines from using OLS to estimate Equation (1) to compare the share of respondents that held the belief post-treatment (left) and the shared that switched into the belief (right) across pairs of graphs (detailed results in Appendix Table D2).

The child penalty graph with raw means yields almost no change in the proportion of respondents with beliefs in the 0–5% range (a 4-percentage-point decrease; see Figure 5 and Appendix Table D1). Closer inspection reveals that an 11.5% outflow from this range occurs after seeing the graph, along with a 7.5% inflow.[17] While the subplot of raw means informed some respondents that the gender gap exceeds 5% before parenthood, a non-trivial fraction still appeared to be swayed into the 0–5% range by the subplot showing the *child penalty graph*

Misunderstanding the level of gender inequality *after* parenthood

We also test whether respondents misinterpret the vertical gap between the trend lines in the child penalty graph as the gender gap in the outcome variable. A second test examines if they interpret the child penalty statistic as the size of this gap. Distributions of answers before and after seeing the *child penalty graph* corroborate these predictions (see Figure 6). Answers shift into the 20–25% range, i.e., into the range corresponding to the approximates distance between the lines in our graph at event time $t = 10$. Within this range, we observe a particularly large increase for the 22% value of the child penalty statistic.

The shifts again involve large fractions of respondents. The share with beliefs in the 20–25% range increased by 26 percentage points from 0.29 to 0.55 after seeing the graph (p-value < 0.001, Appendix Table D1), and the share with beliefs at exactly 22% increased by 23

[17] An over-representation of individuals with PhDs in this inflow indicates that it stems from respondents who feel more confident focusing on the more complex child penalty subplot than the simpler subplot with raw means.

percentage points from 0.01 to 0.24 (p-value < 0.001). Nearly 40% of respondents shifted into the 20–25% belief range, and about 20% to the 22% value (triangles in right side of Figure 6).

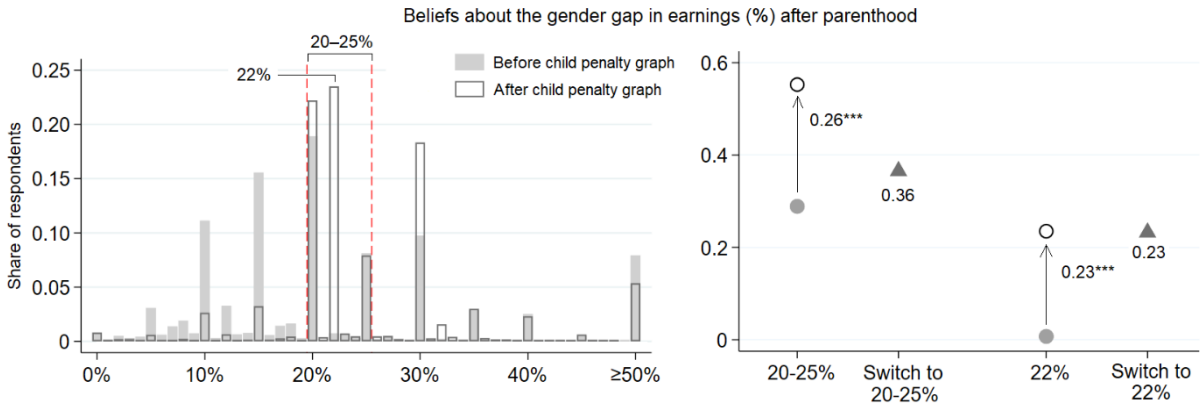


Figure 6. Shift in beliefs about gender inequality after parenthood.

Notes: see notes to Figure 4. Regression output in Table D1, Columns 1–2. *** p<0.01, ** p<0.05, * p<0.1.

Comparing answers across graphs supports our interpretation that they derive from the visual elements of the *child penalty graph*. Viewing the *gender gap graph* reduces rather than increases the proportion of answers in the 20–25% range, and does not budge an initially tiny proportion at exactly 22%. Estimating Equation (1) enforces these results. The share of respondents with beliefs in the 20–25% range is 48.5 percentage points larger for the *child penalty graph* (p-value < 0.001, Figure 7, Appendix Table D4), and the share with beliefs at exactly 22% is 24 percentage points larger (p < 0.001, Table D4).

The results presented in Figure 6 indicate another potential misunderstanding. The *child penalty graph* produces a notable increase in the share of responses for a 30% gender gap (from 0.10 to 0.28). Some respondents appear to interpret the position of the women's trendline, rather than the vertical distance between the two lines, as the gender gap in the outcome.[18]

Respondents with a PhD again shifted their beliefs more than others (not pre-registered; see Table D3), but the difference is smaller than in the analysis of inequality before parenthood. Adding the subplot of raw means to the child penalty graph reduces the shift in beliefs toward the 20–25% range by two-thirds, and the shift toward the 22% value by about one-third relative to the graph without raw means (not pre-registered; Figure 7C, Table D1).

[18] This misunderstanding is similar to *The Economist*'s (30 January 2024) description of Kleven et al. (2024) that "The authors define the 'motherhood penalty' as the average amount by which a woman's probability of being employed declines during the ten years after the birth of her first child."

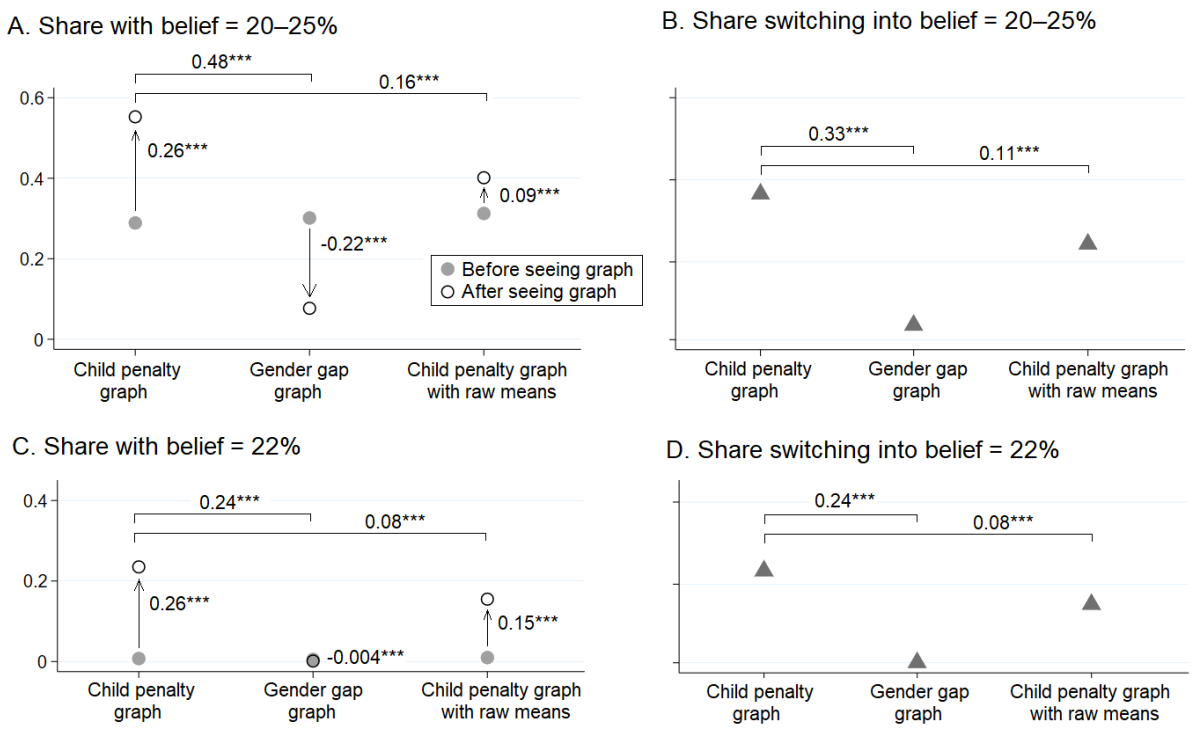


Figure 7. Beliefs about the gender gap in earnings after parenthood across the three graphs.

Notes: see notes to Figure 5. Descriptive statistics and regression outputs in Appendix Tables D1 and D4.

Supplementary evidence from a small survey of researchers

We recruited a convenience sample of mostly active academics to address weaknesses in the main survey. Rather than indirectly measuring the misunderstanding by asking for the size of the gender gap in earnings, we asked direct questions about "what we learn" from the graph, and free-text questions about its content. The survey also compared the prevalence of the misunderstanding between the simplified design from childpenaltyatlas.org and a more technical design typically used in research papers. We randomized respondents to receive either the main graph (Figure 2A) or a version mimicking the original design in Kleven et al. (2019a) (Appendix Figure F1 reproduces both). Compared to the simplified design, the technical design includes the Y-axis label "Earnings relative to t = -1" and a slightly adapted version of the original footnote.

We recruited 766 respondents using social media posts that described the survey as a study of how people "interpret research graphs." Table F1 lists all survey questions. The largest group of respondents was academic economists (41%, N=316), followed by academics from other disciplines (25%; N=193), and "a job requiring a PhD degree" (12%, N=92). Most respondents were Europeans (84%) and had seen the child penalty graph in passing (60%). Nearly one in ten had used the graph in their own research (9%). Distributions of respondents' jobs, continents

of residence, and prior familiarity with the child penalty graph were balanced between these two treatment groups (Appendix Table F2), while the proportions of missing data were not (Table F3).

Figure 8 displays the results of measuring the misunderstanding with direct questions about what the graph shows. About 45% agree that "we learn from the graph" that "in the years just before parenthood, the gender gap in earnings is 0–5%." Nearly 60% agreed with learning that "ten years after parenthood, women earn 20–25% less than men". [19] This approximate prevalence of the misunderstanding differs little between the more simplified and more technical graphs (dark and light-colored bars, respectively). The misunderstanding is less prevalent among academic economists, but this difference becomes small once we exclude respondents who have used the graph in their own research (Appendix Figure F2).

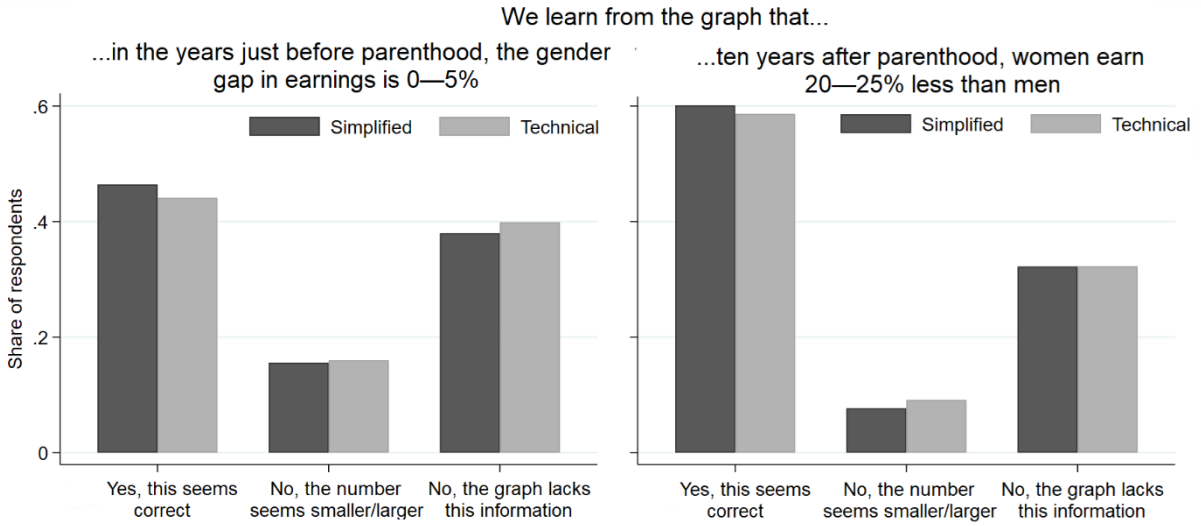


Figure 8. Direct query about the misunderstanding.

Notes: "Don't know" has been set to missing (N=19 for the left plot and N=15 for the right graph). N(Simplified, left)=321, N(Simplified, right)=336, N(Technical, left)=306, N(Technical, right)=307.

A second set of results comes from free-text questions. Table 2 lists these questions and presents the distributions of the answers across three hand-coded categories (coding details in Appendix Table F4, and examples of answers in Table F5). Proportions of missing data around 30% reduce the reliability of this analysis, and more missing responses for the technical graph affect the comparisons of graph designs (Table F3). It is also possible that answering the previous multiple-choice questions affected respondents' answers.

[19] A small proportion chose "No, the graph appears to show a smaller/larger value" (orange bars), which also corresponds to misunderstanding what the graph shows, but not to extracting the information associated with the specific misunderstanding analyzed in this paper.

The results indicate that 25–40% of respondents interpreted the graph's pattern before parenthood as a small or zero gender gap in earnings. This prevalence shrinks by about one-fourth among academic economists (Appendix Table F6). Approximately 15–25% misunderstood the child penalty statistic as the gender gap in earnings (similar for academic economists, but this group is more likely to give the correct answer to both questions, see Appendix Table F6). Differences between graph designs are again negligible.

Table 2. Distribution of free-text answers about the child penalty graph (%).

What does the graph tell us about women's and men's earnings in the years just before parenthood?					
Sample:	All	Atlas	Article	Atlas	Article
Small or zero gender gap in earnings	41	42	41	29	27
Correct answer	33	31	36	22	24
Other	25	28	23	19	15
Missing				30	35
Observations	515	266	249	380	383
How would you interpret the child penalty statistic in the bottom right corner of the graph?					
Sample:	All	Atlas	Article	Atlas	Article
Gender gap in earnings	22	20	25	14	16
Correct answer	31	28	36	21	22
Other	45	50	39	34	24
Missing				31	38
Observations	500	261	239	380	383

Notes: percentages of respondents in hand-coded categories of free-text answers.

Consequences of the misunderstanding

Perceived origins of economic gender inequality. Misunderstanding the child penalty graph may affect perceptions of the origins of gender inequality in the labor market. The perceived importance of parenthood might increase for people who misunderstand the graph as showing that having a child replaces roughly equal outcomes with a large economic disadvantage for women. We return to the main survey to test this hypothesis.

Respondents rated the importance of parenthood for gender inequality on a 5-point Likert scale from "very unimportant" to "very important" (mean=4.27, SD=0.85). Using this rating as the outcome variable in Equation (1) reveals no difference between the *child penalty graph* and the *gender gap graph* (0.01 scale steps, p-value>0.1; Appendix Table D5, column 1). A closer examining suggests an important caveat to this approach. Seeing the *gender gap graph* seems to have (also) raised the importance rating by making respondents perceive a larger earning

gaps after parenthood.[20] Many respondents who saw this graph (60%) switch to stating an earnings gap above 30% for parents (compared to just 18% for the child penalty graph), and these switchers assign a higher importance rating than non-switchers (0.14-scale-steps, p-value<0.001).

We analyze importance ratings among respondents who viewed the *child penalty graph* and split the sample by two proxy variables for the misunderstanding of interest. Respondents who switched to stating a 0–5% earnings gap before parenthood after seeing the graph have a 0.21-scale-steps (0.25 SD) higher rating than non-switchers. Respondents who describe the graph as "helpful" for answering the earnings gap questions (76%) have a 0.36-scale-steps higher rating (0.42 SD) than those who did not. Both differences remain large at 0.18 and 0.29 scale steps (p-values<0.001) when we estimate them in regressions including all controls for individual traits and initial answers about the gender gap in earnings (Appendix Table D6).

We compare importance ratings between the subsamples of interest and all respondents who saw the *gender gap graph*. These results show higher importance ratings for the two subgroups deemed as more likely to have misunderstood the *child penalty graph* (0.14 and 0.09 scale steps, p-values <0.01), and *lower* ratings for the two groups deemed less likely to have misunderstood it (-0.05 and -0.25 scale steps, columns 3 and 4 in Appendix Table D5). Overall, the results in this subsection provide some evidence that misunderstanding the child penalty graph increases the perceived importance of parenthood as a driver of gender inequality in the labor market.

Social media engagement. Viewers who misunderstood the *child penalty graph* might be more willing to engage with it on social media because they find its results more striking and novel. Our respondents rated their likelihood of liking, sharing, or cite-posting the graph on 5-point Likert scales from Very Likely to Very Unlikely. We take the average of these three ratings for each respondent. Among those who self-report using social media (82%), using this average as the dependent variable in Equation (1) shows a 0.09-scale-step *lower* willingness to engage with *the child penalty graph* compared to the *gender gap graph* (p-value <0.001; 0.08 SD).

Repeating the sample split for respondents who are more vs. less likely to have misunderstood the *child penalty graph* yields no coherent picture. Respondents who found the graph "helpful" were more willing to engage (perhaps unsurprising because people generally

[20] Due to this complexity, we do not pursue the planned analytical extensions for the importance ratings of discrimination and gender segregation outlined in the pre-analysis plan.

want to share helpful content), but switchers to a 0–5% earnings gap before parenthood were not (Appendix Table D7). These results provide no clear evidence that misunderstanding the *child penalty graph* propels social media amplification.

Accurate beliefs and graph design

A solution-oriented discussion of graph design might focus on how best to convey the child penalty while avoiding misunderstandings about gender inequality levels. According to the Child Penalty Atlas website, "the child penalty is always expressed as a percentage, ranging from 0 to 100. A child penalty of 20% in employment means that women's employment fell 20% behind men's employment after having children." [21] Our pre-registered analysis erroneously equated the 22% child penalty statistic in our earnings analysis to a 22-percentage-point expansion of the gender gap in earnings after parenthood.[22] This error invalidates our original plan to test for accurate communication about the size of the child penalty using the difference between respondents' answers about the gender gap in earnings before and after parenthood.

We compare how well the three graphs communicate accurate beliefs about (i) gender gap sizes before and after parenthood and (ii) the percentage-point increase in the gender gap after parenthood. Such accurate beliefs can apply to the *adjusted* gender gap in earnings (conditional on age and year fixed effects) as well as the *raw* gender gap in earnings. Our survey did not mention this distinction, and we failed to realize its importance until we had the data in hand. This section summarizes the results of applying the pre-specified analysis for the adjusted gap to both quantities, noting that large and precise treatment effects should limit concerns about multiple testing.

We define intervals of approximately accurate beliefs about gender gap sizes and analyze whether respondents move into these intervals after seeing each graph (details in Appendix E, Figure E1 and Tables E1–E4). The *child penalty graph* strongly reduces accuracy about the gender gap before parenthood, but improves accuracy for parents. The *gender gap graph* doubles accuracy before parenthood and quintuples it for parents. The *child penalty graph with raw means* also improves accuracy for both periods, but with the important caveat that

[21] https://childpenaltyatlas.org/ (accessed on December 18, 2025).
[22] Under a causal interpretation of the event-study regression results (i.e., assuming that the earnings gap between women and men would have evolved similarly in the absence of children), our child penalty statistic indicates that parenthood causes a gender gap in earnings ten years after parenthood that corresponds to 22% of what women would have earned if they did not have any children.

respondents improved their accuracy about the *raw* gender gap in earnings, which is a different and usually smaller quantity than the age- and year-adjusted gender gap typically studied in research on the child penalty.

We analyze accurate beliefs about how inequality changes after parenthood. For each respondent, we subtract their answer about the gap *before* parenthood from their answer about the gender gap *after* parenthood (for example, a respondent stating gaps of 10% and 20% receives a percentage-point value of 10). Respondents who saw the *child penalty graph* or *gender gap graph* both improved their understanding that gender inequality increases with parenthood (compared to no change or a reduction). Seeing the *gender gap graph* makes respondents gain more accurate beliefs about the size of this increase, while seeing the *child penalty graph* makes a substantial fraction of respondents overstate it. The *child penalty graph with raw means* generates some problematic results. The share of respondents that incorrectly perceived inequality as decreasing or remaining constant after parenthood increased modestly to 27% after seeing this graph. At the same time, the share that perceived overstated increases of 20 percentage points or more also grew from 4% to 18%.

We draw on the results in this section to briefly discuss graph design. The results indicate that the design of the *gender gap graph* is a potentially useful alternative. Despite removing the temporal benchmark at $t = -1$, this graph provides more accurate information about both levels and changes in inequality than the *child penalty graph*. Several caveats apply to a wholesale recommendation of the graph design, however. Adding the child penalty statistic might reintroduce the misunderstanding, and fully saturating the interaction model might make the presentation unwieldy. The single trend line also obscures potential comparisons of child penalties for women and men.

The results suggest that adding raw means to the child penalty graph would not be helpful. While doing so largely avoided generating the misperception that gender inequality is small or zero before parenthood (recall Figure 5), people gained more accurate perceptions of the size of the *raw gender gap* rather than the covariate-adjusted gap that is of interest in research on child penalties. The graph also created more misperceptions of overly small and overly large changes in inequality after parenthood. Marginal adjustments to the design are unlikely to overcome this problem, which appears to stem from some people drawing information only from the child penalty subplot, and others drawing information about inequality *before* parenthood from the raw means and *after* parenthood from the child penalty plot.

Finally, the results indicate that a more detailed label on the y-axis might improve comprehension of the *child penalty graph*. Other improvements might add more detailed explanations of (i) the reference year, (ii) the covariate adjustment, and (iii) the percentage normalization relative to the counterfactual outcome "in the absence of children." Adjustments to this graph design might help overcome not just the misunderstanding documented in this paper. Our supplementary survey highlighted other common misinterpretations. Substantial fractions of respondents appear to interpret the graph's y-axis as percentage differences relative to the reference year only, or to interpret the child penalty statistic as children's average impact on all parents regardless of gender.

Conclusions

Many people misunderstand the typical pattern depicted in the *child penalty graph*. They interpret the overlapping lines before parenthood as roughly equal outcomes for women and men, and the vertical distance between the lines after parenthood as the gender gap in the outcome. They also misinterpret the child penalty statistic as showing the gender gap in the outcome. We document this misunderstanding by asking survey respondents to state the gender gap in earnings before and after seeing a child penalty graph for that variable. Supplementary data from a convenience sample demonstrate its existence using alternative measurement strategies based on direct queries and free text questions.

The misunderstanding implies that people will strongly underestimate the large gender inequalities in most pre-parenthood labor market variables. Our results also indicate that misunderstanding the graph increases the perceived importance of parenthood as an explanation for gender inequality in the labor market. Future research might overcome methodological shortcomings in this analysis by measuring these perceptions both before and after viewing the graph.

Contrary to our expectation, respondents with PhDs reacted more strongly to the graph than others. Since our samples skewed toward high levels of education, our results may provide upper rather than lower bounds for how much seeing the child penalty graph affects beliefs about inequality levels in the population. Perhaps more importantly, larger reactions among PhDs raise concerns about academic audiences. Academics who misunderstand the graph might carry these misperceptions into their own research agendas, and reflect them back to audiences inside and outside of academia.

Researchers who use the graph might actively seek to prevent simplifications of their figures in media reproductions. Our comparison of alternative graph designs provides input on potential adjustments to improve comprehension. For example, our results caution against adding raw means to the graph. The event-study plot in our proposed *gender gap graph* is more promising. This graph efficiently communicates levels and changes in covariate-adjusted gender inequality by showing a single trend line for the size of the gender gap in the event window.

References

Adams-Prassl, A., Jensen, M.F. and Petrongolo, B., 2024. Birth Timing and Spacing: Implications for Parental Leave Dynamics and Child Penalties. IZA Discussion Papers (No. 17438)

Aizer, A. (2010). The gender wage gap and domestic violence. *American Economic Review*, *100*(4), 1847-1859.

Anderson, D. J., M. Binder, and K. Krause (2003). The motherhood wage penalty revisited: Experience, heterogeneity, work effort, and work-schedule flexibility. Industrial and Labor Relations Review 56(2), 273–294.

Becker, G. S. (1985). Human capital, effort, and the sexual division of labor. Journal of labor economics, 3(1, Part 2), S33-S58.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. Review of Economic Studies, 81(2), 608-650.

Bensnes, S., Huitfeldt, I., & Leuven, E. (2026). Reconciling estimates of the long-term earnings effect of fertility. ROCKWOOL Foundation Berlin DP 043/26.

Bertrand, M., Goldin, C., & Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. American economic journal: applied economics, 2(3), 228-255.

Lundborg, P., Plug, E. and Rasmussen, A.W., 2024. Is there really a child penalty in the long run? New evidence from IVF treatments. IZA Discussion Papers (No. 16959)

Casarico, A., Schuetz, J., & Uebelmesser, S. (2024). Gender inequality over the life cycle, information provision and policy preferences. CESifo Working Paper No. 10916.

Childpenaltytatlas.org (2025). https://childpenaltyatlas.org/methodology. Accessed on September 2

Cortés, P., Pan, J., Pilossoph, L., Reuben, E. and Zafar, B., 2023. Gender differences in job search and the earnings gap: Evidence from the field and lab. The Quarterly Journal of Economics, 138(4), pp.2069-2126.

Costa Dias, M., Joyce, R. and Parodi, F., 2020. The gender pay gap in the UK: children and experience in work. Oxford Review of Economic Policy, 36(4), pp.855-881.

Folbre, N. (2001). *The invisible heart: Economics and family values*. New York, NY: The New Press.

Gagnon, N., Bosmans, K., & Riedl, A. (2025). The Effect of gender discrimination on labor supply. *Journal of Political Economy*, *133*(3), 1047-1081.

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American economic review*, *90*(4), 715-741.

Goldin, Claudia (1990). *Understanding* the *Gender Gap*: An *Economic History* of *American Women*. *New York*, Oxford: *Oxford University Press*

Hsieh, C. T., Hurst, E., Jones, C. I., & Klenow, P. J. (2019). The allocation of talent and us economic growth. *Econometrica*, 87(5), 1439-1474.

Kleven, H., Landais, C., & Søgaard, J. E. (2019a). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, *11*(4), 181-209.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., & Zweimüller, J. (2019b). Child penalties across countries: Evidence and explanations. In *AEA Papers and Proceedings* (Vol. 109, pp. 122-126).

Kleven, H., Landais, C., & Leite-Mariante, G. (2024). The child penalty atlas. *Review of Economic Studies*, rdae104.

Korenman, S., & Neumark, D. (1992). Marriage, motherhood, and wages. The Journal of Human Resources, Vol. 27, No. 2. pp. 233-255

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies* 76: 1071–1102.

Lundborg, P., Plug, E., & Rasmussen, A. W. (2017). Can women have children and a career? IV evidence from IVF treatments. American Economic Review, 107(6), 1611-1637.

Lundborg, P., Plug, E., & Rasmussen, A. W. (2024). Is there really a child penalty in the long run? New evidence from IVF treatments (No. 16959). IZA Discussion Papers.

Miller, D. L. (2023). An introductory guide to event study models. *Journal of Economic Perspectives*, *37*(2), 203-230.

Reid Margaret G., (1934) *The Economics of Household Production*, New York: John Wiley & Sons Inc.

Settele, S. (2022). How do beliefs about the gender wage gap affect the demand for public policy?. *American Economic Journal: Economic Policy,* 14(2), 475-508.

The Economist (30 Jan, 2024). How motherhood hurts careers. < https://www.economist.com/interactive/graphic-detail/2024/01/30/how-motherhood-hurts-careers> Accessed on December 15, 2025.

Waldfogel, J. (1997). The effect of children on women's wages. American Sociological Review 62(2), 209–217.

Misunderstandings of the Child Penalty Graph

Olle Folke and Johanna Rickne

Appendix A. Usage and dissemination of the child penalty graph

Table A1 lists examples of research that studies Sweden and uses the child penalty graph. Figure A1 presents two examples of the graph as reproduced in the media. The left side shows a simplified version of Kleven et al.'s (2024) graph in *The Economist*, known for its excellence in data visualization. Its Instagram post including this graph on February 12, 2024 was widely circulated and received about 98,000 likes, making it one of the most popular posts of the year. On the right is the Nobel Foundation's post on X (formerly Twitter) on October 9, 2023, describing a research contribution rewarded by Claudia Goldin's Nobel Prize. This post had about 22,000 "likes" as of October 6, 2025. Figure A2 shows reproductions in other media outlets, and Figure A3 examples from government reports.

Table A1. Examples of research studies with child penalty graphs for Sweden.

1	Åslund, O., Karimi, A., & Sundberg, A. (2025). *Origin, norms, and the motherhood penalty* (No. 2025: 1). Working Paper Econstor.
2	Evertsson, M., Moberg, Y., & van der Vleuten, M. (2025). Stimulating (In) equality? The Earnings Penalty in Different-Sex and Female Same-Sex Couples Transitioning to Parenthood in Denmark, Finland, Norway, and Sweden. *American Journal of Sociology*, *130*(6), 1477-1525.
3	Kleven, H., Landais, C., Posch, J., Steinhauer, A., & Zweimüller, J. (2019, May). Child penalties across countries: Evidence and explanations. In *AEA Papers and Proceedings* (Vol. 109, pp. 122-126). 2014 Broadway, Suite 305, Nashville, TN 37203: American Economic Association.
4	Nylin, A. K., Musick, K., Billingsley, S., Duvander, A. Z., & Evertsson, M. (2021). Trends in women's relative earnings within couples across the transition to parenthood in Sweden, 1987–2007. *European Sociological Review*, *37*(3), 349-364.
5	Sundberg, A. (2024). *The child penalty in Sweden: evidence, trends, and child gender* (No. 2024: 12). IFAU-Institute for Evaluation of Labour Market and Education Policy.
6	van der Vleuten, M., & Moberg, Y. (2025). Cash-for-Care and the Cost of Parenthood: Evidence From Same-Sex and Adoptive Parents. *Journal of Marriage and Family*.
7	Van der Vleuten, M., Evertsson, M., & Moberg, Y. (2024). Joint utility or sub-optimal outcomes? Household income development of same-sex and different-sex couples transitioning to parenthood in Denmark, Finland, Norway, and Sweden. *Journal of Family Issues*, *45*(8), 2049-2076.

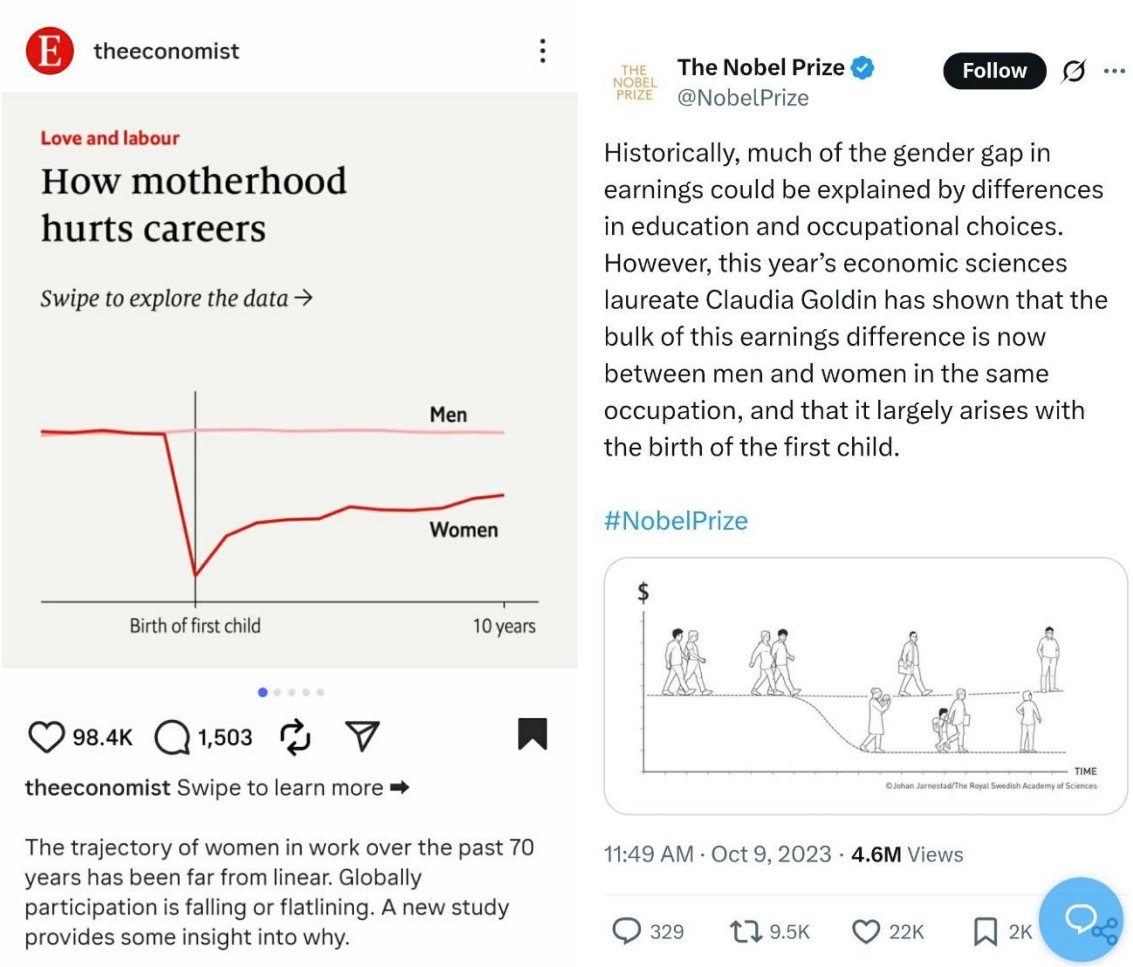


Figure A1. Examples of the child penalty figure circulated widely on social media.

Notes: The left panel shows an Instagram post from February 12, 2024, by *The Economist*, featuring Kleven et al. (2024), the Child Penalty Atlas. The right panel shows a post on X (formerly Twitter) from October 9, 2023 by the Nobel Foundation regarding Claudia Goldin's Nobel Prize. Both posts were retrieved on October 6, 2025.

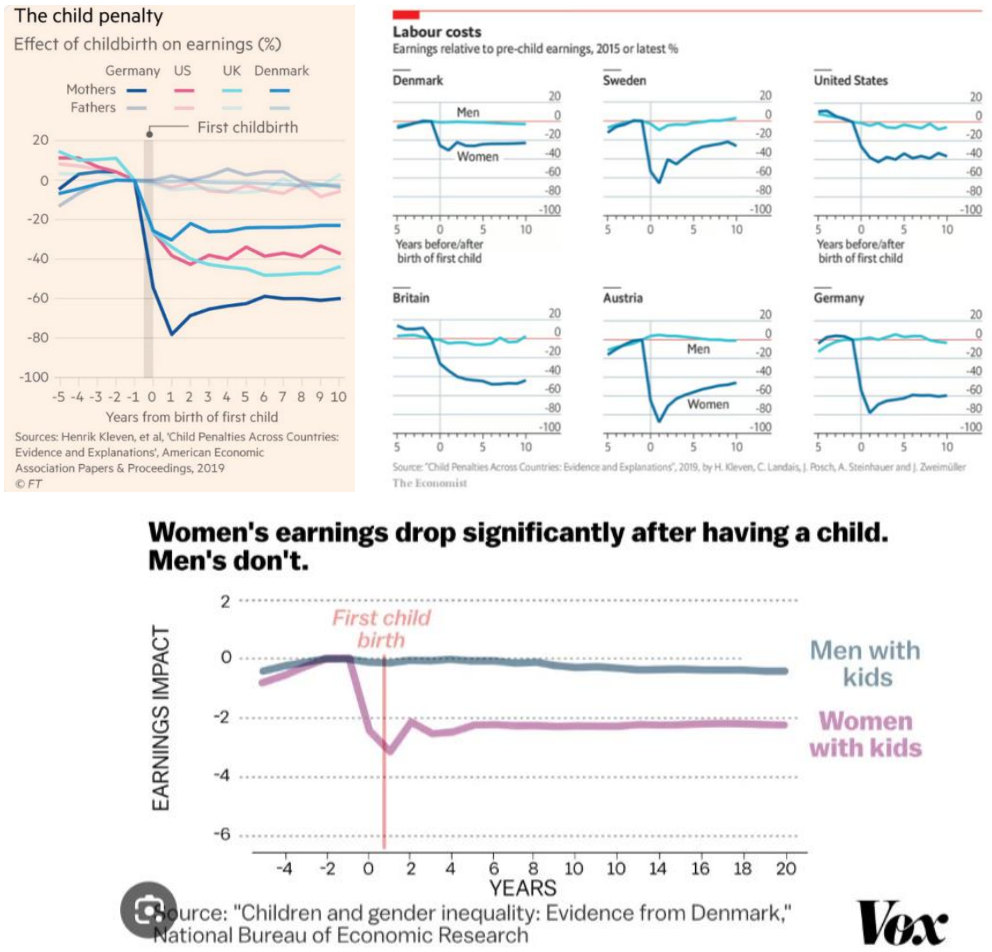


Figure A2. Additional examples of the child penalty graph in the media.

Notes: The top left graph comes from the *Financial Times* "Minouche Shafik: 'I don't have to be the smartest person in the room'" (April 9, 2021, retrieved on December 15, 2025). The upper right graph comes from *The Economist*, "How big is the wage penalty for mothers?" (January 28, 2019, retrieved on December 15, 2025). The bottom graph comes from the *Vox* column "A stunning chart shows the true cause of the gender wage gap" (February 19, 2018, retrieved on December 15, 2025).

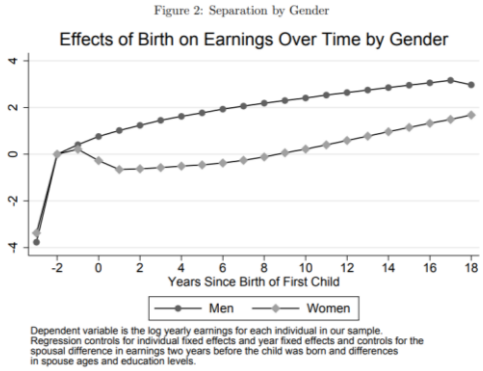


Figure A3a. Examples of child penalty graphs from government reports.

Source: Y. Chung, B. Downs, D. H. Sandler, and R. Sienkiewicz (2017) *The parental gender earnings gap in the United States* (No. 17-68). Technical report from the Census Bureau's Center for Economic Studies.

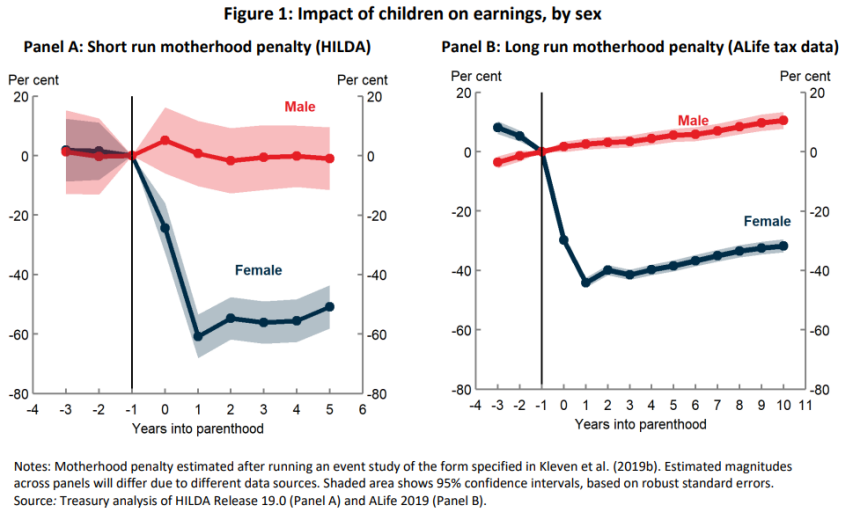
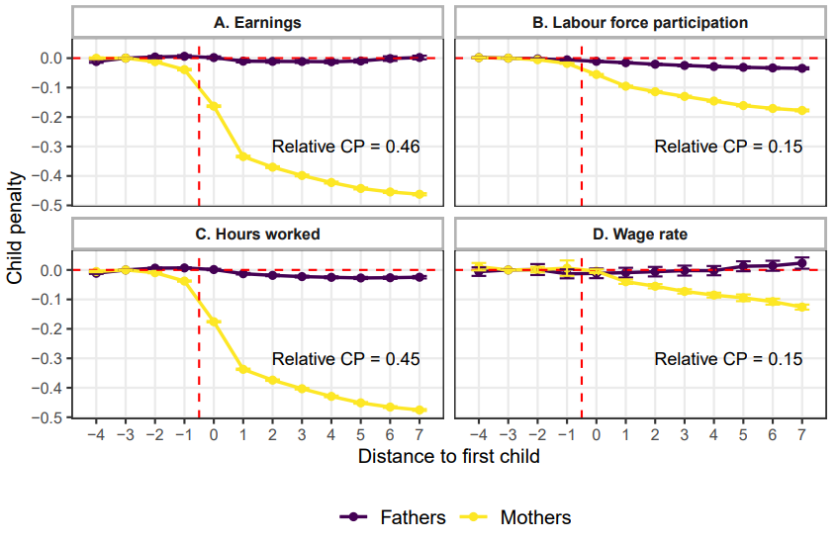


Figure A3b. Examples of child penalty graphs from government reports.

Source: Elif Bahar, Natasha Bradshaw, Nathan Deutscher and Maxine Montaigne (2022) "Children and the gender earnings gap," Treasury Round Up [Australian Treasury]. The same figure appears as Figure 3b in the Australian Productivity Commission's 2024 report "A path to universal early childhood education and care."

Figure 2: Estimated child penalty by distance to first child birth

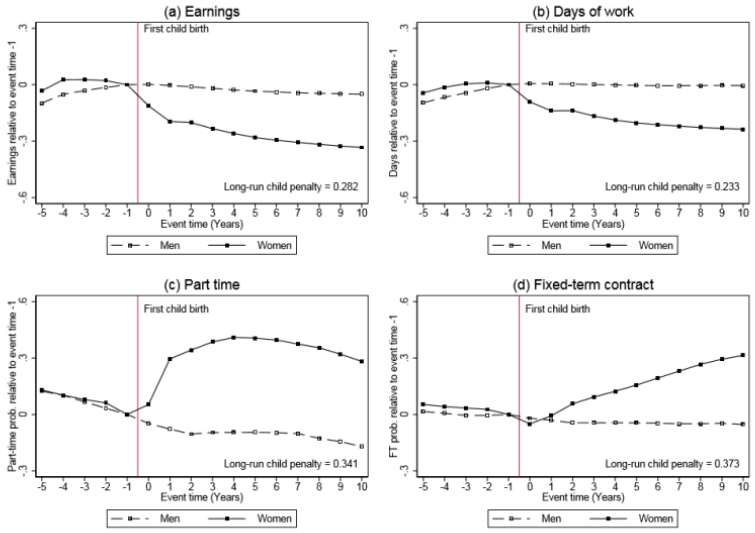


NOTE: This Figure presents the estimated child penalty for each event-time, as defined in equation (2), for mothers and fathers and for different labour market outcomes. The relative child penalty (cf. equation 3) in t=7 is added to each plot.

Figure A3c. Examples of child penalty graphs from government reports.

Source: S. Rabaté and S. Rellstab (2021) *The child penalty in the Netherlands and its determinants* (No. 424). Den Haag: CPB Netherlands Bureau for Economic Policy Analysis.

Chart 1
IMPACTS OF FIRST CHILD



SOURCE: Authors' work from Continuous Sample of Working Histories (MCVL).
NOTES: The figure shows event time coefficients estimated from equation (2), absent children for men and women separately and for different outcomes. Each graph also reports a "long-run child penalty"—the percentage by which women are falling behind men due to children, calculated from equation (3) at event time t = 10. All of these statistics are estimated on a balanced sample of parents who have their first child between 1994 and 2009 and who are observed in the data during the entire period between five years before and ten years after child birth. The effects on earnings and days of work are estimated unconditional on employment status, while the effects on part time and fixed-term contract are estimated conditional on working.

Figure A3d. Examples of child penalty graphs from government reports.

Source: A. De Quinto, L. Hospido, and C. Sanz (2020) *The child penalty in Spain*. Banco de Espana Occasional Paper, (2017).

Appendix B. Main survey and summary statistics

Table B1. English translation of survey screens (A–D) and survey questions (1–9) with cursive and bold font in original (Swedish in Table B2).

A	This is a study in collaboration with researchers from Stockholm University. The study concerns parenthood and gender inequality. The purpose is to investigate how people interpret the content of graphs presenting research results on this topic. If you have any questions about the study, you are welcome to contact the lead researcher Johanna Rickne (johanna.rickne@sofi.su.se), Professor of Economics at Stockholm University. For this study, the same conditions and rights apply as for other studies by the Citizen Panel. These are described in our Terms of Participation and Privacy Policy [HYPERLINKED].[1] To participate in the study, mark "I consent to participate in the study" below and then click "continue" at the bottom right. If you do not want to participate, click “continue” to skip the survey.
B	We will now ask you a few questions about economic equality between women and men in Sweden. Now and then, there's talk of a gap in earnings between women and men, with women typically earning less through work or entrepreneurship over the course of a year.
Q1	Think of adults in Sweden who have not yet had children. If you were to guess, by what percentage do you think women’s annual earnings are lower than men’s? *Answer in numbers in the box below* Among people who are not yet parents, women’s average annual earnings are ___ percent lower than those of men.
Q2	Now think instead about adults in Sweden who have become parents and whose oldest child is about 10 years old. If you were to guess, by what percentage do you think women’s annual earnings are lower than men’s? *Answer in numbers in the box below* Among people with children around 10 years old, women’s average annual earnings are ___ percent lower than those of men.
C	You will now see graphs showing research results on the earnings gap between women and men. The purpose of this research is to examine how having children affects economic gender equality. More precisely, the research examines whether women incur what is referred to as a 'child penalty', i.e., a worsening of their incomes relative to men upon entering parenthood. **After viewing the graphs, we will again ask the questions about a gender gap in earnings.**
Q3	Have you heard about the research on the 'Child Penalty,' which looks at how becoming a parent impacts earnings disparities between women and men? Response categories: Yes, No, Don’t know
Q4	Think of adults in Sweden who have not yet had children. If you were to guess, by what percentage do you think women’s annual earnings are lower than men’s? *If you don’t think the graphs provide useful information, you may enter the same answer as before. Answer in numbers in the box below.* Among people who are not yet parents, women’s average annual earnings are ___ percent lower than those of men.

[1] Deltagarvillkor och integritetspolicy för Medborgarpanelen vid Göteborgs universitet; https://samgu.eu.qualtrics.com/ControlPanel/File.php?F=F_ZKd9AKKLG2Jp8F0

Q5	Now think instead about adults in Sweden who have become parents and whose oldest child is about 10 years old. If you were to guess, by what percentage do you think women's annual earnings are lower than men's? *If you don't think the graphs provide useful information, you may enter the same answer as before. Answer in numbers in the box below.* Among people with children around 10 years old, women's average annual earnings are ___ percent lower than those of men.
Q6	How certain or uncertain are you about the answers you gave regarding the earnings gap between women and men <u>after</u> viewing the graph? Response categories: Very certain; Quite certain; Neither certain nor uncertain; Quite uncertain; Very uncertain
Q7	How helpful did you find the graphs in answering the questions about the difference in earnings between women and men? Response categories: Very helpful, Quite helpful, Not particularly helpful, Not helpful at all
Q8	In public debate about the economic status of women and men, certain factors are often highlighted as explanations for women's lower earnings. How important or unimportant do you consider the following explanations for economic differences between women and men? 8.1: Gender discrimination 8.2: Parenthood 8.3: Occupation and/or workplace choices Response categories: Very important explanation, Quite important explanation, Neither important nor unimportant, Quite unimportant explanation, Very unimportant explanation.
Q9	Think back to the graph you saw earlier and imagine seeing this graph on social media (e.g., Facebook, X, Instagram, or TikTok). How likely to you think you would be to engage with that post in the following ways? 9.1: "Like" the post 9.2: Share the post on your profile 9.3: Quote-post the post (share it with your own comment) Response categories: Very likely, Quite likely, Neither likely nor unlikely, Quite unlikely Very unlikely, Do not use social media
D	Here are a few facts about differences in earnings between women and men related to parenthood in Sweden. Women's average earnings are about 15–20% lower than men's in the years leading up to parenthood, and about 30% lower by the time the first child reaches age 10. So, while the gap grows after parenthood, it is already sizeable before people become parents. The questions you answered on this topic explore how effectively graphs with different designs present this information to viewers. The graphs you saw were based on data from Statistics Sweden and show results for all adults with a Swedish personal identity number who had their first child between 2002 and 2012. The graphs were produced by Olle Folke, Professor of Political Science at Uppsala University. If you have any questions about the study, you are welcome to contact the principal investigator Johanna (johanna.rickne@sofi.su.se), Professor of Economics at Stockholm University. If you have comments about the study, you are welcome to leave them in the text box below:

Table B2. Original Swedish survey screens (A–D) and survey questions (1–9).

A	Nu följer en studie i samarbete med forskare från nationalekonomiska institutionen vid Stockholms universitet. Studien handlar om föräldraskap och ojämställdhet. Syftet med studien är att undersöka hur olika personer tolkar innehåll i grafer för forskningsresultat. Om du har några frågor om studien är du välkommen att kontakta huvudansvarig forskare Johanna Rickne (johanna.rickne@sofi.su.se), professor i nationalekonomi vid Stockholms universitet. För denna studie gäller i övrigt samma villkor och rättigheter för dig som i andra studier från Medborgarpanelen, dessa beskrivs i våra Deltagarvillkor och integritetspolicy. För att delta i studien, klicka på "Jag godkänner att delta i studien" nedan samt klicka på "gå vidare" nere till höger. Om du inte vill delta i studien kan du istället gå vidare till nästa sida.
B	Nu följer några frågor om ekonomisk jämställdhet mellan kvinnor och män i Sverige. Ibland talas det om ett inkomstgap mellan kvinnor och män, det vill säga att kvinnor har lägre samlade inkomster från arbete eller företagande under ett år.
Q1	Tänk dig vuxna personer i Sverige som ännu inte har blivit föräldrar. Om du skulle gissa, hur många procent lägre årsinkomst tror du att kvinnor har i jämförelse med män? *Svara i siffror i rutan nedan* Bland personer som ännu inte har blivit föräldrar har kvinnor ☐ procent lägre genomsnittlig årsinkomst än män.
Q2	Tänk nu i stället på vuxna personer i Sverige som har fått barn och vars barn är i tioårsåldern. Om du skulle gissa, hur många procent lägre årsinkomst tror du att kvinnor har i jämförelse med män i den gruppen? *Svara i siffror i rutan nedan* Bland personer med barn i tioårsåldern har kvinnor ☐ procent lägre genomsnittlig årsinkomst än män.
C	Du kommer nu att få se grafer som visar forskningsresultat när det gäller inkomstsgapet mellan kvinnor och män. Syftet med forskningen är att undersöka hur den ekonomiska jämställdheten påverkas av att bli förälder. Mer specifikt undersöks om kvinnor drabbas av ett så kallat "child penalty", det vill säga försämrade inkomster jämfört med män i samband med föräldraskapet. **Efter att du har sett graferna kommer vi återigen att ställa frågorna om inkomstgapet.**
Q3	Har du tidigare hört talas om forskningen om ”child penalty” som handlar om hur föräldraskap påverkar inkomstskillnaderna mellan kvinnor och män? Svarsalternativ: Ja, Nej, Vet ej
Q4	Tänk dig vuxna personer i Sverige som ännu inte har blivit föräldrar. Om du skulle gissa, hur många procent lägre årsinkomst tror du att kvinnor har i jämförelse med män? *Om du inte anser att graferna ger hjälpsam information kan du ange samma svar som förut. Svara i siffror i rutan nedan.* Bland personer som ännu inte har blivit föräldrar har kvinnor ☐ procent lägre genomsnittlig årsinkomst än män.
Q5	Tänk nu i stället på vuxna personer i Sverige som har fått barn och vars barn är i tioårsåldern. Om du skulle gissa, hur många procent lägre årsinkomst tror du att kvinnor har i jämförelse med män i den gruppen? *Om du inte anser att graferna ger hjälpsam information kan du ange samma svar som förut. Svara i siffror i rutan nedan.*

	Bland personer med barn i tioårsåldern har kvinnor ☐ procent lägre genomsnittlig årsinkomst än män.
Q6	Hur säker eller osäker är du på dina svar som du angav efter du sett graferna angående inkomstgapen mellan kvinnor och män? Svarsalternativ: Mycket säker, Ganska säker, Varken säker eller osäker, Ganska osäker, Mycket osäker
Q7	Hur hjälpsamma tyckte du att graferna var för att besvara frågorna om inkomstgapen mellan kvinnor och män? Svarsalternativ: Mycket hjälpsamma, Ganska hjälpsamma, Inte särskilt hjälpsamma, Inte alls hjälpsamma
Q8	I debatten om kvinnors och mäns ekonomiska ställning brukar man diskutera olika faktorer som kan bidra till ekonomiska skillnader. Hur viktiga eller oviktiga anser du att följande orsaker är till ekonomiska skillnader mellan kvinnor och män? 8.1: Könsdiskriminering 8.2: Föräldraskap 8.3: Människors val av yrke och/eller arbetsplats Svarsalternativ: Mycket viktig orsak, Ganska viktig orsak, Varken viktig eller oviktig orsak, Ganska oviktig orsak, Mycket oviktig orsak
Q9	Tänk tillbaka på graferna du såg tidigare och föreställ dig att du såg denna information på sociala medier (exemeplvis Facebook, X, Instagram eller TikTok). Enligt din bedömning, hur sannolikt skulle det vara att du interagerade med inlägget på följande sätt? 9.1: ”Gillat” inlägget 9.2: Delat inlägget på din egen profil 9.3: Citatpostat inlägget (delat tillsammans med egen kommentar) Svarsalternativ: Mycket sannolikt, Ganska sannolikt, Varken sannolikt eller osannolikt, Ganska osannolikt, Mycket osannolikt, Använder inte sociala medier
D	Nu följer några fakta om skillnaden i inkomst i samband med föräldraskap i Sverige. Kvinnors inkomst är i genomsnitt 15—20% lägre än mäns under åren innan de blir föräldrar och ungefär 30% lägre när det första barnet hunnit bli tio år. Gapet är alltså större efter föräldraskapet än innan, men dess storlek är betydande redan innan föräldraskapet inträffar. De frågor du besvarat om det här ämnet undersöker hur framgångsrikt grafer med olika utseende redovisar dessa fakta för en läsare. Graferna baseras på data från statistiska centralbyrån och redovisar resultat för alla vuxna personer med svenskt personnummer som fick sitt första barn mellan 2002 och 2012. Graferna har framtagits av Olle Folke, professor i statskunskap vid Uppsala universitet. Om du har några frågor om studien är du välkommen att kontakta huvudansvarig forskare Johanna Rickne (johanna.rickne@sofi.su.se), professor i nationalekonomi vid Stockholms universitet. Om du har några kommentarer kring studien är du välkommen att lämna dem här:

Table B3. Non-response rates for each question and by covariates.

		Gender (0% missing)		Age (0.1% missing)			Education (0% missing)	
	All	Men	Women	<40	40—59	>=60	No university degree	University degree
Survey question (text in Table B1)								
Q1	0.5%	0.5%	0.5%	0.1%	0.2%	1.0%	0.7%	0.4%
Q2	0.6%	0.6%	0.6%	0.1%	0.3%	1.2%	0.9%	0.6%
Q3	0.1%	0.1%	0.1%	0.0%	0.1%	0.2%	0.0%	0.1%
Q4	4.8%	4.8%	4.7%	3.9%	4.9%	4.9%	5.3%	4.6%
Q5	5.8%	6.0%	5.6%	4.4%	5.9%	6.2%	7.5%	5.5%
Q6	4.7%	4.7%	4.7%	4.0%	5.2%	4.3%	5.1%	4.6%
Q7	5.1%	5.1%	5.2%	4.3%	5.6%	4.8%	5.7%	5.0%
Q8.1	5.6%	5.7%	5.5%	4.8%	5.8%	5.7%	6.6%	5.4%
Q8.2	5.5%	5.6%	5.3%	4.8%	5.8%	5.3%	6.4%	5.3%
Q8.3	5.6%	5.7%	5.5%	4.8%	5.7%	5.7%	6.5%	5.4%
Q9.1	6.0%	5.9%	6.0%	5.2%	6.1%	5.9%	7.3%	5.7%
Q9.2	6.4%	6.4%	6.5%	5.1%	6.3%	7.1%	8.0%	6.1%
Q9.3	6.6%	6.6%	6.7%	5.3%	6.2%	7.5%	8.2%	6.3%
	Bracket of monthly income (SEK) (0.17% missing)				Current economic activity (2.3% missing)			
	< 36 999	37 000—44 999	45 000—64 999	> 65 000	Working	Un-employed	Student	Retired
Q1	0.7%	0.4%	0.2%	0.5%	0.2%	0.5%	0.0%	1.4%
Q2	1.1%	0.3%	0.4%	0.6%	0.3%	0.8%	0.5%	1.6%
Q3	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.2%
Q4	5.1%	4.9%	4.2%	4.8%	4.6%	7.1%	2.4%	4.9%
Q5	6.5%	6.1%	5.0%	5.8%	5.5%	8.6%	3.4%	6.4%
Q6	4.8%	5.1%	4.3%	4.8%	4.9%	6.3%	3.4%	4.1%
Q7	5.3%	5.3%	4.6%	5.5%	5.2%	6.6%	3.9%	4.8%
Q8.1	6.1%	6.1%	4.9%	5.6%	5.6%	6.6%	4.9%	5.9%
Q8.2	5.9%	5.9%	4.6%	5.6%	5.5%	6.1%	4.9%	5.4%
Q8.3	6.0%	6.0%	4.9%	5.6%	5.6%	6.3%	4.4%	5.8%
Q9.1	6.3%	6.2%	5.3%	6.1%	6.0%	7.1%	3.9%	5.9%
Q9.2	7.4%	6.5%	5.5%	6.2%	6.2%	7.1%	3.9%	7.3%
Q9.3	7.7%	6.6%	5.5%	6.4%	6.2%	7.1%	3.9%	8.2%

Notes: The table shows proportions of missing values (non-response rates) across survey questions and split by respondent traits.

Table B4. Summary statistics and balance tests for prior beliefs.

	Treatment arm			Balance tests	
	Child penalty graph (Figure 2A)	Gender gap graph (Figure 2B)	Child penalty graph with raw means (Figure 2C)	Difference (2)-(1)	Difference (3)-(1)
Prior beliefs about the gender gap in earnings before parenthood					
Quartile 1	0.247	0.255	0.009	0.261	0.014
Quartile 2	0.252	0.245	-0.008	0.244	-0.009
Quartile 3	0.322	0.325	0.003	0.323	0.001
Quartile 4	0.179	0.175	-0.004	0.173	-0.006
Prior beliefs about the gender gap in earnings after parenthood					
Quartile 1	0.246	0.268	0.022	0.254	0.008
Quartile 2	0.402	0.398	-0.004	0.400	-0.002
Quartile 3	0.101	0.103	0.002	0.114	0.013
Quartile 4	0.251	0.231	**-0.020**	0.232	-0.019
Observations	3,912	3,831	3,969	7,743	7,881

Notes: Columns 1–3 show the shares of observations across quartiles in the three treatment arms. Columns 4 and 5 report estimated coefficients from binary OLS regressions with the quartile dummy as the dependent variable and a dummy for the gender gap graph (column 4) and the child penalty graph with raw means (column 5) as the independent variable. The standard child penalty graph (Figure 2A) is the reference condition for both comparisons. Bold font indicates statistical significance at the 5% level for the estimated coefficient.

Table B5. Correlations between individual traits and prior beliefs (columns 1 and 2) and importance ratings of parenthood (column 3).

Dependent variable:	Gender gap in earnings (%)		Importance rating of parenthood for economic gender inequality
	before parenthood	after parenthood	
Standard deviation of dependent variable:	16.6	15.0	0.85
	(1)	(2)	(3)
(Reference: man)			
Woman	3.07***	4.38***	0.25***
	(0.31)	(0.29)	(0.016)
(Reference: Age < 40)			
Age 40–59	1.89***	0.77	-0.13***
	(0.49)	(0.48)	(0.03)
Age > 59	5.41***	3.71***	-0.07***
	(0.51)	(0.49)	(0.03)
(Reference: No tertiary education)			
Tertiary education	-3.66***	-1.69***	0.17***
	(0.40)	(0.39)	(0.02)
(Reference: Currently working)			
Not working	4.39***	3.80***	0.05
	(0.85)	(0.82)	(0.04)
Student	1.07	2.19*	0.15**
	(1.16)	(1.13)	(0.06)
Retired	4.14***	3.02***	0.03*
	(0.36)	(0.35)	(0.02)
(Reference: Single, never married)			
Divorced or widowed	2.81***	1.95***	0.13***
	-0.63	(0.61)	(0.03)
In a relationship	0.54	0.53	0.10***
	-0.46	(0.45)	(0.02)
(Reference: income < 36,999 SEK/Month)			
37,000–44,999	-3.53***	-2.94***	0.02
	(0.45)	(0.43)	(0.02)
45,000–64,999	-4.36***	-3.62***	-0.02
	(0.381)	(0.37)	(0.02)
> 65,000	-4.21***	-3.65***	-0.09***
	(0.45)	(0.43)	(0.02)
Observations	11,654	11,053	10,923

Notes: see next page.

Table B5 (continued). Correlations between individual traits and prior beliefs (columns 1 and 2) and importance ratings of parenthood (column 3).

	Gender gap in earnings (%)		Importance rating of parenthood for economic gender inequality
Dependent variable:	before parenthood	after parenthood	
Standard deviation of dependent variable:	16.6	15.0	0.85
	(1)	(2)	(3)
(Reference category: managers)			
Occupations requiring higher education qualifications or equivalent	3.12	3.95	0.06
	(2.59)	(2.50)	(0.13)
Admin. and customer service clerks	4.83*	4.48*	-0.12
	(2.63)	(2.54)	(0.14)
Service, care and shop sales workers	7.91***	7.37***	0.02
	(2.77)	(2.67)	(0.14)
Agricultural/horticultural/forestry/fishery workers	7.95***	7.44***	-0.01
	(2.70)	(2.61)	(0.14)
Building and manufacturing workers	4.33	4.27	-0.07
	(3.65)	(3.52)	(0.19)
Mechanical manufacturing and transport workers	6.82**	4.84*	-0.19
	(2.87)	(2.77)	(0.15)
Elementary occupations	3.34	3.32	-0.23
	(2.94)	(2.83)	(0.15)
Armed forces	7.09**	7.58**	-0.14
	(3.55)	(3.43)	(0.19)
Observations	11,654	11,053	10,923

Notes: estimates in each section come from regressing the dependent variable in the column heading on the dummies for each demographic or socioeconomic trait. Missing value indicators were used in each regression. Proportions of missing data in Table B3. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table B6. Comparison of respondents who had heard about child penalty research or not.

	Gender gap in earnings (%)				Importance rating of parenthood for economic gender inequality	
	Before parenthood		After parenthood			
	(1)	(2)	(3)	(4)	(5)	(6)
Heard about the child penalty research=1	-2.322***	-0.978***	-0.861***	0.0364	0.127***	0.110***
Reference: did not hear	(0.329)	(0.340)	(0.318)	(0.328)	(0.0174)	(0.0179)
Don't know=1	-0.923	-1.079*	-0.0968	-0.439	-0.0154	-0.0358
	(0.610)	(0.609)	(0.590)	(0.587)	(0.0323)	(0.0321)
Demographic and socioeconomic traits	No	Yes	No	Yes	No	Yes
Observations	11,654	11,449	11,638	11,431	11,053	10,923
SD of the dependent variable	16.6		15.0		0.85	

Notes: OLS regression estimates. Controls are dummies for all categories of demographic and socioeconomic traits in Table 1. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table B7. Ventiles of pre-treatment beliefs about the gender gap in earnings in percent.

Beliefs about non-parents				Beliefs about parents			
Ventile	Lower	Upper	Share	Ventile	Lower	Upper	Share
1	0	3	6.4%	1	0	5	5.4%
2	3.5	5	12.0%	2	5.9	9	4.7%
3				3	9.2	10	11.5%
4	5.7	7	3.1%	4			
5	7.5	9	3.6%	5	10.5	12	3.8%
6	10	10	20.9%	6	12.5	15	16.7%
7				7			
8				8			
9				9	16	18	3.4%
10	11	13	3.6%	10	19	20	19.6%
11	14	15	16.0%	11			
12				12			
13				13			
14	16	20	16.1%	14	21	25	10.5%
15				15			
16				16	26	30	10.1%
17	21	25	4.5%	17			
18	26	30	4.8%	18	31	40	6.2%
19	31	60	3.6%	19	41	60	2.6%
20	62	100	4.5%	20	63	100	4.7%
missing	.i	.i	0.5%	missing	.i	.i	0.6%

Appendix C. Methods for creating the graphs

Child penalty graph. Following the replication files from Kleven et al. (2019b), we estimate

$$Earnings_{ist}^{g} = \sum_{j \neq -1} \alpha_j^g \cdot \mathbf{I}[j = t] + \sum_k \beta_k^g \cdot \mathbf{I}[k = age_{is}] + \sum_y \gamma_y^g \cdot \mathbf{I}[y = s] + v_{ist}^g \quad \text{(C1)}$$

where $Earnings_{ist}^{g}$ is annual earnings in Swedish Kronor (SEK) for individual *i* of gender *g* in year *s* at event time *t*. We include dummy variables for each event year from -5 to +10 except for the year just prior to parenthood (*t* = -1), making this year the reference category. The coefficients α_j^g on the event-time dummies capture the difference in average annual earnings between each event time and the year just prior before parenthood, residualized by age and year by including fixed effects for each in the regression.

An important recalculation of the estimates from Equation (1) takes place before creating the event-study graph. It normalizes the coefficients on the event-time dummies ($\hat{\alpha}_t^g$) "as a percentage of the counterfactual outcome, absent children" (Kleven et al. 2019, 190). The formula $P_t^g \equiv \hat{\alpha}_t^g / \mathrm{E}\left[\widetilde{Earnings}_{ist}^{g} | t\right]$ calculates these percentages P_t^g in each event year. Kleven et al (2019, 188) describe this counterfactual outcome ($\mathrm{E}\left[\widetilde{Earnings}_{ist}^{g} | t\right]$) as the "predicted outcome when omitting the contribution of the event dummies, i.e., $\tilde{Y}_{ist}^{g} = \sum_k \hat{\beta} E_k^g * I|k = age| + \sum_y \hat{\gamma} E_y^g * I|k = s|$." While this appears complicated, the number equates the average of the outcome variable plus the coefficient on the event-year dummy.

Figure C1 illustrates the normalization for women in our child penalty graph. The trend line with black circles shows women's average annual earnings in each event year in 1000s SEK. Average earnings increase until parenthood occurs at *t = 0* and drop to a lower level after the first child is born (note that women with zero earnings are included in the sample). Five years after parenthood, women in our sample made on average 176,709 SEK.

The gray squares in Figure C1 denote the estimated coefficients on the event-year dummies from Equation (C1). The coefficient for event-year five ($\hat{\alpha}_5^w$ = -98,002) indicates that women's average covariate-adjusted earnings were 98,002 SEK lower 5 years after parenthood compared to the year prior. The gray diamonds show women's counterfactual earnings without children, i.e., the sum of average earnings and the estimated coefficient in each event year. The normalized number plotted in the child penalty for event-year five is P_5^w≡(-98,002/(274,711) = -0.357. Ascribing a causal interpretation to the event-study

coefficients, it shows that parenthood depressed women's earnings by 36% in the fifth year after parenthood.

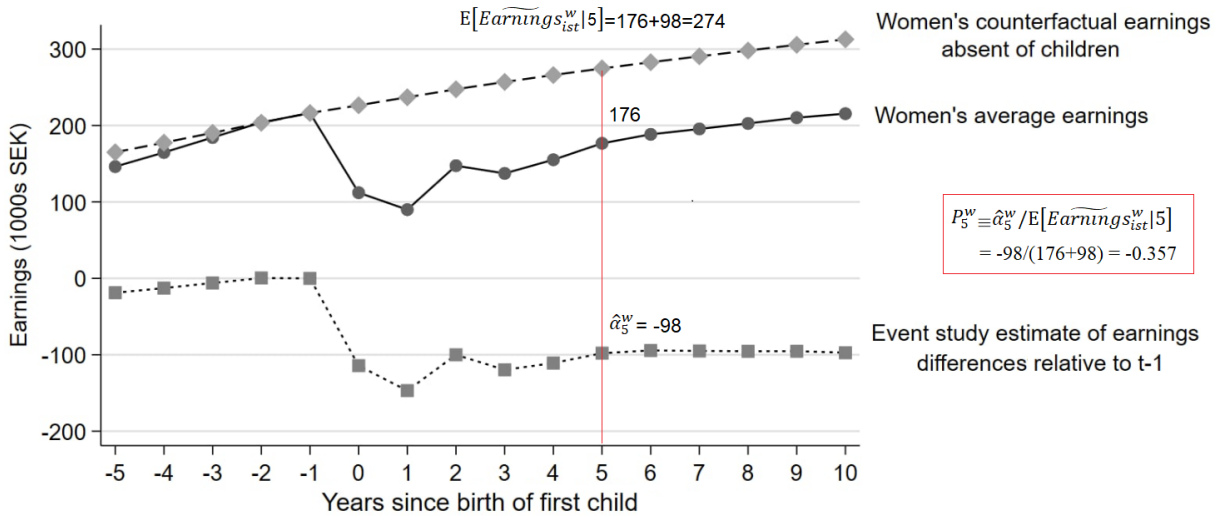


Figure C1. Normalization method for the trend line percentages in the child penalty graph.

The *child penalty graph* from Kleven et al. (2019a) displays a calculated number for the child penalty in event year 10. This calculation takes the difference between women's and men's alphas from Equation (1) at event year 10 and divides that difference by women's predicted *counterfactual* earnings in that event year ($P_t \equiv (\hat{\alpha}^{w}_{t} - \hat{\alpha}^{m}_{t}) / \mathrm{E}[\widetilde{Earnings}^{w}_{ist}|t]$).

Gender gap graph. The second graph combined a plot with trends in women's and men's average earnings in the event window with a plot estimating women's relative earnings. We obtain these estimates by pooling data for women and men and estimating

$$Earnings_{istg} = \sum_{j \neq -1} \alpha_j \cdot \mathbf{I}[j = t] \cdot \mathbf{I}[g = w] + \sum_{k} \beta_k \cdot \mathbf{I}[k = age_{isg}] + \sum_{y} \gamma_y \cdot \mathbf{I}[y = s] + \nu_{istg} \quad (2)$$

which differs from Equation (1) in two ways. We interact the event-year dummies with a dummy for female sex at birth and keep all event-year dummies. The estimated alphas now capture the female–male gap in the outcome variable in each event year, residualized for age and year using the same fixed effects as in Equation (C1).

To keep the formats similar to the child penalty graph, we continue using the recalculation method from Kleven et al. (2019a) to convert the point estimates into percentages. But instead of dividing the gender gap in earnings in each event year by women's counterfactual average earnings without children, we divide it by the counterfactual value for *men*. This is, arguably, a closer approximation of the quantity of interest because it captures how much women's earnings fall behind *men's* (without children). Dividing the earnings difference by men's earnings is the

standard approach used, for example, by the International Labour Organization and the Organisation for Economic Co-operation and Developement.

Appendix D. Supplementary tables and figures

Table D1. Within-respondent comparisons of beliefs about the gender gap in earnings.

Sample:	Child penalty graph		Gender gap graph		Child penalty graph with means	
Before or after seeing the graph:	Before	After	Before	After	Before	After
	(1)	(2)	(3)	(4)	(5)	(6)
0–5% before parenthood	0.18	0.48	0.19	0.04	0.18	0.14
Change in share	0.293***		-0.150***		-0.039***	
	(0.008)		(0.006)		(0.007)	
20–25% after parenthood	0.29	0.55	0.30	0.08	0.31	0.40
Change in share	0.264***		-0.224***		0.089***	
	(0.010)		(0.008)		(0.010)	
22% after parenthood	0.01	0.24	0.01	0.00	0.01	0.16
Change in share	0.228***		-0.004***		0.145***	
	(0.007)		(0.001)		(0.006)	
Observations	7,572		7,608		7,302	
Individuals	3,786		3,804		3,651	
Proportions of respondents switching into belief						
0–5% before parenthood	0.321		0.006		0.076	
20–25% after parenthood	0.366		0.035		0.244	
22% after parenthood	0.233		0.001		0.151	

Notes: Shares of respondents with beliefs about the gender gap in earnings in pre-specified ranges and values (left column) by treatment arm (top row) and split into columns for responses given before and after seeing the graph. Regression estimates from bivariate regressions with two observations per respondent, one for the pre-treatment and one for the post-treatment observation, and regressing a dummy for each belief of interest on a dummy for the post-treatment observation. Standard errors clustered at the respondent level in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table D2. Beliefs about a 0–5% gender gap in earnings of prior to parenthood.

DV: dummy for beliefs about the gender gap in earnings of 0–5% prior to parenthood	Static post-treatment	Pre-post switch	Static post-treatment	Pre-post switch
	(1)	(2)	(3)	(4)
Child penalty graph	0.448***	0.341***		
(reference=gender gap graph)	(0.008)	(0.010)		
Child penalty graph			0.251***	0.319***
(reference=child penalty graph with raw means)			(0.009)	(0.007)
Controls for				
Individual traits	Yes	Yes	Yes	Yes

Prior beliefs	Yes	Yes	Yes	Yes
Observations	7,486	7,341	7,341	7,486

Notes: Estimates from equation (1) with controls selected by a double LASSO procedure. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table D3. Within-respondent comparisons of beliefs about the gender gap in earnings; heterogeneity analysis by education level and occupation.

Sample:	All respondents		No university degree		PhD degree		Journalist, communicator, etc.	
Before or after graph:	Before	After	Before	After	Before	After	Before	After
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0–5% before parenthood	0.18	0.48	0.13	0.33	0.23	0.61	0.16	0.41
Change in share	0.293***		0.200***		0.372***		0.254***	
	(0.008)		(0.018)		(0.023)		(0.033)	
20—–5% after parenthood	0.29	0.55	0.28	0.47	0.30	0.60	0.32	0.51
Change in share	0.264***		0.181***		0.297***		0.185***	
	(0.010)		(0.024)		(0.027)		(0.040)	
22% after parenthood	0.01	0.24	0.01	0.14	0.01	0.27	0.01	0.17
Change in share	0.228***		0.132***		0.264***		0.153***	
	(0.007)		(0.013)		(0.019)		(0.025)	
Observations	7,572		1,362		1,076		496	
Individuals	3,786		681		538		248	
Proportions of respondents switching into belief								
0–5% before parenthood	0.321		0.225		0.400		0.295	
20–25% after parenthood	0.366		0.304		0.392		0.306	
22% after parenthood	0.232		0.137		0.268		0.165	

Notes: See notes to Table D1.

Table D4. Beliefs about the gender gap in earnings after parenthood.

DV: dummy for beliefs about the gender gap in earnings of 20–25% after parenthood	Static post-treatment	Pre-post switch	Static post-treatment	Pre-post switch
	(1)	(2)	(3)	(4)
Child penalty graph (reference=gender gap graph)	0.484***	0.329***		
	(0.009)	(0.008)		
Child penalty graph (reference=child penalty graph with raw means)			0.158***	0.114***
			(0.011)	(0.009)
DV: dummy for beliefs about the gender gap in earnings of 22% after parenthood				
Child penalty graph (reference=gender gap graph)	0.237***	0.235***		
	(0.007)	(0.007)		
Child penalty graph (reference=child penalty graph with raw means)			0.081***	0.082***
			(0.009)	(0.009)
Controls for individual traits	Yes	Yes	Yes	Yes
Controls for prior beliefs	Yes	Yes	Yes	Yes
Observations	7,486	7,341	7,272	7,272

Notes: Estimates from Equation (1) with controls selected by a double LASSO procedure. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table D5. Comparison of importance ratings of parenthood across subsamples more vs. less likely to have misunderstood the child penalty graph.

DV: Importance rating of parenthood as an explanation of economic gender inequality (1–5 scale)					
	(1)		(2)		(3)
Child penalty graph= 1 (Reference: gender gap graph)	0.011	CPG*Graph was helpful	0.090***	CPG*Switch to 0—5% belief	0.139***
	(0.020)		(0.021)		(0.028)
		CPG*Graph not helpful	-0.254***	CPG*No switch	-0.053**
			(0.032)		(0.022)
		(Reference: gender gap graph)		(Reference: gender gap graph)	
Controls for					
Individual traits	Yes		Yes		Yes
Prior beliefs	Yes		Yes		Yes
Standard deviation of the DV	0.85		0.85		0.85
Observations	7,312		7,312		7,312

Notes: Column (1) reports the results for estimating Equation (1) using the importance rating of parenthood as the outcome variable. Columns (2) and (3) run the same regression but split the dummy for the child penalty graph into two dummies. The split in column (2) subdivides it by whether the respondents reported the child penalty graph was "helpful" or "very helpful," and column (3) by whether they switched their belief about the gender gap in earnings prior to parenthood into the 0–5% range. The gender gap graph is the reference category in both of these regressions. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table D6. Comparison of importance ratings of parenthood across subsamples of respondents who saw the child penalty graph in the survey experiment.

DV: Importance rating of parenthood as an explanation of economic gender inequality (1–5 scale)				
	(1)	(2)	(3)	(4)
Graph was helpful=1 (Reference: not helpful)	0.206***	0.175***		
	(0.029)	(0.031)		
Switch to 0–5% earnings gap=1 (Reference: no switch)			0.362***	0.290***
			(0.032)	(0.032)
Controls for individual traits		Yes		Yes
Controls for prior beliefs		Yes		Yes
Standard deviation of DV	0.85	0.85	0.85	0.85
Observations	3,718	3,634	3,718	3,634

Notes: The sample is respondents who viewed the child penalty graph (Figure 2A) in the survey experiment. Estimates from regressing the respondent's rating of how important parenthood is for economic gender inequality on a dummy for answering that the graph was "helpful" or "very helpful" for answering the questions about the gender gap in earnings (columns 1 and 2) or a dummy for switching their belief about the size of the gender gap in earnings before parenthood into the 0–5% range after viewing that graph (columns 3 and 4). Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table D7. Comparison of self-reported social media engagement with the child penalty graph across subsamples more likely to have misunderstood its content.

DV: Index for the willingness to engage with the graph (like, share, or cite-post) on social media (1—5).				
	(1)	(2)	(3)	(4)
Graph was helpful (Reference: not helpful)	0.465***	0.471***		
	(0.044)	(0.045)		
Switch to 0—5% earnings gap prior to parenthood (Reference: no switch)			0.012	-0.014
			(0.041)	(0.044)
Controls for individual traits		Yes		Yes
Controls for prior beliefs		Yes		Yes
Standard deviation of DV	1.07	1.07	1.07	1.07
Observations	3,072	3,005	3,072	3,005

Notes: The sample is respondents who viewed the child penalty graph (Figure 2A) in the survey experiment, restricted to self-reported social media users. Estimates from regressing an index of the respondent's rating of their likelihood to "like," share, or cite-post the graph on their social media on a dummy for stating the graph was "helpful" or "very helpful" for answering the questions about the gender gap in earnings (columns 1 and 2) or a dummy for switching their belief about the size of the gender gap in earnings before parenthood into the 0–5% range after viewing that graph (columns 3 and 4). Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Appendix E. Analysis of approximately accurate beliefs

Accurate beliefs about levels of gender inequality. We define intervals of approximately accurate beliefs about the gender gap in earnings using a method that accounts for practical

difficulties in extracting exact and correct numerical values from the graphs, as well as vague time indications in our survey questions (e.g., asking about "*people who are not yet parents*" instead of "*people two years away from having their first child*"). Allowing imbalanced intervals around exact gap sizes from the microdata further reflects rounding to numbers in their vicinity, i.e., reflecting people's likelihood to gravitate toward multiples of 5 and 10 (Talbot et al. 2014). Table E1 presents our intervals.

Figure E1 shows belief distributions before and after seeing each graph, with dashed lines indicating the intervals of interest. This section describes the results from our sequence of within- and across-individual analyses, and Tables E2–E3. The results are insensitive to using exactly balanced intervals of 5 or 10 percentage points (see Table E4).

Viewing the *child penalty graph* reduces the accuracy of respondents' belief about the level of gender inequality before parenthood. The share of respondents with approximately accurate beliefs shrinks by 19 percentage points from 0.31 to 0.12 for the *adjusted gender gap*, and by 4 percentage points from 0.08 to 0.04 for the *raw gender gap.* By contrast, the share with approximately accurate beliefs about the *adjusted gender gap* among parents increases by 10 percentage points from 0.10 to 0.20. Accuracy remains constant for the *raw gender gap* after parenthood.

The *gender gap graph* produces large increases in approximately accurate beliefs about gender inequality before and after parenthood. These improvements mainly occur for the *adjusted gender gap* in earnings. The proportion with accurate beliefs increases by 29 percentage points from 0.32 to 0.61 for that gap before parenthood, compared to 5 percentage points (from 0.09 to 0.14) for the raw gap. Accuracy for the *adjusted gender gap* after parenthood increases dramatically by 53 percentage points (from 0.10 to 0.63), compared to 3 percentage points for the raw gap. Respondents appear to have easily extracted information from the subplot with regression estimates for women's relative earnings, and largely ignored the subplot with raw means.

The *child penalty graph with raw means* generates modest increases in belief accuracy for inequality levels before and after parenthood. This time, improvements occur for the *raw gender gap*, with the share of approximately accurate beliefs increasing by 19 percentage points from 0.09 to 0.28 before parenthood, and by 15 percentage points from 0.06 to 0.21 after

parenthood.[2] Accuracy for the *adjusted gender gap* drops by 7.5 percentage points before parenthood, and improves slightly by 4 percentage points after parenthood. Some additional observations indicate that respondents found this graph less clear than the others. A larger variance in post-treatment gender gap sizes may reflect struggles to correctly calculate ratios of average earnings from the trend lines. Only 18% of respondents rated this graph as "very helpful" for assessing the earnings gap, compared to 30% for the *child penalty graph*, and 37% for the *gender gap graph*.

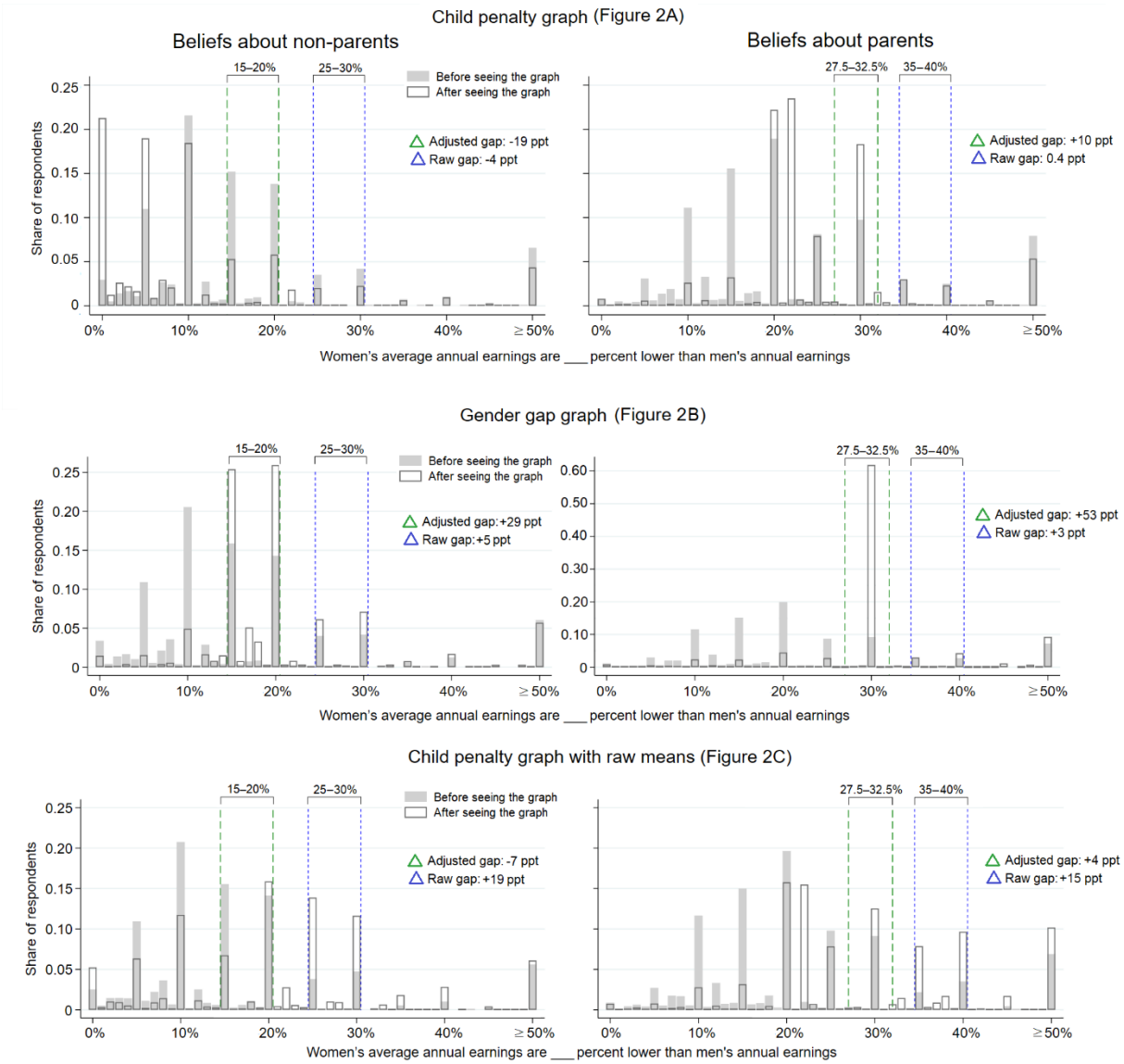


Figure E1. Approximately accurate beliefs about the gender gap in earnings.

Notes: Green dashed lines indicate intervals for approximately accurate beliefs about the gender gap in earnings adjusted for age and year fixed effects. Blue dashed lines indicate approximately accurate beliefs about the raw gender gap in earnings. Data for each graph is restricted to respondents who answered the questions both before

[2] We can speculate that larger improvements might have resulted from starting the vertical axis at zero in the subplot with raw means.

and after seeing the graph. N(child penalty graph)=3,786; N(gender gap graph)=3,804, N(child penalty graph with raw means)=3,651.

Table E1. Intervals of approximately accurate beliefs about the gender gap in earnings.

	Exact value		Approximate interval		Exact 5ppt interval	
	Adjusted	Raw	Adjusted	Raw	Adjusted	Raw
Before parenthood	17.1	27.4	15–20	25–30	14.6–19.6	24.7–29.7
After parenthood	29.4	36.2	27.5–32.5	35–40	26.9–31.9	32.9–38.9

Notes: the exact values before parenthood were calculated based on the percentage normalization from Kleven et al. (2019a) for the 5 years before parenthood, and for t=10 for the value after parenthood.

Table E2. Within-respondent analysis of approximately accurate beliefs.

Sample:	Child penalty graph		Gender gap graph		Child penalty graph with raw means	
Before or after graph:	Before	After	Before	After	Before	After
	(1)	(2)	(3)	(4)	(5)	(6)
A. Before parenthood						
15–20% adjusted gap	0.312	0.121	0.322	0.608	0.318	0.243
Change in share	-0.191***		0.287***		-0.075***	
	(0.008)		(0.010)		(0.009)	
25–30% raw gap	0.080	0.045	0.086	0.140	0.089	0.278
Change in share	-0.036***		0.054***		0.190***	
	(0.005)		(0.006)		(0.008)	
Observations	7,572		7,608		7,302	
Individuals	3,786		3,804		3,651	
Proportions of respondents switching into belief:						
15–20% adjusted gap	0.046		0.372		0.127	
25–30% raw gap	0.026		0.109		0.234	
B. After parenthood	Before	After	Before	After	Before	After
27.5–32.5% adjusted gap	0.103	0.205	0.097	0.629	0.094	0.139
Change in share	0.101***		0.532***		0.045***	
	(0.007)		(0.009)		(0.007)	
35–40% raw gap	0.057	0.060	0.054	0.080	0.060	0.208
Change in share	0.004		0.027***		0.148***	
	(0.005)		(0.005)		(0.007)	
Observations	7,572		7,608		7,302	
Individuals	3,786		3,804		3,651	
Proportions of respondents switching into belief:						
27.5–32.5% adjusted gap	0.159		0.559		0.111	
35–40% raw gap	0.043		0.065		0.186	

Notes: Shares of respondents with beliefs about the gender gap in earnings in ranges of approximately correct values based on either covariate-adjusted or raw numbers (left column) by treatment arm (top row) and split into columns for responses given before and after seeing the graph. Regression estimates from bivariate regressions with two observations per respondent (one for the pre-treatment and one for the post-treatment observation) and

regressing a dummy for each belief of interest on a dummy for the post-treatment observation. Standard errors clustered at the respondent level in parentheses, * p<0.01, ** p<0.05, * p<0.1.

Table E3. Comparison across graphs for accurate beliefs; approximate intervals around exact values from the microdata (see Table B5).

DV: dummy for accurate beliefs before parenthood	Adjusted gap = 15–20%		Raw gap = 25–30%	
	Static post-treatment	Pre-post switch	Static post-treatment	Pre-post switch
	(1)	(2)	(3)	(4)
Child penalty graph	-0.485***	-0.099***	-0.099***	-0.401***
(reference=gender gap graph)	(0.009)	(0.006)	(0.006)	(0.010)
	7,486	7,486	7,486	7,424
Child penalty graph	-0.124***	-0.234***	-0.234***	0.048***
(reference=child penalty graph with raw means)	(0.009)	(0.008)	(0.008)	(0.008)
Observations	7,341	7,341	7,341	7,272
DV: dummy for accurate beliefs after parenthood	Adjusted gap = 27.5—32.5%		Raw gap = 35—40%	
	Static post-treatment	Pre-post switch	Static post-treatment	Pre-post switch
	(1)	(2)	(3)	(4)
Child penalty graph	-0.430***	-0.401***	-0.021***	-0.022***
(reference=gender gap graph)	(0.010)	(0.010)	(0.006)	(0.005)
Observations	7,424	7,424	7,424	7,424
Child penalty graph	0.063***	0.048***	-0.150***	-0.145***
(reference=child penalty graph with raw means)	(0.009)	(0.008)	(0.008)	(0.007)
Observations	7,272	7,272	7,272	7,272
Controls for				
Individual traits	Yes	Yes	Yes	Yes
Prior beliefs	Yes	Yes	Yes	Yes

Notes: Estimates from Equation (1) with controls selected by a double LASSO procedure. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table E4. Comparison of accurate beliefs across graphs, exact ranges (see Table D5).

DV: static dummy for post-treatment beliefs about the gender gap in earnings	Adjusted gap: 17.1% midpoint		Raw gap: 27.4% midpoint	
	5 ppt interval	10 ppt interval	5 ppt interval	10 ppt interval
	(1)	(2)	(3)	(4)
Child penalty graph (reference=gender gap graph)	-0.287***	-0.495***	-0.048***	-0.102***
	(0.009)	(0.010)	(0.005)	(0.007)
	7,486	7,486	7,486	7,486
Child penalty graph (reference=child penalty graph with raw means)	-0.023***	-0.137***	-0.141***	-0.243***
	(0.006)	(0.009)	(0.006)	(0.008)
Observations	7,341	7,341	7,341	7,341
DV: dummy for accurate beliefs after parenthood	Adjusted gap: 29.4% midpoint		Raw gap: 36.2% midpoint	
	5 ppt interval	10 ppt interval	5 ppt interval	10 ppt interval
Child penalty graph (reference=gender gap graph)	-0.440***	-0.371***	0.001	-0.006
	(0.010)	(0.011)	(0.004)	(0.006)
Observations	7,424	7,424	7,424	7,424
Child penalty graph (reference=child penalty graph with raw means)	0.057***	0.060***	-0.073***	-0.152***
	(0.009)	(0.010)	(0.006)	(0.008)
Observations	7,272	7,272	7,272	7,272
Controls for individual traits	Yes	Yes	Yes	Yes
Controls for prior beliefs	Yes	Yes	Yes	Yes

Notes: Estimates from Equation (1) with controls selected by a double LASSO procedure. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Change in inequality with parenthood. Figure E2 displays the results of analyzing percentage-point differences between respondents' answers about the gender gap in earnings before and after parenthood. Before seeing their graph, most respondents stated similar gender gaps before and after parenthood. A quarter (24%) of the percentage-point differences between the two answers indicate perceptions of either no change or a reduction in inequality. Nearly two-thirds (63%) of respondents reported small differences of 5 percentage points or less, and only 4% perceived large gender gap expansions of 20 percentage points or more. These priors understate the actual change in the microdata. The covariate-adjusted earnings gap grows by 15 percentage points between $t = -1$ and $t = 10$, or by 12 percentage points if we use the average gap 5 years before parenthood as the initial value (i.e., from 17% to 29%). Corresponding increases in the raw gender gap are 9 and 12 percentage points.

Seeing the *child penalty graph* improves respondents' understanding that inequality increases with parenthood. After seeing the graph, perceptions of constant or declining

inequality drop to 14% of the sample, and perceptions of small increases (≤ 5 percentage points) drop to 21%. The median post-treatment value of 15 percentage points perfectly matches the change in the adjusted earnings gap between *t = -1* and *t = 10* in the microdata. At the same time, 40% perceive exaggerated increases of the gender gap of 20 percentage points or more.

The *gender gap graph* increases perceptions that inequality rises with parenthood and provides more accurate perceptions about magnitudes. Perceptions of constant or declining inequality drop to 17%, and perceptions of small increases fall to 26%. A clear majority of respondents have values close to the accurate numbers for the adjusted gender gap in earnings (i.e., 12 and 15 percentage points; see hollow bars in Figure 9) and we observe nearly no increase in overstated gap sizes above 20 percentage points.

The *gender gap with raw means* produces two types of inaccurate beliefs. The share of respondents that believes inequality decreases or remains constant after parenthood *increases* modestly to 27% after seeing this graph. At the same time, the share perceiving overstated increases of 20 percentage points or more grows from 4% to 18%. While the median post-treatment response of 10 percentage points is close to the accurate number, fewer answers fall in the vicinity of those numbers compared with the other graphs.

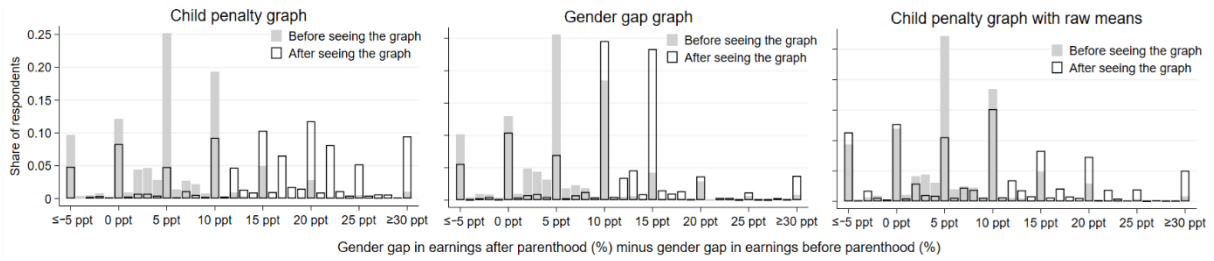


Figure E2. Perceived changes in the earnings gap with parenthood.

Notes: Each respondent's answer for the gender gap in earnings after parenthood is subtracted from their answer about the gender gap before parenthood. This gives one pre-treatment and one post-treatment observation per respondent. Data for each graph is restricted to respondents who answered those pre- and post-treatment questions. N(child penalty graph)=3,786; N(gender gap graph)=3,804, N(child penalty graph with raw means)=3,651.

Appendix F. Supplementary survey and summary statistics

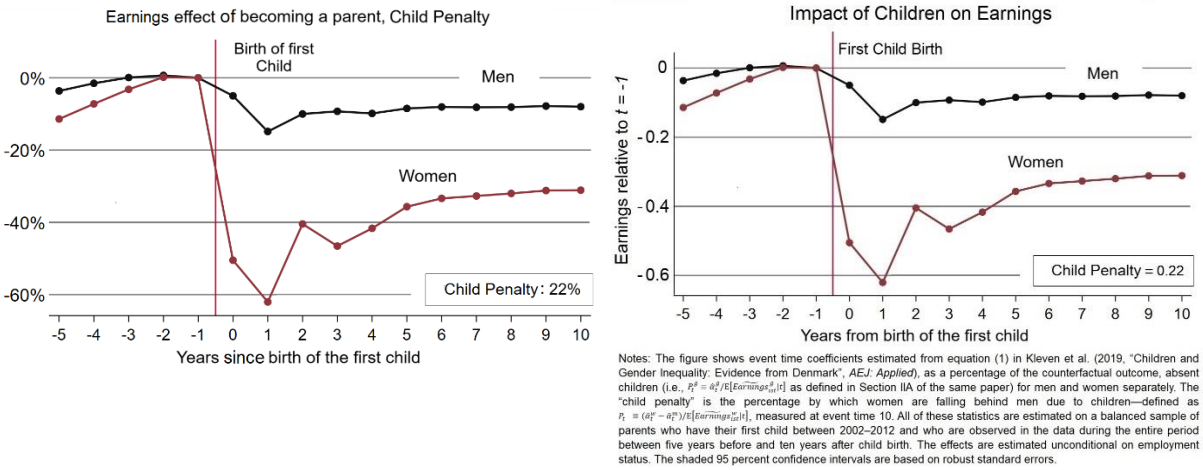


Figure F1. Simplified (left) and technical (right) graphs in the supplementary survey.

Notes: The simplified graph seeks to mimic the format from childpenaltyatlas.org (right side of Figure 1), and the technical graph seeks to mimic the format from Kleven et al. (2019a, left side of Figure 1).

Table F1. Survey text and questions.

	Welcome! This short survey asks questions about a specific research graph and two background questions about you. Answers are completely anonymous, and you can skip any question by clicking the arrow button. Please direct questions about the survey to olle.folke@hhs.se. **Do you agree to take this voluntary and anonymous survey?** Response categories: I consent, begin the study; I do not consent, I do not wish to participate.
Q1	Which category best describes your current job? Response categories: Economics research or PhD studies; Other research or PhD studies; A job requiring a PhD degree; Other job; Student or not working
Q2	Where do you live? Response categories: Europe; North America; Asia; Africa; South America; Oceania
Q3	This graph examines the effect of parenthood on women relative to men. [GRAPH HERE]. Have you seen a graph for the "child penalty" before? Response categories: Yes, in passing; Yes, I use it in my research; No.
Q4	We ask you to evaluate three statements about the results in the graph. Please choose the answer closest to your own opinion. We learn the from this graph that... Items a) ...in the years just before parenthood, the gender gap in earnings is 0–5%. b) ...10 years after parenthood, women earn 20–25% less than men. c) ...parenthood explains nearly all gender inequality in earnings. Response categories: Yes, this seems correct; No, the number seems smaller/larger; No, the graph lacks this information; Don't know
Q5	The last two questions are more demanding. **What does the graph tell us about women's and men's earnings in the years just before parenthood?** *Please write your answer in the text box below. Note that answers like "I don't know" or "I'm not sure" are just as valuable as more elaborate answers. Click the arrow button to skip the question.* *For the years just before parenthood, the graph shows that...* [TEXT BOX]
Q6	This is the last question. **How would you interpret the child penalty statistic in the bottom right corner of the graph?** *Please write your answer in the text box below.* *The child penalty statistic of 22% implies that...* [TEXT BOX]
End	Thank you very much for participating! Please help us by sharing the survey link with others! You can send questions to olle.folke@hhs.se and leave comments here [TEXT BOX]

Table F2. Summary statistics and balance tests.

	Simplified graph	Technical graph	Difference (OLS)	P-value
Current job				
Job requiring a PhD degree	0.12	0.13	0.01	0.70
Economics research or PhD studies	0.43	0.40	-0.03	0.39
Other job	0.11	0.16	0.04	0.08
Other research or PhD studies	0.26	0.24	-0.02	0.51
Student or not working	0.08	0.08	0.00	0.97
Familiarity with the graph				
No	0.29	0.32	0.03	0.41
Yes, I use it in my research	0.09	0.10	0.01	0.54
Yes, in passing	0.62	0.58	-0.04	0.26
Current continent of residence				
Europe	0.84	0.83	-0.01	0.62
North America	0.12	0.13	0.01	0.79
Other global regions	0.04	0.05	0.01	0.63

Table F3. Proportions of missing data (%).

Question	1	2	3	4a	4b	4c	5	6
Simplified graph	0.52	1.05	2.36	13.61	13.61	13.35	30.1	31.41
Technical graph	0.00	0.26	2.34	17.19	16.67	17.19	35.16	37.50

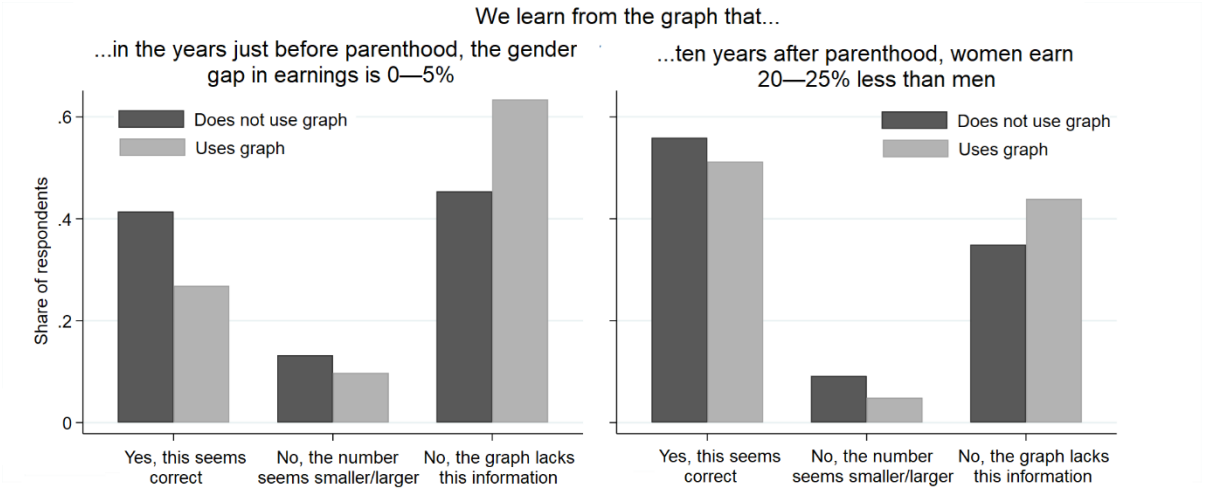


Figure F2. Comparison of graph users and non-users among academic economists.

Notes: "Don't know" has been set to missing (N=4 for the left plot and N=3 for the right plot). N(non-users, left)=227, N(non-users, right)=229 N(users, left and right)=41

Table F4. Coding of free-text answers.

What does the graph tell us about women's and men's earnings in the years just before parenthood?	
Group	Definition
Correct answer	Answers close to the following: - [Parallel/converging] pre-trends in earnings after controlling nonparametrically for lifecycle and time trends. - the outcome for women and men evolves [in parallel/by women catching up] until the birth of their first child. - [similar/parallel/converging] earnings trajectories of men and women are prior to having children Answers less technical in nature, such as saying that trend lines overlap due to the temporal benchmark. Excluding cases where it is clear that the person interprets the lines as the level of earnings, for example by saying that the level of earnings "peaks" the year just before parenthood.
Small or no gender gap in earnings	Women and men have equal/similar earnings; the gender gap is very small/close to zero/goes to zero.
How would you interpret the child penalty statistic in the bottom right corner of the graph?	
Group	Definition
Correct answer	Relatively close to the definition in Kleven et al. (2019) that "This child penalty measures the percentage by which women are falling behind men due to children at event time t." We count an answer as correct regardless of saying "percent" or "percentage points." We count an answer as correct if it also includes misunderstood aspects, such as the number representing the distance between the lines at a certain time point, or a percentage normalization relative to the level of earnings in the baseline year.
Gender gap in earnings	The gender gap in earnings after parenthood.

Table F5. Examples of answers in hand-coded categories.

What does the graph tell us about women's and men's earnings in the years just before parenthood?	
Correct	**Small or no gender gap in earnings**
Women's income was rising faster (as a share of their year b_-1 income) than men's.	There is a gender gap in earnings that gets smaller and smaller
Changes in earnings are relatively similar, with women having somewhat lower wage growth (or rather somewhat higher wage reductions)	Two years before the first child is born, wages are equal for men and women. Before that wages of men are slightly larger than wages of women
Women's earnings are growing slightly faster than men's in the years just before parenthood.	That women and men start differently 5 years before parenthood but seem to earn about the same just about becoming a parent. (Although the heading of the graph does not state that this is the focus)
The earnings evolved parallel between men and women.	Earnings seem to be similar but we need the CIs to claim this.
Wages for men were quite flat; women wages increased slightly	The gender gap is about zero one and two years before the birth of a child. Before that, the gender pay gap ranges between about 5 (year -5) and 3 (year -3)
How would you interpret the child penalty statistic in the bottom right corner of the graph?	
Correct	**Gender gap in earnings**
Very unclear. Seems to be the difference between men and women, in percentage points, in earnings effect of becoming a parent, ten years after birth of first child.	Women earn 22 percentage points less than men after giving birth on average
Women earn 22 percent less than man due to child birth	That the gender gap appears to be 22% when it reaches a relatively stable state after 10 years or so.
The negative percentage effect of parenthood on women's earnings is 22 percentage points larger than the negative effect on men's earnings	Ten years after first child women earn 22 percent less then men.
The average difference in women's penalty and men's penalty for having children is 22%, relative to people who don't have kids.	For all the years after parenthood taken together the earnings of women are 0.22 less them men.
significant negative impact parenthood has on women's careers, causing their earnings, employment, and career progression to fall behind men's after having children	Relative to men who also had their first child 10 years prior, women earn 22% less.

Table F6. Distribution of free-text answers about the child penalty graph (%) among academic economists.

What does the graph tell us about women's and men's earnings in the years just before parenthood?					
Sample:	All	Atlas	Article	Atlas	Article
Small or zero gender gap in earnings	33	32	35	22	24
Correct answer	45	43	48	29	34
Other	22	22	18	17	13
Missing				32	30
Observations	217	110	107	162	152
How would you interpret the child penalty statistic in the bottom right corner of the graph?					
Sample:	All	Atlas	Article	Atlas	Article
Gender gap in earnings	23	20	25	14	16
Correct answer	40	37	42	25	27
Other	38	43	32	29	21
Missing				33	35
Observations	208	109	99	162	152

Notes: percentages of respondents in hand-coded categories of free-text answers.

References

Kleven, H., Landais, C., & Søgaard, J. E. (2019a). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics, 11*(4), 181-209.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., & Zweimüller, J. (2019b). Child penalties across countries: Evidence and explanations. In *AEA Papers and Proceedings* (Vol. 109, pp. 122-126).