Children, Household Specialization and Relationship Quality*

Belén Rodríguez Moro† Olatz Román‡

(Job Market Paper)

This version: December 20, 2024
Link to latest verison here

Abstract

We investigate how having children impacts the quality of couples' relationships, a proxy of the non-material gains from being in a relationship. Using a novel measure of relationship quality (RQ), we perform a dynamic difference-in-differences estimation around the birth of the first child. We find a sharp and lasting decrease in RQ immediately after birth. We attribute this effect to changes in household specialization. Traditional gender-based specialization prevails after birth, regardless of the baseline distribution of tasks within the couple. Leveraging heterogeneous changes in household specialization after birth, we find that couples undergoing larger rearrangements also suffer larger RQ drops.

Keywords: Fertility, Marital decisions, Time allocation, Household specialization

JEL Codes: J12, J13, J22, D13

*We are grateful to Sule Alan, Thomas F. Crossley, David K. Levine, Libertad González and Alexander Monge-Naranjo for their invaluable guidance and support. We would also like to thank Alessandro Tarozzi, Andrea Ichino, Jonas Jessen, Peter Levell, Yannick Reichlin, and seminar and conference participants at the European University Institute, Universitá di Firenze, Universitat Pompeu i Fabra, the 18th Doctorissimes Conference, Universitat de Barcelona, Universidad Carlos III de Madrid, Universidad de Salamanca, the BSE Jamboree, the EDP Jamboree, the Understanding Society Scientific Conference 2023, the 2023 EEA-ESEM congress, the 2023 EALE conference, the 48th SAEe symposium, IZA Institute for Labor Economics, SOLE 2024, University of the Basque Country, the XVI AEET Labor Economics Conference, the 17th UniTO-Collegio Carlo Alberto Ph.D. Workshop in Economics and the 2024 European Winter Meeting of the Econometric Society for helpful comments and suggestions.

†European University Institute, belen.rodriguez@eui.eu.
‡European University Institute, olatz.roman@eui.eu.

1 Introduction

Having children changes peoples' lives. With more responsibilities at home, including housework and childcare, couples need to reorganize how they allocate their time. Mothers experience a significant reduction in the time spent in the labor market (Goldin, 2021), home production is readjusted accordingly (Aguilar-Gomez et al., 2019; Siminski and Yetsenga, 2022), and leisure and resting become second order for both parents (Aguiar and Hurst, 2007; Costa-Font and Flèche, 2020). Despite these significant adjustments that couples undergo when having children, the deeper consequences that parenthood and these accompanying changes may have on couples have received limited attention.

This paper studies the impact of having children on the quality of couples' relationships. Relationship quality is a measure of the non-material benefits that individuals derive from their relationships, and is widely used in the theoretical literature of family economics as a component of utility that influences marital decisions. While models studying family formation and dissolution widely consider the notion of relationship or match quality, its underlying determinants are little understood (Browning et al., 2014; Chiappori, 2020). To our knowledge, only a few papers have attempted to empirically measure relationship quality to guide such models (Weiss and Willis, 1997; Chiappori et al., 2018).

Studying relationship quality in the context of children is of particular relevance for several reasons. Firstly, theoretical assessments have highlighted that the quality of the parental relationship plays a critical role in decisions regarding investments in child education. In this context, children are considered public goods and investments in them rely on the perceived stability of the parental relationship, dictated by match quality (Chiappori and Weiss, 2006, 2007). Secondly, poor relationship quality can lead to couple dissolution. This negatively influences children's long-term outcomes directly through parental separation (Gruber, 2004; Björklund et al., 2007), but even before separation, from exposure to a low-quality parental relationship (Piketty, 2003; Björklund and Sundström, 2006). Understanding the changes in relationship quality when a child is born can improve our knowledge of these issues and provide insights into avenues to mitigate the negative impact.

To investigate the effect of having children on parents' relationships, we construct a novel measure of relationship quality (RQ). We use a questionnaire from Understanding Society, a household longitudinal panel in the United Kingdom, which collects information about relationships with cohabiting partners. The questionnaire items provide insights into subjective assessments about the relationship, such as the degree of happiness with the couple or the frequency with which they consider splitting, and couple time use, like how often they work together on a project or kiss. This rich set of information allows us to gain a nuanced understanding of the non-material benefits from being in the current relationship. To capture the multiple dimensions of couples' relationship quality in a parsimonious way, we combine the

responses to the questionnaire in a factor analysis and construct the RQ measure. We conduct several validation exercises that ensure the capacity of this measure to predict marital outcomes such as couple dissolution and marriage decisions. Our companion paper, Rodríguez-Moro and Román (2024), provides more details on the construction, validation and analysis of the measure.

We leverage variation in the timing of the birth of the first child to estimate the dynamic impact of having children on the quality of parents' relationships. Using a dynamic difference-in-differences approach, we remove individual unobserved heterogeneity and focus on the evolution of RQ with time relative to birth. We estimate this specification following the method proposed by Callaway and Sant'Anna (2021).

We find a sharp and steady decrease in RQ in the four years after the birth of the first child. This decline stabilizes at approximately half a standard deviation below pre-birth levels. Illustratively, the RQ of individuals who ranked in the 75th percentile of RQ before having their first child is reduced to median RQ within the first four years after birth, and such a decrease in RQ is associated with a 12% higher probability of dissolution. This impact lasts for the entire seven-year period examined. In contrast to most of the child penalty literature, this impact is largely symmetric for mothers and fathers, challenging the notion that fathers are unaffected by parenthood.

The magnitude and persistence of the estimated effect are sizeable when considering the large and positive association between RQ and couple dissolution, and in comparison to the effect that fertility has on general happiness. Notably, the effect on general happiness is positive, although about half the size of the impact on RQ, and it is short-lived, lasting only one period. We consider RQ a component of happiness exclusively attached to the couple and distinct from the happiness derived from having a child (Chiappori et al., 2018). Based on that, we interpret the impact on happiness as the sum between a decrease in happiness from one's partner and an increase in happiness from the child. Moreover, the impact of childbirth on RQ is larger than other major life events, such as becoming unemployed, which has virtually no impact on RQ.

The causal interpretation of these results is contingent on two assumptions: limited anticipation and conditional parallel trends. We provide evidence supporting the plausibility of these assumptions. Crucially, we observe flat trends in RQ in the periods leading up to birth, which is not masking any significant heterogeneity in pre-trends. Moreover, delays in fertility, such as miscarriages or the need for fertility treatments, do not lead to differences in the impact of childbirth on RQ.

We establish the robustness of our findings addressing four potential concerns. First, we demonstrate that our results are not solely driven by a reduction in time spent together as a couple by conducting separate analyses on the different item blocks, all of which are impacted by the arrival of the first child. Second, we rule out the possibility that subsequent children drive the lasting impact on RQ repeating the analysis separately for couples with different total

number of children. Third, we address concerns about potential selection bias associated with the unavailability of RQ data after a couple dissolves, showing that attrition from the sample due to separation is minimal and that the results do not change repeating the analysis solely on couples who do not dissolve. Lastly, we verify that the timing of birth, in terms of parental age and relationship tenure, does not influence our results.

In light of these findings, we delve into the mechanisms at play. We start by documenting that the arrival of a child introduces significant and unprecedented changes in how couples allocate their time, substantially increasing the demand for routine housework, which does not include childcare. This increase is almost fully borne by women, who devote five additional hours to housework every week on average, which is offset by a corresponding reduction in their labor market participation. At the same time, men's housework time increases in one weekly hour, on average, while their labor market time use remains largely unaffected. These findings align with previous research (Goldin, 2021), including Kuziemko et al. (2018), who find that mothers often underestimate the extent of this time rearrangement.

To measure the extent of household specialization, we compute the female share of the total paid and unpaid work hours within the couple and categorize couples into distinct groups based on the division of responsibilities before birth. We establish that, regardless of the division of paid and unpaid work before child birth, all types of couples adopt gender-based household specialization after child birth. While prior studies have identified an average effect on household specialization, our research is the first to reveal that this trend prevails across various degrees of pre-birth household specialization. This finding is not consistent with household specialization being a product of comparative advantages within the couple, pointing to the existence of labor market frictions or identity considerations after birth (Akerlof and Kranton, 2000; Ichino et al., 2019).

Leveraging the documented heterogeneity in the redistribution of labor and home production after first child birth, we study how changes in household specialization induced by children may mediate the effects on relationship quality. In doing so, we conduct a separate analysis for different couple types based on their degree of specialization before birth. We find that couples experiencing the largest changes in the share of housework done by women tend to experience more pronounced declines in RQ.

Our findings support the notion that household specialization driven by the arrival of children may have adverse implications for couples' relationship quality. These results have implications for the design and choice of policies aimed at promoting fertility. Policies that aim to support more equitable distributions of responsibilities within households may mitigate this negative effect. Such policies have been linked to favorable female labor market outcomes (Olivetti and Petrongolo, 2017) and have been shown to increase men's contribution to housework (Farré and González, 2019). Moreover, Avdic and Karimi (2018) find that such policies may anticipate couple separation, possibly due to a quicker revelation of the fully unravelled

decrease in RQ. Therefore, implementing these policies could yield both immediate benefits, by improving parents' RQ, and long-term advantages, by perpetuating more equitable arrangements in the relationships of future generations.

Related literature. This paper connects to the extensive literature on the economics of family formation and dissolution. Standard models studying family formation and dissolution commonly integrate the concept of relationship or match quality into the decision-making process (Weiss and Willis, 1997; Browning et al., 2014; Chiappori and Mazzocco, 2017; Chiappori, 2020). However, the underlying determinants of this variable are little understood.[1] Empirical efforts to guide such models are limited, proposing general characteristics of partners or subjective well-being as proxies for match quality (Weiss and Willis, 1997; Bertrand et al., 2015; Chiappori et al., 2018). We develop a novel measure that explicitly integrates information about relationships, which has the potential to yield a more precise and reliable assessment in this context, shedding light on the process followed by this variable, in terms of its evolution over the life cycle, and on how it reacts to life events.

Previous studies in psychology have proposed and tested various measures of marital satisfaction and conflict (e.g., Norton, 1983; Busby et al., 1995; Joel et al., 2020). These variables have previously been studied across social sciences in relation to different outcomes. For instance, Carlson and VanOrman (2017) study the relationship quality trajectories of married and unmarried parents, only after birth. Our measure and analysis present a number of advantages over prior research. First, RQ integrates multiple aspects of relationships, for which only separate measures have been available, into a single, parsimonious measure of match quality (Busby et al., 1995). Second, the combination of a larger sample and longitudinal data allows us to employ causal identification methods that were previously unfeasible, overcoming limitations acknowledged in earlier studies (Amato and Booth, 2001; Hassebrauck and Fehr, 2002; Amato and Patterson, 2017).

This paper adds to our understanding of the consequences of having children. Seminal papers in gender economics have documented the differential effect of fertility on women relative to men on several labor market outcomes (Blau and Kahn, 2017; Bertrand, 2020; Goldin, 2021). A great volume of work has focused on accounting for the share of the gender wage gap explained by fertility and on studying the channels through which it operates (Goldin, 2014; Adda et al., 2017; Kleven et al., 2019b; Cortés and Pan, 2020; Cavapozzi et al., 2021; Kleven,

[1]The models of couple formation and dissolution incorporate match quality in two ways: through a match quality parameter and through partner quality. A set of papers have modelled relationship quality as a parameter that is known in the current period but that receives stochastic shocks in the future (Chiappori and Weiss, 2006, 2007; Gemici and Laufer, 2011; Bruze et al., 2015; Voena, 2015; Goussé et al., 2017; Low et al., 2018). Others consider it to be deterministic but unobserved, where couple members only get a noisy signal each period about true quality (Brien et al., 2006; Blasutto et al., 2020; Antler et al., 2022). A second strand of models consider match quality to be contingent on partner quality in terms of different socio-economic characteristics (Greenwood et al., 2017; Eckstein et al., 2019; Low, 2024).

2022; Albanesi et al., 2023), in different countries (Kleven et al., 2019a, 2023) and across different cohorts (Goldin, 2021). Other work has discussed the deeper implications of having children on outcomes like identity (Akerlof and Kranton, 2000, 2010; Bertrand et al., 2015), well-being (Dolan et al., 2008; Clark et al., 2008; Blanchflower, 2009; Ferrer-i Carbonell, 2013) and mental health (Ahammer et al., 2023).[2]

The findings across these studies consistently reveal a stark asymmetry: fathers experience little to no impact, while mothers face significant consequences. In contrast, this paper identifies a consequence of childbirth shared by both parents, offering novel insights into the experiences of fathers. Importantly, this differs from studies of couple-level outcomes like relationship tenure or divorce, which inherently involve both partners (Lillard and Waite, 1993; Svarer and Verner, 2008). By focusing on an individual-level outcome like RQ, we can capture how childbirth uniquely affects each parent. Moreover, the implications of RQ extend beyond parents to future outcomes of their children, who are affected by exposure to a poor-quality parental relationships (Piketty, 2003; Björklund and Sundström, 2006) and by parental divorce (Gruber, 2004; Björklund et al., 2007).

Finally, this paper contributes to our understanding of household time allocation, which is a long standing issue (Lundberg and Rose, 2000; Lundberg, 2005). This topic has gained even greater relevance during and after the COVID-19 pandemic, which introduced unprecedented disruptions to housework and childcare (Sevilla and Smith, 2020; Hupkau and Petrongolo, 2020; Alon et al., 2020; Farré et al., 2020; Del Boca et al., 2020). The dynamic impact of the birth of the first child on paid and unpaid work has previously been studied by Aguilar-Gomez et al. (2019) in Mexico and Siminski and Yetsenga (2022) in Australia. We replicate their findings for the United Kingdom and extend them to uncover heterogeneity in this impact based on household time arrangements before birth, providing the novel finding that gender-based specialization occurs after birth, irrespective of pre-birth arrangements. Furthermore, we use these findings to shed light on the impact of this reallocation in couple outcomes.

Roadmap. The rest of the paper is organized as follows. Section 2 describes the data and presents our measure of relationship quality. Section 3 describes the empirical strategy and discusses identification. Section 4 presents the main results and considers their robustness. Section 5 explores changes to household specialization as the potential mechanism at play. Section 6 concludes.

[2]The papers cited on identity and well-being do not directly address the influence of children on those outcomes, but they simply discuss their potential implications.

2 Data and Measures

The analysis is based on data from Understanding Society, a longitudinal household survey in the United Kingdom representative of the country's population (University of Essex, Institute for Social and Economic Research, 2022). Information regarding the primary outcome spans from 2009 to 2021 and is based on data collected about couples in waves 1, 5, 7, 9 and 11. We source additional data from the British Household Panel Survey (BHPS), which predates Understanding Society and covers the period from 1991 to 2008. Crucially, the BHPS allows us to identify both past and current partners, along with detailed characteristics of the relationships, including tenure.

2.1 Measure of Relationship Quality

Understanding Society conducts a 10-question survey every other data collection wave. These questions revolve around individuals' relationships with their cohabiting partners, including questions like "How often do you and your partner quarrel?". Respondents rate these questions on a scale from "All of the time" to "Never" on a six-point Likert scale. The survey also includes questions regarding relationship happiness and shared interests. Table 1 contains the full set of questions.[3] These questions are asked individually to all respondents living with a partner, regardless of marital status. This data is available in alternating waves spanning from 2009 to 2022.[4]

Table 1: Questions in the Understanding Society Partner module.

(a) Subjective assessment	(b) Couple time use
How often do you... ?	*How often do you... ?*
discuss or consider splitting	work together on a project
regret that you married or lived together	have stimulating exchanges of ideas
quarrel	calmly discuss something
get on each other's nerves	kiss your partner
What is the... ?	*Do you and your partner... ?*
degree of happiness of your relationship	engage in outside interests together

Notes: This table displays the full set of questions in the Understanding Society Partner module. This self-completion questionnaire is included in waves 1, 5, 7, 9, 11 and 13, and spans the period from 2009 to 2021. We divide these questions in (a) subjective assessments of the relationship and (b) couple time use, depending on the information they convey. The first four questions in each block are responded in terms of frequency in a six-point Likert scale, the happiness question is answered in terms of degree in a seven-point scale and the responses to the outside interest question correspond to amounts in a five-point scale. The exact phrasing of the questions is reported in Appendix A.1.

[3] This questionnaire, excluding the questions on happiness and kissing, was originally used in psychology research to construct the cohesion and satisfaction subscales of the Revised Dyadic Adjustment Scale (Busby et al., 1995). Appendix A.1 provides the exact phrasing of the questions and the response options of the questionnaire.

[4] An implementation error in wave 3, makes the item "Do you and your partner engage in outside interests together?" incompatible with the rest of the waves. Hence, we cannot construct the outcome using this wave.

We categorize the items in Table 1 in two blocks based on the information they convey: (a) subjective assessments of the relationship and (b) couple time use items. The items in the first block contain information related to the degree of happiness and conflict in the relationship, or individual perceptions that may not be shared by both partners. In the second block, the items refer to joint activities and how couples spend their time together. Including both types of items provides a nuanced overview of the relationship.

To construct the main outcome, we first transform all the items such that lower values correspond to poorer couple behaviors. Using responses from the complete dataset, we conduct a factor analysis and retain the first factor as the comprehensive measure of relationship quality (RQ). All items have positive loadings and the factor accounts for 40.61% of the variation in the data.[5] The resulting variable is standardized with a mean of zero and a standard deviation of one. In this context, higher values of RQ indicate higher quality relationships. A more in-depth discussion of the construction, validation and analysis of the RQ measure can be found in the companion paper, Rodríguez-Moro and Román (2024).

Figure 1 presents the distribution of RQ in the sample of individuals who become parents, separately for the periods (a) before and (b) after the birth of their first child. In both cases, the distribution is skewed towards the left, indicating a higher frequency of high-quality relationships. The average RQ before birth is remarkably higher than the average for the complete data and highly concentrated. After birth, instead, the mean RQ drops below the full data average. This is preliminary evidence of a decrease in RQ after having a child.

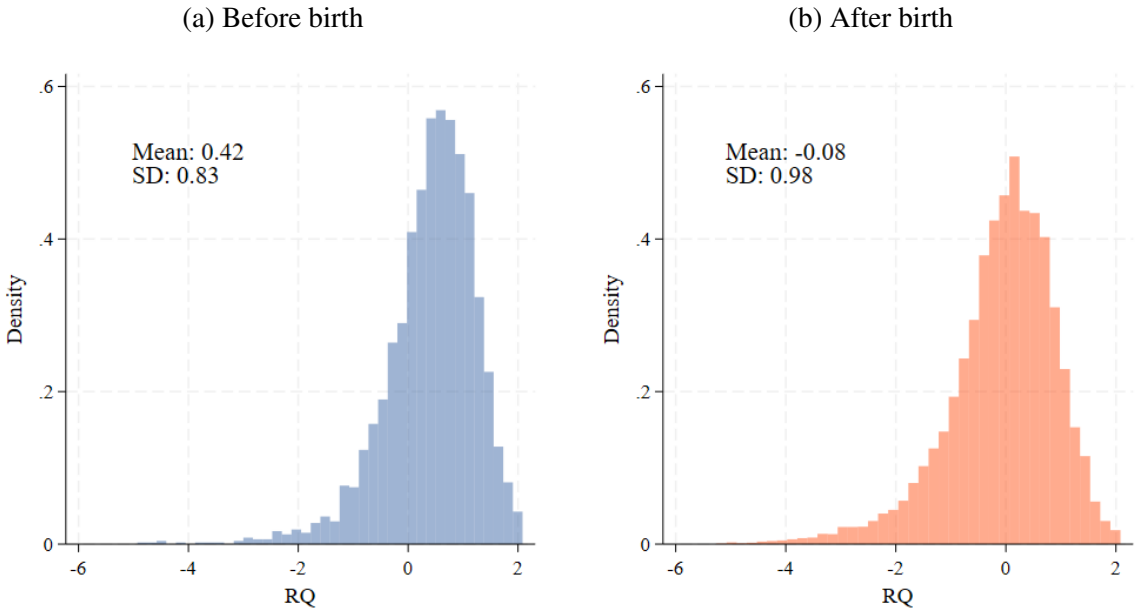
Validity of the measure. Given the novelty of the measure, in Appendix A.3 we conduct a series of tests to verify that RQ provides sensible information about the quality of a relationship. We follow the life satisfaction literature and aim to validate two fundamental theoretical assumptions: (i) informativeness and (ii) interpersonal comparability (Ferrer-i Carbonell and Frijters, 2004). The validation exercises are computed using all the available information on RQ (full data), instead of restricting to parents like in the remainder of the paper.

First, we evaluate the informativeness of RQ by assessing its capacity to predict couples' decisions. Specifically, we test whether the distribution of RQ in periods preceding marital transitions (marriage and separation) differs from the distribution in the full data. We further explore the association between RQ and couple dissolution in the regression analysis in Table A.3.[6] We find that, in the full data, a standard deviation increase in RQ is associated with a 0.9 percentage point lower probability of splitting in the subsequent period. Similarly, we compare the distribution of RQ before the birth of the first child with our complete analysis sample.

[5]The retained factor has an eigenvalue of 4.06, while the next one has an eigenvalue of 1.45. All the factor loadings are reported in Table A.1.

[6]We regress a binary variable that equals one on the period before dissolution on RQ, and control for age, sex, college education, employment status, log monthly income, presence of children, relationship tenure, marital status, area of residence and period.

Figure 1: Distribution of RQ in the sample.



Notes: This figure plots the distribution of RQ in the sample of individuals who become parents for the periods (a) before and (b) after the birth of their first child. The mean RQ in the complete data is 0 and its standard deviation is 1.

We find strong evidence that, before couple dissolution, the distribution of RQ is below the distribution for the complete data. The opposite is true for periods preceding marriage and fertility events. Hence, our results are consistent with the theoretical assumption of informativeness.

Second, we substantiate the interpersonal comparability of the measure, that is, we evaluate whether there is a degree of commonality in the concept captured by RQ. We study the correlation of responses between partners by regressing women's RQ on their partner's RQ. Our analysis reveals a high level of correlation across responses, which is robust to controlling for individual and couple characteristics. This evidence supports the notion of objectivity in the measured concept.

Additionally, we verify that RQ evolves smoothly with age and relationship tenure, and examine its correlation with observable individual and couple characteristics. We assess the stability of the measure by estimating a two-way fixed effects regression that includes age and tenure non-parametrically. This allows us to obtain the age and tenure profiles of RQ in the full data. Both profiles are largely smooth and do not present any noticeable jumps. While the age profile is an increasing and almost linear, the tenure profile of RQ displays a rapidly decrease over the first relationship years to only slow down at longer tenures. Lastly, we find that women and individuals with children consistently report lower RQ levels, whereas college educated and married individuals report higher levels of this measure.

2.2 Household specialization

Household specialization refers to how each couple member contributes to home and market production. Understanding Society provides information on the the number hours worked in the labor market and the number of hours spent on routine housework by each individual in a week.[7] Using this information, we compute the share of home and market production out of the couple total carried out by women as a proxy to household specialization. We refer to this as the *female share* of unpaid and paid hours. A 50% share of both types of work indicates no specialization, larger shares imply female specialization and lower ones male specialization.

Figure E.1 plots the distribution of the female share of paid and unpaid hours (a) before and (b) after the birth of the first child. Before birth, there exists considerable variation in household specialization. The distribution of the paid hours share is centered around 50%, whereas the distribution of unpaid work is more evenly spread, primarily above the 50% mark. However, after the birth of the first child, both distributions experience significant changes. The mass of paid hours shifts below the 50% threshold, while the share of unpaid hours becomes more concentrated above this point.

We classify couples according to the female share of paid and unpaid hours *before* the birth of their first child. We distinguish four types of couples: (i) traditional couples, where women contribute mostly to housework and men to paid work; (ii) unbalanced couples, where women take the largest share of work in both cases, (iii) egalitarian couples, where the split of both types of work is equal across couple members; and (iv) counter-traditional couples, where men take the largest share of housework.[8]

2.3 Sample restrictions and descriptives

The population of interest consists of individuals in cohabiting relationships, whether married or not, who experience parenthood. Our analysis focuses on the subset of individuals who become parents for the first time within the observation period. To ensure the robustness of our sample, we impose several criteria. First, we restrict our analysis to individuals who were aged 18 to 45 (women) or 50 (men) when they first became parents. Second, we exclude couples cohabiting with children from previous relationships. Third, we use only heterosexual couples.

[7] Specifically, the questions posed are "About how many hours do you spend on housework in an average week, such as time spent cooking, cleaning and doing the laundry?" and "Thinking about your (main) job, how many hours, excluding overtime and meal breaks, are you expected to work in a normal week?", respectively. The responses to these questions are available for approximately half of the sample of parents. Following Borra et al. (2021) we do not consider childcare part of routine housework. Time use information on childcare is unavailable.

[8] The distribution to labor market work is classified as equal if the female share of paid work lies between 45% and 60%, whereas an egalitarian distribution of housework is based on a female share of unpaid work between 45% and 55%. Female shares above those intervals are indicative of a larger contribution of women to that sort of production, and shares below that refer to a larger contribution of men.

The resulting sample is an unbalanced panel of 1,760 individuals observed up to 6 times.[9]

Table 2 describes the analysis sample in the period preceding childbirth. The left panel summarizes individual characteristics separately for fathers and mothers. Fathers are, on average, 32 years old and mothers are four years younger. Fewer mothers than fathers are either active in the labor market or employed. Mothers spend four hours less on labor market work and three and a half more hours on housework than fathers. The RQ of both is above the complete data average. The second panel gives an overview of couple characteristics. Respondents cohabit for four years before having their first child, on average. Consequently with the time use description, mothers take a share of paid work slightly below 50% and a share of housework above 60%.

Table E.1 summarizes the same characteristics separately for the different types in terms of household specialization.[10] Traditional couples are formed by less educated partners and are on average poorer than the rest of the types. Traditional couples comprise the largest difference in paid and unpaid work hours, differing by 17 and 9 hours, respectively.

Table 2: Summary statistics for the period before the birth of the first child.

(a) Individual characteristics			(b) Couple characteristics	
	(1) Fathers	(2) Mothers		(3) Couples
Age	32.00	28.39	Tenure	4.173
	(6.326)	(6.065)		(3.307)
College educated (%)	33.86	36.38	Married (%)	42.67
	(47.33)	(48.12)		(49.22)
Active in labor mkt (%)	86.92	84.26	Monthly household income	4039.4
	(33.71)	(36.40)		(2986.8)
Employed (%)	82.47	78.06	Female share of paid work	0.472
	(37.99)	(41.37)		(0.210)
Weekly work hours	31.43	27.31	Female share of housework	0.630
	(17.02)	(16.12)		(0.204)
Weekly housework hours	5.148	8.594		
	(4.045)	(6.261)		
RQ	0.353	0.388		
	(0.859)	(0.895)		
Observations	2718	3266	Observations	4134

Notes: This table presents mean values of a set of individual and couple characteristics in the sample the period before the birth of the first child. Standard deviations in parentheses.

[9] This sample captures 97% of the fertility events in the data.

[10] The reduced sample sizes in this table reflect the limited availability of responses to the time use questions and the unbalanced sample, where not all individuals can be observed on the period right before birth.

3 Empirical Strategy

Our primary objective is identifying the causal effect of first child birth on RQ. We leverage variation in the timing at which individuals have their first child and take a dynamic difference-in-differences (DiD) approach. We estimate the following two-way fixed effects (TWFE) regression model:

$$y_{i,t} = \alpha_i + \mu_t + \sum_j \mathbb{1}\{j = t - G_i\}\delta_j + u_{i,t} \tag{1}$$

where $y_{i,t}$ denotes the RQ of individual i in period t. We denote as G_i the year in which the first child is born to individual i. Thus, $t - G_i$ denotes time since i's first child was born. We refer to this as event-time. Equation 1 includes the full set of event-time dummies, as well as individual and period fixed effects. We estimate this regression using the estimator proposed by Callaway and Sant'Anna (2021).[11] We use this method to overcome the issues derived from estimating two-way fixed effects regressions through ordinary least squares (OLS) pointed out in the *new differences-in-differences* literature.[12] We cluster the standard errors at the couple level.

The causal interpretation of the estimated coefficients corresponding to the event-time dummies as average treatment effects on the treated (ATT) relies on two assumptions. First, changes in RQ do not predict when individuals have their first child (*no anticipation*).[13] This assumption would be violated if couples decided to have children in response to some periods of high RQ, for example.[14] Second, in absence of treatment (having a child) RQ would have evolved similarly regardless of the period when couples had their first child (*parallel trends*). Although we cannot directly test whether these assumptions are satisfied, we provide evidence about their plausibility in Appendix B.

First, we verify that the evolution of RQ is flat during the periods preceding the birth of the first child, which would be evidence in favor of both assumptions. However, flat pre-trends may arise from averaging out couples with varying circumstances: some may decide to have a child to address relationship issues, while others may do so because of their happiness. We provide evidence that this is not the case showing that the standard deviation of RQ is not larger during the periods preceding birth, and that average deviations from the individual-specific mean are zero before birth. Second, we show that parents who experienced interruptions in fertility, such

[11] We provide an extensive summary of this methodology in Appendix B. We implement it using the `csdid` Stata package.

[12] A large portion of the recent econometrics literature has devoted its attention to the problematic deriving from estimating TWFE regression models in a context of staggered treatment introduction and have proposed solutions to overcome the derived biases (De Chaisemartin and d'Haultfoeuille, 2020; Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021; Sun and Abraham, 2021; Borusyak et al., 2022; Gardner, 2022; Roth et al., 2023).

[13] This assumption can be relaxed to limited anticipation if the anticipation horizon can be specified. We relax this assumption in complementary analyses to allow for outcomes to react during pregnancy.

[14] Importantly, this assumption relates to *changes* in RQ before treatment, pertaining the longitudinal dimension of the data. This differs from the finding in Figure A.2 (b) that the average *level* of RQ is higher before having a child, in cross-sectional comparisons.

as those who have used fertility treatments or faced involuntary pregnancy interruptions, do not differ from those who did not. Third, our results hold when restricting comparisons to individuals with similar RQ levels before birth. Finally, we do not find evidence of heterogeneity in treatment effects for parents whose first children were born in different years.

4 Impact of Children on Relationship Quality

Figure 2 depicts the estimated event-time dummy coefficients from Equation 1, along with 95% confidence intervals. These coefficients correspond to the effect of the birth of the first child on RQ at each period relative to birth. Since RQ is a standardized measure, the estimates are interpreted in terms of standard deviations. The available information allows us to look at four periods before the birth of the first child and up to seven periods after.

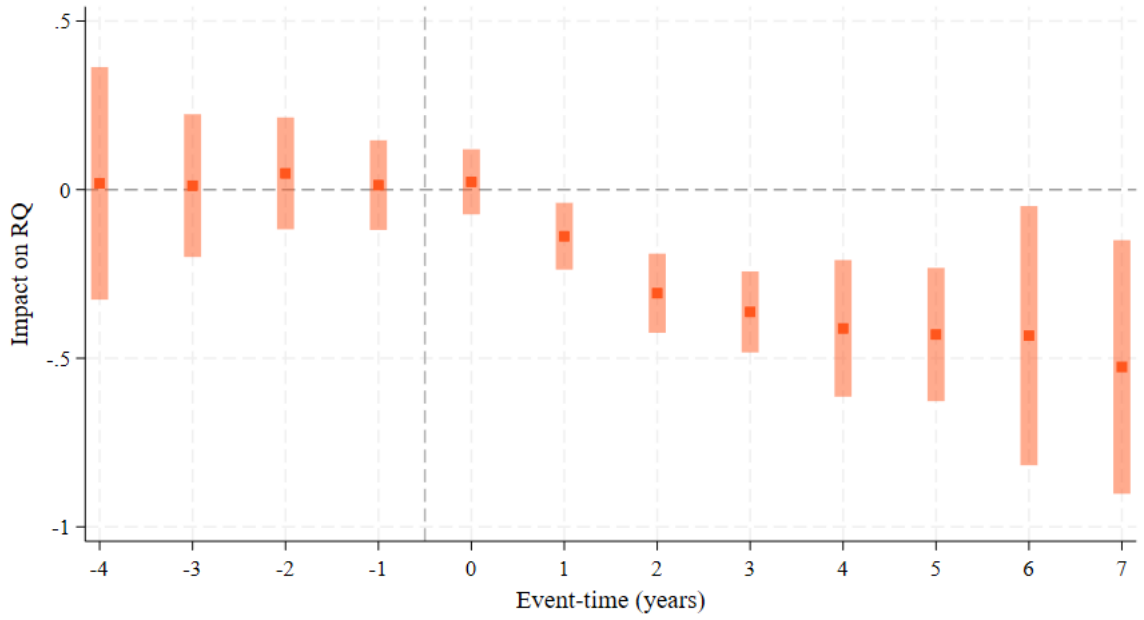
The coefficients corresponding to periods before birth are not significantly different from zero. These flat pre-trends are consistent with the assumptions that the decision to have a child is not endogenous to changes in RQ and that, in absence of treatment, RQ would have evolved in parallel for individuals having their first child in different periods. The coefficient corresponding to the period of birth is also statistically non-distinguishable from zero. This means that RQ does not immediately react to the birth of the first child.

There is a significant decrease in RQ during the first four years after birth. At this time, the decrease in RQ stabilizes at around half a standard deviation. As an illustration, individuals ranked in the 75^{th} percentile of the RQ distribution before the birth of their first child are pushed down to median RQ over this time period. This negative impact stabilizes after four years and until seven years after child birth. This suggests that, although having a child shifts RQ down, it does not change the evolution of this variable over time.

Impact on mothers and fathers. Most literature on child penalties finds strongly asymmetric effects of first child birth on mothers and fathers. Studies on labor market outcomes find that women reduce labor force participation, work fewer hours, and experience declines in wages and earnings following childbirth, while men show no such impact. This contributes to increasing within-household inequality and widens the gender wage gap (Adda et al., 2017; Blau and Kahn, 2017; Kleven et al., 2019b; Cortés and Pan, 2020; Goldin, 2021; Albanesi et al., 2023). Beyond labor market outcomes, gender differences in the impact of childbirth also emerge on time spent on housework (Aguilar-Gomez et al., 2019; Siminski and Yetsenga, 2022), leisure (Aguiar and Hurst, 2007), sleep (Costa-Font and Flèche, 2020) and even mental health (Ahammer et al., 2023).

Motivated by this, we study potential gender differences in RQ around first child birth. Despite RQ being a shared concept, as discussed in Subsection 2.1, individual perceptions

Figure 2: Dynamic effect of first child birth on RQ.



Notes: This graph plots the event-time dummy coefficients obtained from estimating Equation 1 with RQ as an outcome and using the method proposed by Callaway and Sant'Anna (2021), and the 95% confidence intervals.

can vary. However, as shown in Table 2, mothers and fathers report nearly identical levels of RQ prior to childbirth. In Figure C.1 (a), we repeat the main analysis for men and women separately. We find a comparable decline in RQ for both parents during the first three years post-birth. After three years, RQ stabilizes for men, while it declines further for women in year four. Figure C.1 (b) examines whether these differences are statistically significant, comparing RQ responses between husbands and wives. The results show no significant gender difference.

The symmetric effect of childbirth on both mothers' and fathers' RQ makes our results distinct from prior child penalty literature. While most penalties primarily affect women, RQ reflects a broader effect that involves both partners, indicating a utility loss on both sides. Notably, this decline occurs for fathers, even though they typically do not see significant changes in many outcomes traditionally studied, providing valuable insight into their experiences and contrasting with the notion that fathers are unaffected by parenthood.

Implications for couple dissolution. To better understand the relevance of these results, we compute a back-of-the-envelope calculation of the implications for couple dissolution. In Section 2 we provide evidence of the capacity of RQ to predict marital transitions, finding lower values of RQ preceding couple separation. The regression analysis in Table A.3 indicates that, in the parent sample used in the main analysis, a standard deviation increase in RQ is associated with a 0.7 percentage point lower probability of separation.

The magnitude of this coefficient highlights the relevance of RQ in dictating couple disso-

lution. Based on that, the observed half standard deviation decline in RQ following childbirth would be correlated with a 0.35 percentage point higher probability of separating. This association, while illustrative, is notable given that 2% of the existing couples in our data separate each year.[15] Furthermore, this result is in line with the findings of previous research on the impact of children on couple dissolution (Lillard and Waite, 1993; Svarer and Verner, 2008).

As argued in Browning et al. (2014), RQ is just one aspect of how the presence of children influences couple dissolution decisions. Having a child tends to raise separation costs for parents while also increasing the value of maintaining the relationship through the creation of economies of scale within the couple, which would reduce the probability of dissolution. While exploring these factors jointly is crucial given the detrimental effects that separation can have on children, studying them separately may be equally relevant, especially considering the adverse effects of being exposed to a low parental RQ.

General happiness and RQ. We next study the effect of the birth of the first child on general happiness to obtain a more comprehensive picture of the impact of children on individual utility. RQ can be interpreted as part of the utility exclusively attached to the couple, distinct from the utility derived from having a child. Therefore, the impact of childbirth on happiness and RQ would not necessarily align in the same direction.

In Figure C.3 (a) we repeat the main analysis using standardized general happiness.[16] We find that the birth of the first child induces an immediate increase in general happiness during the pregnancy and birth periods. However, the absolute magnitude of this impact is smaller than the one documented for RQ. Moreover, the impact on general happiness only lasts one period, not mirroring the persistence observed in the impact on RQ.[17]

The depicted patterns align with the hypothesis that happiness encompasses a component attached to the couple (RQ) and another attached to the child.[18] The main finding of this paper indicates that the couple component decreases after birth. Hence, the fact that the sum of both components - couple and child - does not decrease suggests that the child component is positive. Conclusively, while the event of childbirth may indirectly reduce happiness, entailing negative consequences for RQ, this evidence suggests that the direct contribution of children to

[15] Figure C.2 plots the share of couples dissolving out of all the couples observed each year in the full data and in the parent sample. According to the Office for National Statistics (2022), on average 1.44% of the married couples in fertility ages (20-45 year-olds) living in England and Wales divorce every year, during the period 2009-2021. We do not have the equivalent data for cohabiting couples.

[16] The question used is "Have you recently been feeling reasonably happy, all things considered?". We take two periods before birth as the baseline to allow for anticipation during the pregnancy period.

[17] There is a decreasing trend in general happiness after birth. Blanchflower and Oswald (2004) find that general happiness is U-shaped in age, reaching the minimum at ages 37-41 in the United Kingdom. In our sample, the average age at birth is 30, as seen in Table 2. Thus, the documented decline corresponds to a life-cycle trend that our methodology does not allow accounting for.

[18] In Table C.1, we quantify the association between happiness and RQ for individuals who have not become parents by age 50, the end of the fertility cycle. We find that a one standard deviation higher RQ is associated with a 0.2 standard deviation higher happiness measure, on average.

happiness is not negative.

Other measures and shocks. We compare the magnitude and persistence of our main finding with the effect of this event on other subjective outcomes at the individual level. Figure C.3 (b) displays the results for subjective well-being.[19] This effect, although also negative, only amounts to a tenth of the magnitude of our main finding, and dissipates one period after birth. Ahammer et al. (2023) use a similar approach to study the impact on mental health and find a significant increase in antidepressant prescriptions for the sub-population of individuals that visit the psychiatrist. They find a greater effect on mothers, which they attribute to the higher childcare burden borne by them.

Finally, we examine the significance of the birth of the first child as a distinct shock to RQ by comparing it to another major life event: becoming unemployed. Previous studies have found a positive correlation between male unemployment and divorce probabilities (Jensen and Smith, 1990; Charles and Stephens, 2004; Eliason, 2012; Doiron and Mendolia, 2012). Figure C.4 displays the results obtained from computing our main estimation around unemployment events on RQ.[20] Unlike childbirth, unemployment is not significantly associated with changes in RQ, which contrasts with the findings on divorce.

These insights can inform models of family formation and dissolution in multiple ways. As mentioned earlier, RQ is defined as a measure of the non-material benefits of being in a relationship. The difference between the association of unemployment with RQ and with divorce may indicate that, while the job loss of a partner reduces the material gains from being in a relationship with them, it may not have such impact on non-material gains. Furthermore, our main finding evidences the smooth and path-dependent nature of RQ, which could guide the modelling of the distribution of this factor in these family formation and dissolution models.

4.1 Robustness

In Appendix D we address four potential concerns that could affect the interpretation of our results. First, we explore whether the impact on RQ is driven by specific items within the measure. This could result from an uneven response of items across different blocks, or from a change in how the items are valued after the birth of a child. Second, we consider whether the persistent impact on RQ is solely due to subsequent children or if it represents a lasting effect of the first child's arrival. Third, we take into account the limitation that RQ cannot be observed

[19] The subjective well-being measure combines the responses to the 12 General Health Questionnaire, related to physical and mental health. Some of the questions relate to sleep and being under strain, which are mechanically affected by having a child. Previous literature has discussed the association between children and well-being, although not in a causal way (Dolan et al., 2008; Clark et al., 2008; Blanchflower, 2009; Ferrer-i Carbonell, 2013).

[20] We focus on the first instance of unemployment for an individual within our observation period and exclusively consider individuals who have experienced unemployment as our control group. We do not make any causal claims about these results.

after a couple dissolves, which may introduce some selection bias that could distort our results. Fourth, we examine whether the timing of child birth, in terms of the age of the parents and the duration of the relationship, has any influence on our findings.

Time invariance of RQ. Our first concern is whether the decrease in RQ might be solely attributed to specific items within the measure, particularly those related to couple time use. This concern arises from the fact that, after the birth of the first child, the available time for couples to spend together decreases significantly due to the new responsibilities that occupy a significant portion of their time. Given that individuals highly value spending leisure time with their partners (Georges-Kot et al., 2024), this decrease could directly reduce RQ, potentially being the only driver of the documented decrease.

To explore this, we repeat the factor analysis excluding the time use items and using solely the items containing subjective assessments of the relationship. Figure D.1 (a) plots the results of conducting our main analysis on the resulting measure. The documented decrease is much more pronounced than what we observed on RQ. For completeness, we repeat this procedure using only the time use items in Figure D.1 (b). In this case, the impact is less pronounced and not statistically different from zero after three years. This suggests that our primary finding is not only driven by reduced time together, but is mainly influenced by changes in how individuals assess their relationship.[21]

However, the possibility remains that individuals change how they value different items within the RQ measure after becoming parents. To address this, we construct a new RQ measure using data solely from individuals who are already parents. We use those observations to obtain the factor loadings in the factor analysis and construct the measure for the entire sample. Figure D.2 displays the results of our main analysis using this measure, showing that our primary result remains consistent.

Subsequent fertility. Our analysis focuses exclusively on the birth of the first child. It is unclear whether the persisting impact on RQ represents a lasting effect of the first child's arrival or if it is merely a consequence of subsequent children. In Figure D.5 we split the analysis sample in two groups: (a) individuals with only one child or (b) those with subsequent children. About 45% of the individuals in our sample are observed having an only child. The impact occurs slightly faster and is somewhat more pronounced for individuals with more than one child, but they experience a recovery four years after the first child's birth. Thus, we cannot attribute the documented persistence on RQ to parents of more than one child.

However, this analysis masks censorship in the sample, as many individuals have not

[21] Figure D.3 and Figure D.4 display the impact of first child birth on each subjective assessment and couple time use item separately. The item responses are standardized for comparability. The degree of happiness with the couple experiences the largest impact.

reached the end of their fertility cycle at the last observation. We restrict the analysis to individuals observed at age 40 or older, which better approximates their lifetime fertility. This restriction remarkably reduces the sample and does not allow for a dynamic analysis. Table D.1 displays the static DiD results obtained through the usual estimation method. Note that the obtained ATT is an average of the impact all periods after, where periods with more observations get higher weights (Callaway and Sant'Anna, 2021). The results suggest that subsequent children do not have an additional negative impact on parental RQ. In fact, they may somewhat offset the initial decline in RQ following the birth of the first child.

It is important to note, however, that this could be a selected sample, as couples who choose to have a second child may have had a different experience compared to those who did not. We focus exclusively on second-time parents to examine whether the birth of a second child has any impact on RQ. Figure D.6 shows the results of the main analysis around the birth of the second child. Although we observe a pre-trend, likely reflecting the decline in RQ following first child birth, there is no significant effect associated to the birth of the second child.

Selected sample. Due to the nature of RQ, it cannot be observed once a couple dissolves, potentially introducing selection bias that could distort our results. We investigate the attrition from the sample due to separation, by estimating the impact of birth on the likelihood of separation. Note that, by construction of the sample, we do not observe couples dissolving before the first child is conceived. Thus, the results should not be interpreted as causal. Figure D.7 (a) indicates that dissolution probabilities are around 2% over the periods analyzed after birth, suggesting that the attrition from the sample is not a significant concern. We repeat the analysis excluding from the sample couples who separate over the observation period, as shown in Figure D.7 (b). These results largely mirror our primary findings, alleviating concerns about selection bias.

Age and tenure at birth. In addition to the birth of the first child, the timing of this event may significantly affect RQ (Kleven et al., 2019b). We explore this by separately examining individuals who had children at different ages and relationship tenures. Figure D.8 presents the average RQ at each event-time period by (a) age bin and (b) tenure bin, normalizing RQ to zero in the period before birth. While baseline RQ levels differ for individuals who had children at different ages, the post-birth evolution of RQ remains consistent across groups.

Baseline levels also differ across tenure groups. In this case, individuals who had their first child at the earliest stages of their relationship experience a slightly greater decrease in RQ . However, this more substantial decrease could be partially attributed to the general trend of RQ decreasing during the early stages of a relationship, as seen in Figure A.4.

To ensure the robustness of our main findings, we include these variables as controls, as presented in Figure D.9. Note that Callaway and Sant'Anna (2021) does not allow to study the

dynamic impact of age and tenure on RQ, since it only allows to control for pre-birth levels of covariates. The differences for all sets of controls remain minimal until six years after the birth of the first child. At that point the coefficients of the specification including both age and tenure become larger. Thus, our main results could be interpreted as a lower bound.

To account for potential dynamic effects across age and tenure, we employ the event-study analysis as popularized by Kleven et al. (2019b). This method involves estimating the impact through OLS while substituting individual fixed effects with the full set of age and relationship tenure dummies. The results displayed in Figure D.10 indicate that our main finding is sustained even accounting for the dynamic evolution of RQ with age and tenure.

Additional checks. We test the robustness of our findings in other ways. First, we use an alternative control group to the one used so far. Instead of using individuals who will become parents but have not yet, we consider those who never become parents. We classify individuals as never parents if they have not had children by age 50, the end of the fertility cycle. This group was initially excluded due to concerns about their comparability to future parents - the identification assumptions required appeared to be more plausible within the parent sample. However, as shown in Figure D.11, the results from this alternative specification are remarkably similar to the main findings. The estimates are slightly larger in magnitude, with improved precision.

Second, we analyze heterogeneity in the results by child sex. Previous literature has found differences in the time use and the gender norm attitudes of parents after the birth of their first child (Lundberg, 2005; Grinza et al., 2017). We document that there are no such differences in the impact on RQ in Figure D.12.

5 Mechanism: Changes in Household Specialization

The birth of a child introduces a significant and unprecedented shift in how couples allocate their time. New responsibilities related to childcare emerge, and the demand for routine housework substantially increases. Couples must adapt to these additional tasks by redistributing the time each member allocates to paid work in the labor market and unpaid work within the household. This division of responsibilities dictates the roles that each partner assumes within the relationship, such as the primary earner or the caregiver. Research shows that men's labor market outcomes remain largely unaffected by the arrival of a child, while women significantly reduce their labor force participation and working hours (e.g., Kleven et al., 2019b; Goldin, 2021) and take on longer home production times (Aguilar-Gomez et al., 2019; Siminski and Yetsenga, 2022).

Kuziemko et al. (2018) show that the magnitude of these changes is often unforeseen. Women, particularly high educated ones, underestimate the impact of having a child on their labor market outcomes and do not envision themselves as full-time housewives after motherhood. We argue that the unanticipated, structural changes in the division of paid and unpaid work times within the couple may be mediating the deterioration in relationship quality that follows the birth of the first child.

To investigate this, we first study how household specialization changes following birth, in terms of the contribution of each partner to paid and unpaid work. We then examine how the documented changes in household specialization translate into changes in RQ, categorizing couples based on their initial distribution of tasks and assessing whether the effects of childbirth vary depending on those couple categories.

5.1 Impact of children on household specialization

We start by documenting how mothers and fathers adjust their time use in response to the arrival of a child, focusing on the time spent by each partner on paid employment and unpaid housework, as defined in Section 2.[22] Figure 3 (a) plots the findings on paid work and Figure 3 (b) does it for unpaid work.

The impact of the birth of the first child on working time substantially differs for mothers and fathers. As shown in Table 2, fathers spend about 31 hours in paid work weekly before birth, while mothers work roughly four hours less. There is a steep reduction in the number of hours worked by mothers immediately after birth, reaching and stabilizing at a 14 hour reduction in two years. In contrast, fathers display only a minimal decrease in their working hours later in the child's life, which could be the product of a life-cycle trend unaccounted for in our analysis.

Turning our attention to housework time, one year after the birth of their first child mothers increase their weekly unpaid work by about five hours in comparison to their pre-birth baseline of eight hours. This increase is sustained for about four years, stabilizing around seven hours above the baseline. Fathers increase their housework time from five to six weekly hours. Although small, this increase is statistically significant. Our findings are consistent with previous results on the impact of having children on labor market outcomes (Kleven, 2022) and home production (Aguilar-Gomez et al., 2019; Siminski and Yetsenga, 2022).

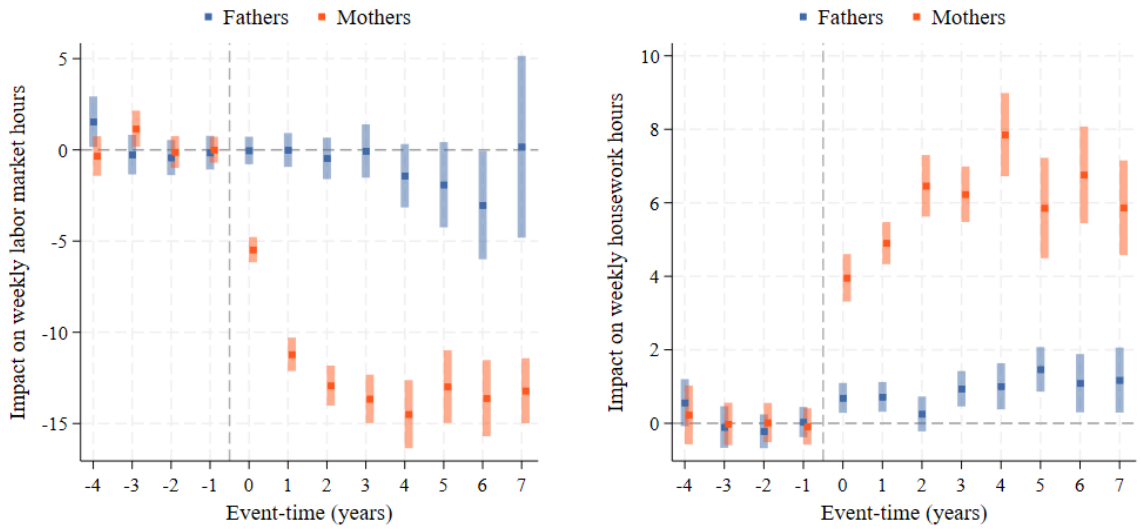
Overall, following the birth of their first child, women give up one and a half days of full-time work and take on nearly a full day of unpaid work weekly. Conversely, men increase

[22] We acknowledge the trends of employment and home production over the life cycle and include age as a control in Equation 1 to account for any potential bias. The causal interpretation of these results is subject to the identifying assumptions outlined in Section 3, with the nuance that the parallel trends assumption needed is conditional on age. However, the Callaway and Sant'Anna (2021) method can only account for baseline levels of this variable, overlooking non-linear trends. Thus, we may still observe trends attributable to the life-cycle.

their housework time by one hour and maintain their labor market hours.[23] These findings indicate that women assume most of the additional housework generated by children, effectively changing how they use their time. Furthermore, this new allocation alters the established roles within the couple, widening gender gaps in home and labor market duties, and assigning new production responsibilities to each partner.

Figure 3: Impact of first child birth on paid and unpaid hours.

(a) Weekly labor market hours (b) Weekly housework hours



Notes: This figure plots the impact of first child birth on weekly (a) labor market work and (b) housework hours separately for men and women. We estimate Equation 1 using unpaid and paid hours as outcomes and including age as a control to partially out any potential life-cycle bias. We use the estimator proposed by Callaway and Sant'Anna (2021). The plotted coefficients are the effects on each lead item and lag around the event. The 95% confidence intervals are plotted.

To study changes in roles and household specialization, we look at the evolution of the relative contribution of women to market and home production. To that end, we compute the female share of the total paid and unpaid work hours within the couple and use the pre-birth values of these two shares to classify couples into four categories, as outlined in Section 2: traditional, unbalanced, egalitarian and counter-traditional. The pre-birth values of the female shares provide information about the roles previously adopted by each partner and the existing degree of household specialization, which could be indicative of comparative advantages within the couple.

We investigate the association between the birth of the first child and these shares for each type of couple separately. Figure 4 (a) plots the average female share of paid work at each period relative to birth. The contribution of women to paid work decreases after birth in all types of couples, accounting for less than half of the household total. Unbalanced couples experience the largest changes, with the female share decreasing by almost 15 percentage points. The con-

[23] Andrew et al. (2021) document that mothers also spend more time on childcare than fathers do.

tribution of women to paid work stabilizes at similar levels in egalitarian and counter-traditional couples, at around 35% of the household total, representing a decrease of approximately 10 percentage points. Traditional couples are the least affected by parenthood in this dimension, although the female share of paid work remains low compared to the other groups.

Figure 4 (b) displays the same figure for the female share of unpaid hours. Regardless of the initial housework allocation, women in all couples assume a larger share of unpaid work than men do after birth. Counter-traditional and egalitarian couples experience the largest changes, increasing the female shares by approximately 20 and 15 percentage points, respectively. Women's contribution accounts for more than 60% in these couples. In contrast, traditional and unbalanced couples, where the female share of housework already constitutes more than three quarters of the total before birth, do not experience significant changes.[24]

This evidence suggests that couples of all types adjust their degree of household specialization after childbirth, transitioning towards more traditional, gender-based roles. However, the specific changes they undergo to establish these new roles and division of duties vary considerably. Traditional couples experience the smallest changes, sustaining the predominant roles of men in the labor market and of women in home production. Unbalanced couples do not change the distribution of housework, but the predominant role in the labor market is transferred from mothers to fathers. Notably, women in unbalanced couples are the only ones reducing their total time contribution after birth. Egalitarian couples experience substantial changes in the division of both types of work, abandoning the equal distribution of tasks. After birth, women adopt a dominant role at home whereas men specialize in the labor market. Finally, counter-traditional couples experience the largest changes in the division of housework. Men abandon their predominant role in this sphere, which is taken on by women, but preserve or strengthen their role in the labor market.

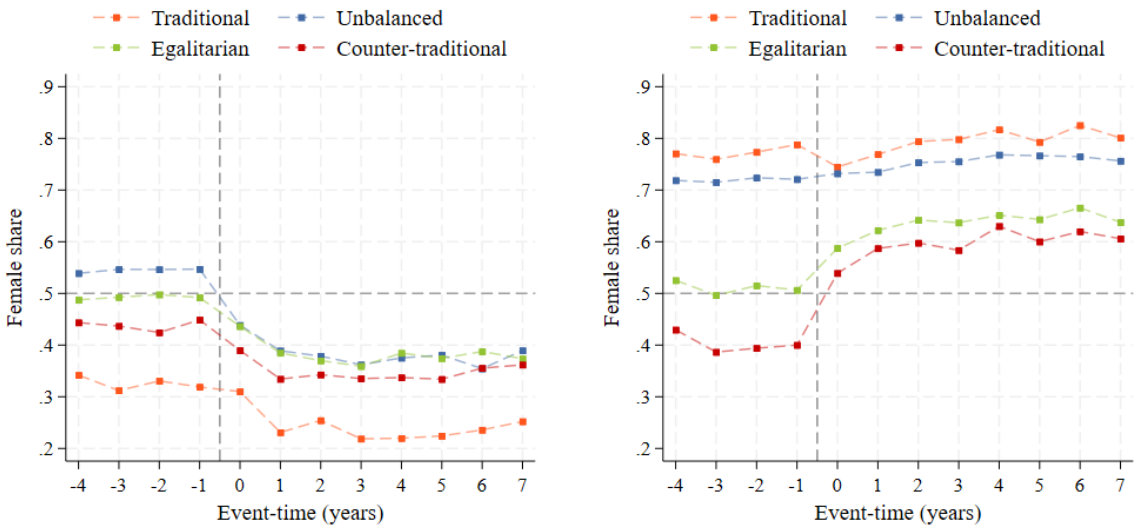
In traditional models of the household, paid and unpaid work are assumed to be allocated among household members based on the degree of substitutability between the labor and domestic inputs of the different members (Becker, 1991). Our findings challenge the plausibility of this assumption. If the distribution of tasks before the birth of the first child is indicative of potential comparative advantages within the couple, then the exchange of responsibilities documented after is not behaving according to the degree of skill substitutability. Instead, our results suggest the presence either of frictions in the labor market, for instance in the shape of statistical discrimination against mothers (Petit, 2007; Becker, 2010) or of identity considerations of individuals (Akerlof and Kranton, 2000; Ichino et al., 2019; Farré et al., 2020). This evidence suggests that such factors become prevalent with parenthood, confirming the difficulties that individuals may face to anticipate changes in roles and in household specialization (Kuziemko et al., 2018). We argue that this mismatch in expectations might affect RQ.

[24] In Figure E.2, we examine changes in labor market work and housework hours by couple type and gender, confirming that the shifts in female shares are driven by changes in women's time use.

Figure 4: Impact of first child birth on female time shares.

(a) Paid work share

(b) Unpaid work share



Notes: These graphs plot the average female shares of (a) paid labor market work and (b) unpaid housework time at each time around the birth of the first child by couple type.

5.2 Household specialization and relationship quality

Leveraging the heterogeneity in the reallocation of market and home tasks uncovered in Figure 4, we study how changes in household specialization induced by the birth of the first child may mediate the effects on RQ. To do so, we repeat the main analysis on the impact of child birth on RQ separately for each couple type. The reduced sample size of each group separately does not allow us to carry out a dynamic analysis as stated in Equation 1. Thus, we adopt a static approach and estimate the following regression for each couple type:

$$y_{i,t} = \alpha_i + \mu_t + \delta D_{i,t} + u_{i,t} \tag{2}$$

where $D_{i,t}$ is a binary variable equal to one if individual i already had a child at time t, and α_i and μ_t denote the full set of individual and time dummies.

We estimate this regression using the estimator proposed by Callaway and Sant'Anna (2021). The causal interpretation of δ in a static setting requires the additional assumption that treatment effects are homogeneous with time relative to birth.[25] The results in Figure 2 evidenced heterogeneity in treatment effects with time relative to childbirth, meaning that this assumption is not plausible in our context. Due to this and since we cannot disentangle the impact of the birth of the first child from unobserved characteristics of each couple type, we do not make

[25] As explained in Appendix B, the static aggregation constitutes a weighted average of the treatment effects for different treatment cohorts at different times relative to treatment. These weights are assigned based on the number of observations in each time bin, with more populous bins receiving higher weights. Hence, without this assumption, the estimates would be skewed towards the event-times with a larger number of observations.

causal claims about these results.

In Table E.1 we present the baseline levels of RQ for each type of couple. All couples report RQ values above the average in the full dataset. The highest pre-birth RQ levels are observed in egalitarian couples, followed by counter-traditional ones. The lowest levels of RQ are found in unbalanced couples, although they are quite similar to those of traditional couples. The greatest difference in RQ between partners is observed in traditional couples, with men reporting lower values.[26] Table 3 contains the results from estimating Equation 2 separately for each couple type. All the estimated coefficients are negative, meaning that the negative impact of first child's birth is sustained across couple types. However, the coefficient is only significantly different from zero for individuals in egalitarian and counter-traditional couples.[27] As seen in Figure 4, these are also the couples experiencing the largest changes in household specialization, where women start adopting primary roles in housework and men specialize in paid work after birth. The smallest coefficient corresponds to individuals in unbalanced couples. Although these couples experience the largest drop in the female share of paid hours, the distribution of housework remains largely constant. Furthermore, overall time is more evenly distributed between partners after birth.

Table 3: Impact of first child birth on RQ by couple type.

	(1) Traditional	(2) Unbalanced	(3) Egalitarian	(4) Counter-traditional
ATT	-0.107	-0.0992	-0.175*	-0.243**
	(0.180)	(0.086)	(0.069)	(0.075)
Observations	273	876	611	856

Notes: This table reports the results from estimating Equation 2 using Callaway and Sant'Anna (2021) separately for each couple type. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

There are two factors that could jointly explain these results. First, RQ might be sensitive to unequal distributions of the total household production time. Unbalanced couples, where women shoulder the bulk of labor and home responsibilities before birth, are the only ones distributing overall time more evenly between partners after they become parents. Although the time rearrangements they undergo are substantial, the increased equality in the contribution to household production may be attenuating the associated tensions and the decrease in relationship quality.

Second, the redistribution of housework may be the key factor triggering the deterioration of couples' relationships, by being less foreseeable than changes in labor market time. While

[26] Equation 2 assumes that the impact of first child birth on RQ is linear for all types of couples, and hence it does not depend on the pre-birth values of RQ.

[27] Note that the event-time periods with a larger number of observations are those closest to the time of birth. As seen in Figure 2, the dynamic impact of children on RQ is increasing with time since birth. Thus, these estimates should be considered a lower bound to the full impact.

labor interruptions and protecting mothers in the labor market are a central concern of the most common fertility policies, changes in home production are arguably less salient. Thus, the unforeseen and unequal distribution of housework may have a dominant role in decreasing relationship quality.

6 Concluding Remarks

This paper investigates the causal relation between having children and RQ, a proxy of the non-material gains from being in a relationship. The arrival of children triggers a series of decisions related to time allocation in both the labor market and at home. These decisions have implications not only for individual well-being, but also for the quality of the couple's relationship. While having children can lead to increased overall happiness, it does not necessarily guarantee happiness in every aspect of one's life.

We find that having a child significantly reduces RQ, and that this effect endures over time. In other words, the level of RQ is shifted downward after the birth of a child, maintaining the pre-existing life-cycle trend. This impact on RQ is substantial and persistent compared to other life events, such as unemployment, and other well-being indicators, like general happiness.

The paper posits that the arrival of children increases the demands on couples, including childcare responsibilities and additional housework. Consistent with existing literature, we find that women largely bear the increased housework at the expense of their participation in the labor market, a consequence not typically foreseen by mothers before childbirth. We exploit heterogeneity in changes to household specialization to explore the extent to which these changes mediate the impact of child birth on RQ. The results indicate that couples undergoing the most significant changes in household specialization also experience the greatest declines in RQ.

These findings open avenues for further research on the consequences of having children. From a policy perspective, several measures addressing the impact of children on various outcomes have been explored. Policies like paternity leave, shared parental leaves, and childcare provisions have been shown to enhance female labor market outcomes, potentially by redistributing household tasks and responsibilities more equitably among couple members. Our findings suggest that such policies may mitigate the negative effects of children on RQ, an outcome closely tied to couple dissolution and the decisions surrounding investments in children, which ultimately influence their future well-being.

Moreover, this paper introduces a novel factor for consideration in the study of decisions to have children. As noted earlier, the impact on relationship quality is, to some extent, unforeseen when deciding to have the first child. Nevertheless, the observed decline in RQ and the associated utility losses, despite potential utility gains from the child itself, may become

a significant factor in deciding to have subsequent children. This factor could contribute to explaining the decline in fertility observed on the intensive margin.

References

Adda, J., Dustmann, C., and Stevens, K. (2017). The career costs of children. *Journal of Political Economy*, 125(2):293–337. 4, 12

Aguiar, M. and Hurst, E. (2007). Measuring trends in leisure: The allocation of time over five decades. *The Quarterly Journal of Economics*, 122(3):969–1006. 1, 12

Aguilar-Gomez, S., Arceo-Gomez, E., and De la Cruz Toledo, E. (2019). Inside the black box of child penalties: Unpaid work and household structure. *Available at SSRN 3497089*. 1, 5, 12, 18, 19

Ahammer, A., Glogowsky, U., Halla, M., and Hener, T. (2023). The parenthood penalty in mental health: Evidence from Austria and Denmark. 5, 12, 15

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753. 3, 5, 21

Akerlof, G. A. and Kranton, R. E. (2010). *Identity economics: How our identities shape our work, wages, and well-being*. Princeton University Press. 5

Albanesi, S., Olivetti, C., and Petrongolo, B. (2023). Families, labor markets, and policy. In *Handbook of the Economics of the Family*, volume 1, pages 255–326. Elsevier. 5, 12

Alon, T., Doepke, M., Olmstead-Rumsey, J., and Tertilt, M. (2020). The impact of COVID-19 on gender equality. Technical report, National Bureau of Economic Research. 5

Amato, P. R. and Booth, A. (2001). The legacy of parents' marital discord: consequences for children's marital quality. *Journal of personality and social psychology*, 81(4):627. 4

Amato, P. R. and Patterson, S. E. (2017). The intergenerational transmission of union instability in early adulthood. *Journal of Marriage and Family*, 79(3):723–738. 4

Andrew, A., Bandiera, O., Dias, M. C., and Landais, C. (2021). The careers and time use of mothers and fathers. *Institute for Fiscal Studies. Available at: https://ifs. org. uk/publications/careers-and-time-use-mothers-and-fathers*. 20

Antler, Y., Bird, D., and Fershtman, D. (2022). *Learning in the marriage market: The economics of dating*. Centre for Economic Policy Research. 4

Avdic, D. and Karimi, A. (2018). Modern family? paternity leave and marital stability. *American Economic Journal: Applied Economics*, 10(4):283–307. 3

Becker, G. S. (1991). *A treatise on the family: Enlarged edition*. Harvard university press. 21

Becker, G. S. (2010). *The economics of discrimination*. University of Chicago press. 21

Bertrand, M. (2020). Gender in the twenty-first century. In *AEA Papers and Proceedings*, volume 110, pages 1–24. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203. 4

Bertrand, M., Kamenica, E., and Pan, J. (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics*, 130(2):571–614. 4, 5

Björklund, A., Ginther, D. K., and Sundström, M. (2007). Family structure and child outcomes in the USA and Sweden. *Journal of Population Economics*, 20:183–201. 1, 5

Björklund, A. and Sundström, M. (2006). Parental separation and children's educational attainment: A siblings analysis on Swedish register data. *Economica*, 73(292):605–624. 1, 5

Blanchflower, D. G. (2009). International evidence on well-being. In *Measuring the subjective well-being of nations: National accounts of time use and well-being*, pages 155–226. University of Chicago Press. 5, 15

Blanchflower, D. G. and Oswald, A. J. (2004). Well-being over time in Britain and the USA. *Journal of public economics*, 88(7-8):1359–1386. 14

Blasutto, F., Kozlov, E., et al. (2020). (Changing) Marriage and cohabitation patterns in the US: Do divorce laws matter? Technical report, Job Market Papers. 4

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865. 4, 12

Bögl, S., Moshfegh, J., Persson, P., and Polyakova, M. (2024). The economics of infertility: Evidence from reproductive medicine. Technical report, National Bureau of Economic Research. 48

Borra, C., Browning, M., and Sevilla, A. (2021). Marriage and housework. *Oxford Economic Papers*, 73(2):479–508. 9

Borusyak, K., Jaravel, X., and Spiess, J. (2022). Revisiting event study designs: Robust and efficient estimation. *Available at SSRN 2826228*. 11

Brien, M. J., Lillard, L. A., and Stern, S. (2006). Cohabitation, marriage, and divorce in a model of match quality. *International Economic Review*, 47(2):451–494. 4

Browning, M., Chiappori, P.-A., and Weiss, Y. (2014). *Economics of the Family*. Cambridge University Press. 1, 4, 14

Bruze, G., Svarer, M., and Weiss, Y. (2015). The dynamics of marriage and divorce. *Journal of Labor Economics*, 33(1):123–170. 4

Busby, D., Christensen, C., Crane, D., and Larson, J. (1995). A Revision of the Dyadic Adjustment Scale for Use with Distressed and Nondistressed Couples: Construct Hierarchy and Multidimensional Scales. *Journal of Marital and Family Therapy*, 21:289–308. 4, 6

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230. 2, 11, 13, 17, 19, 20, 22, 23, 45, 46, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 61

Carlson, M. J. and VanOrman, A. G. (2017). Trajectories of relationship supportiveness after childbirth: Does marriage matter? *Social Science Research*, 66:102–117. 4

Cavapozzi, D., Francesconi, M., and Nicoletti, C. (2021). The impact of gender role norms on mothers' labor supply. *Journal of Economic Behavior & Organization*, 186:113–134. 4

Charles, K. K. and Stephens, Jr, M. (2004). Job displacement, disability, and divorce. *Journal of Labor Economics*, 22(2):489–522. 15

Chiappori, P.-A. (2020). The theory and empirics of the marriage market. *Annual Review of Economics*, 12:547–578. 1, 4

Chiappori, P.-A. and Mazzocco, M. (2017). Static and intertemporal household decisions. *Journal of Economic Literature*, 55(3):985–1045. 4

Chiappori, P.-A., Radchenko, N., and Salanié, B. (2018). Divorce and the duality of marital payoff. *Review of Economics of the Household*, 16:833–858. 1, 2, 4

Chiappori, P.-A. and Weiss, Y. (2006). Divorce, remarriage, and welfare: A general equilibrium approach. *Journal of the European Economic Association*, 4(2-3):415–426. 1, 4

Chiappori, P.-A. and Weiss, Y. (2007). Divorce, remarriage, and child support. *Journal of Labor Economics*, 25(1):37–74. 1, 4

Clark, A. E., Frijters, P., and Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature*, 46(1):95–144. 5, 15

Cortés, P. and Pan, J. (2020). Children and the remaining gender gaps in the labor market. Technical report, National Bureau of Economic Research. 4, 12

Costa-Font, J. and Flèche, S. (2020). Child sleep and mother labour market outcomes. *Journal of Health Economics*, 69:102258. 1, 12

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996. 11

Del Boca, D., Oggero, N., Profeta, P., and Rossi, M. (2020). Women's and men's work, housework and childcare, before and during COVID-19. *Review of Economics of the Household*, 18:1001–1017. 5

Doiron, D. and Mendolia, S. (2012). The impact of job loss on family dissolution. *Journal of Population Economics*, 25:367–398. 15

Dolan, P., Peasgood, T., and White, M. (2008). Do we really know what makes us happy? a review of the economic literature on the factors associated with subjective well-being. *Journal of economic psychology*, 29(1):94–122. 5, 15

Eckstein, Z., Keane, M., and Lifshitz, O. (2019). Career and family decisions: Cohorts born 1935–1975. *Econometrica*, 87(1):217–253. 4, 41

Eliason, M. (2012). Lost jobs, broken marriages. *Journal of Population Economics*, 25(4):1365–1397. 15

Farré, L., Fawaz, Y., González, L., and Graves, J. (2020). How the COVID-19 lockdown affected gender inequality in paid and unpaid work in Spain. 5, 21

Farré, L. and González, L. (2019). Does paternity leave reduce fertility? *Journal of Public Economics*, 172:52–66. 3

Ferrer-i Carbonell, A. (2013). Happiness economics. *SERIEs*, 4(1):35–60. 5, 15

Ferrer-i Carbonell, A. and Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? *The economic journal*, 114(497):641–659. 7, 37

Gardner, J. (2022). Two-stage differences in differences. *arXiv preprint arXiv:2207.05943*. 11

Gemici, A. and Laufer, S. (2011). Marriage and cohabitation. *New York University, mimeo*. 4

Georges-Kot, S., Goux, D., and Maurin, E. (2024). The value of leisure synchronization. *American Economic Journal: Applied Economics*, 16(1):351–376. 16

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119. 4

Goldin, C. (2021). *Career and Family: Women's Century-Long Journey toward Equity*. Princeton University Press. 1, 3, 4, 5, 12, 18

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277. 11, 45

Goussé, M., Jacquemet, N., and Robin, J.-M. (2017). Marriage, labor supply, and home production. *Econometrica*, 85(6):1873–1919. 4

Greenwood, J., Guner, N., and Vandenbroucke, G. (2017). Family economics writ large. *Journal of Economic Literature*, 55(4):1346–1434. 4

Grinza, E., Devicienti, F., Rossi, M. C., and Vannoni, D. (2017). How entry into parenthood shapes gender role attitudes: New evidence from longitudinal UK data. 18

Gruber, J. (2004). Is making divorce easier bad for children? the long-run implications of unilateral divorce. *Journal of Labor Economics*, 22(4):799–833. 1, 5

Hassebrauck, M. and Fehr, B. (2002). Dimensions of relationship quality. *Personal relationships*, 9(3):253–270. 4

Hupkau, C. and Petrongolo, B. (2020). Work, care and gender during the COVID-19 crisis. *Fiscal Studies*, 41(3):623–651. 5

Ichino, A., Olsson, M., Petrongolo, B., and Thoursie, P. S. (2019). Economic incentives, home production and gender identity norms. 3, 21

Jensen, P. and Smith, N. (1990). Unemployment and marital dissolution. *Journal of Population Economics*, 3(3):215–229. 15

Joel, S., Eastwick, P. W., Allison, C. J., Arriaga, X. B., Baker, Z. G., Bar-Kalifa, E., Bergeron, S., Birnbaum, G. E., Brock, R. L., Brumbaugh, C. C., et al. (2020). Machine learning uncovers the most robust self-report predictors of relationship quality across 43 longitudinal couples studies. *Proceedings of the National Academy of Sciences*, 117(32):19061–19071. 4

Kleven, H. (2022). The geography of child penalties and gender norms: Evidence from the United States. Technical report, National Bureau of Economic Research. 4, 19

Kleven, H., Landais, C., and Leite-Mariante, G. (2023). The child penalty atlas. Technical report, National Bureau of Economic Research. 5

Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimüller, J. (2019a). Child penalties across countries: Evidence and explanations. In *AEA Papers and Proceedings*, volume 109, pages 122–126. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203. 5

Kleven, H., Landais, C., and Søgaard, J. E. (2019b). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209. 4, 12, 17, 18, 60

Kuziemko, I., Pan, J., Shen, J., and Washington, E. (2018). The mommy effect: Do women anticipate the employment effects of motherhood? Technical report, National Bureau of Economic Research. 3, 18, 21

Lillard, L. A. and Waite, L. J. (1993). A joint model of marital childbearing and marital disruption. *Demography*, 30(4):653–681. 5, 14

Low, C. (2024). The human capital–reproductive capital trade-off in marriage market matching. *Journal of Political Economy*, 132(2):552–576. 4, 41

Low, H., Meghir, C., Pistaferri, L., and Voena, A. (2018). Marriage, labor supply and the dynamics of the social safety net. Technical report, National Bureau of Economic Research. 4

Lundberg, S. and Rose, E. (2000). Parenthood and the earnings of married men and women. *Labour Economics*, 7(6):689–710. 5

Lundberg, S. J. (2005). The division of labor by new parents: does child gender matter? 5, 18

Norton, R. (1983). Measuring marital quality: A critical look at the dependent variable. *Journal of Marriage and the Family*, pages 141–151. 4

Office for National Statistics (2022). Divorces in England and Wales: 2021. 14, 50

Olivetti, C. and Petrongolo, B. (2017). The economic consequences of family policies: lessons from a century of legislation in high-income countries. *Journal of Economic Perspectives*, 31(1):205–230. 3

Petit, P. (2007). The effects of age and family constraints on gender hiring discrimination: A field experiment in the French financial sector. *Labour Economics*, 14(3):371–391. 21

Piketty, T. (2003). The impact of divorce on school performance: evidence from France, 1968-2002. *Available at SSRN 482863*. 1, 5

Rodríguez-Moro, B. and Román, O. (2024). Measuring relationship quality. *Working Paper*. 2, 7, 33

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*. 11

Sevilla, A. and Smith, S. (2020). Baby steps: The gender division of childcare during the COVID-19 pandemic. *Oxford Review of Economic Policy*, 36(Supplement_1):S169–S186. 5

Siminski, P. and Yetsenga, R. (2022). Specialization, comparative advantage, and the sexual division of labor. *Journal of Labor Economics*, 40(4):851–887. 1, 5, 12, 18, 19

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199. 11

Svarer, M. and Verner, M. (2008). Do children stabilize relationships in denmark? *Journal of Population Economics*, 21:395–417. 5, 14

University of Essex, Institute for Social and Economic Research (2022). Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009. Technical Report 15th Edition, UK Data Service. http://doi.org/10.5255/UKDA-SN-6614-16. 6

Voena, A. (2015). Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples? *American Economic Review*, 105(8):2295–2332. 4

Weiss, Y. and Willis, R. J. (1997). Match quality, new information, and marital dissolution. *Journal of Labor Economics*, 15(1, Part 2):S293–S329. 1, 4, 41

A RQ measure

This appendix outlines the construction and validity of RQ. The companion paper, Rodríguez-Moro and Román (2024), provides details on these exercises, and extends the analysis of the measure to discuss its value to guide theoretical studies.

A.1 Self-Completion Partner Module

Survey text. *For each of the following questions, please indicate which best describes your relationship with your partner at the moment. Please select only one answer per question.*

How often do you have a stimulating exchange of ideas?

1	Never
2	Less than once a month
3	Once or twice a month
4	Once or twice a week
5	Once a day
6	More often

How often do you calmly discuss something?

1	Never
2	Less than once a month
3	Once or twice a month
4	Once or twice a week
5	Once a day
6	More often

How often do you work together on a project?

1	Never
2	Less than once a month
3	Once or twice a month
4	Once or twice a week
5	Once a day
6	More often

How often do you discuss or consider divorce, separation or terminating your relationship?

1	All of the time
2	Most of the time
3	More often than not
4	Occasionally
5	Rarely
6	Never

Do you ever regret that you married or lived together?

1	All of the time
2	Most of the time
3	More often than not
4	Occasionally
5	Rarely
6	Never

How often do you and your partner quarrel?

1	All of the time
2	Most of the time
3	More often than not
4	Occasionally
5	Rarely
6	Never

How often do you and your partner "get on each other's nerves"?

1	All of the time
2	Most of the time
3	More often than not
4	Occasionally
5	Rarely
6	Never

Do you kiss your partner?

1	All of the time
2	Most of the time
3	More often than not
4	Occasionally
5	Rarely
6	Never

Do you and your partner engage in outside interests together?

1	All of them
2	Most of them
3	Some of them
4	Very few of them
5	None of them

The responses below represent different degrees of happiness in your relationship. The middle point, "happy", represents the degree of happiness of most relationships. Please select the number which best describes the degree of happiness, all things considered, of your relationship.

1	Extremely unhappy
2	Fairly unhappy
3	A little unhappy
4	Happy
5	Very happy
6	Extremely happy
7	Perfect

Responding to the module. Due to the personal nature of the questionnaire, preserving the privacy of respondents is a priority of Understanding Society. The questionnaire is part of a self-completion module, meaning that it is not administered by an interviewer but completed individually, either on paper or a computer, regardless of the interview mode.[28]

Although the questionnaire is answered individually, respondents may still be concerned about others being present during the interview. Individuals are given the option to refuse to answer the self-completion part of the survey, and if they do so due to the presence of another person, this reason is recorded separately. However, this is rarely reported: in the most recent wave, no one mentioned it, and in the previous wave, only three out of 363 refusals were due to another person being present.[29] Additionally, when an interviewer is present, they record whether another person is present and whether that person influences the interview. In the majority of cases, no one else is present during the interview.

[28]Understanding Society conducts interviews in three modes: face-to-face, by phone, and online. Until wave six, all interviews were conducted either face-to-face (98%) or by phone. Web mode was introduced in wave seven, and its use has steadily increased, reaching 87% in the most recent wave (13).

[29]Other reasons for refusal include disliking the computer format, a child needing attention, confidentiality concerns, lack of motivation, or time constraints.

A.2 Factor Analysis

Factor analysis is a statistical dimension reduction method that seeks to disentangle latent associations between different items. The goal is to identify a set of latent variables, or factors, that explain the shared variance among a set of observed items. The variance of each item is assumed to be influenced by these common factors, as well as by a unique error term specific to each item. The factors, which are unobservable, are inferred from their effects on the items. They are identified by solving a system of equations that produces a set of weights, called factor loadings, which indicate how strongly each item is associated with each factor.

Crucially, factor analysis differs from principal component analysis (PCA), another common data reduction method, in both assumptions and interpretation. PCA aims to reduce dimensionality by finding the linear combinations of items that capture the largest possible variance, driven purely by the data and without assuming an underlying structure. In contrast, factor analysis assumes that observed variables are influenced by latent factors and explicitly models the unique variance of each variable. Its goal is to uncover the latent constructs rather than simply reducing the number of dimensions. In our case, the responses to the questionnaire items in Table 1 are the observed variables. Instead of treating relationship quality as linear combinations of the items, as PCA would, we consider that this variable is a latent factor influencing respondents' answers to each item of the questionnaire.

Before conducting the factor analysis, we transform the responses to the items in Appendix A.1 so that a value of zero represents the worst couple behaviors and habits, like quarrelling all of the time or never kissing one's partner, to ensure that all item responses are increasing in the quality of the relationship. We then perform factor analysis on the ten items, and we retain the first factor as the measure of RQ. This decision is made based on the eigenvalue of the factors. The eigenvalue of the first factor is 4.05, more than double that of the next factor, while all subsequent factors have eigenvalues below 1. This first factor explains 40.49% of the variation in the items.

Figure A.1 displays the factor loadings obtained. All factor loadings are positive and greater than 0.5, indicating a strong association between RQ and each questionnaire item. The highest correlations with the RQ measure are found in items related to having stimulating exchanges of ideas and regretting getting married, while the lowest correlations are with items related to the frequency of kisses and of calm discussions.

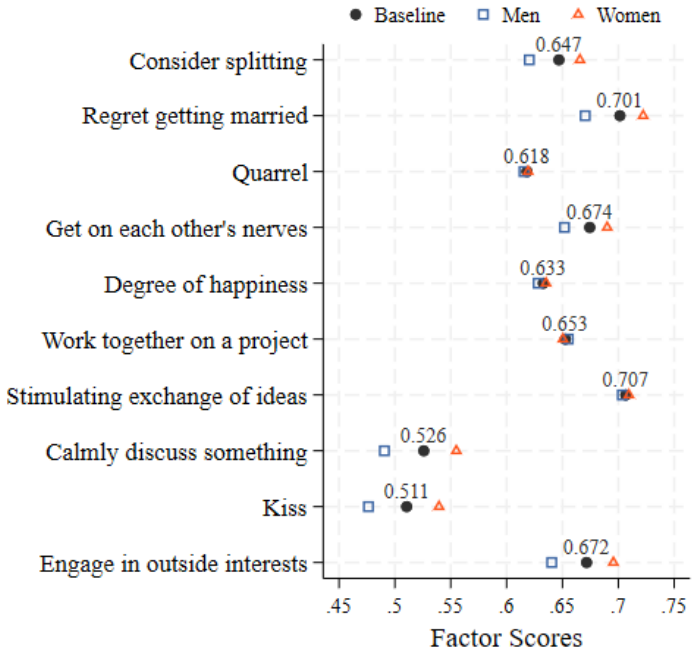
In Figure A.1, we plot the factor scores derived from computing the factor analysis separately for men and women, constructing a man-specific and a woman-specific RQ measure. The factor loadings for men are lower, with an eigenvalue of 3.84, indicating weaker correlations compared to the main RQ measure. For women, the loadings are higher, with a correspondingly larger eigenvalue. Despite these differences, the item rankings remain largely consistent across genders, with women valuing the items more highly and men consistently lower.

Table A.1: Factor loadings of RQ.

(a) Subjective assessment		(b) Couple time use	
How often do you... ?		*How often do you... ?*	
consider splitting	0.647	work together on a project	0.653
regret getting married	0.701	stimulating exchange of ideas	0.707
quarrel	0.618	calmly discuss something	0.526
get on each others nerves	0.674	kiss partner	0.510
What is the... ?		*Do you and your partner... ?*	
degree of happiness w/ relationship	0.633	engage in outside interests	0.672

Notes: This table reports the factor loadings resulting from computing a factor analysis on the 10 items in the Understanding Society Partner module. The first factor is the measure of relationship quality used in the analysis, which we call RQ. It has eigenvalue 4.06 and explains 40.61% of the variation in the data. The left panel shows the subjective assessment items and the right panel displays the couple time use items.

Figure A.1: Factor loadings by sex.



Notes: This graph displays the factor loadings from computing a factor analysis on the complete data, only on women and only on men. The eigenvalues are 4.01, 3.84 and 4.23 for each sample, respectively.

A.3 Validity

Given the novelty of the measure, we conduct a series of tests to verify that RQ provides sensible information about the quality of a relationship. We follow the life satisfaction literature (Ferrer-i Carbonell and Frijters, 2004) and substantiate two fundamental theoretical assumptions: informativeness and interpersonal comparability. Additionally, we verify that RQ evolves smoothly overtime and does not experience abrupt jumps, and that it correlates with observed characteristics of individuals and partners in expected ways.

Informativeness. First, we verify that the information provided by RQ is meaningful. To do so, we assess the predictive capacity of RQ for couple decisions: (a) marriage and separation and (b) fertility decisions. Marriage increments separation costs, serving as a commitment mechanism. Hence, couples transitioning into marriage should report higher levels of RQ. In contrast, separations typically result from poor-quality relationships. Thus, we expect lower than average RQ levels on couples on the brink of dissolution. Finally, we hypothesise that couples deciding to have a child exhibit higher RQ levels.

To evaluate the predictive power of RQ on these decisions, we compare the distribution of the measure in the period before the decision with the overall distribution of RQ. Figure A.2 plots the empirical cumulative distribution function (CDF) of RQ for different samples. Panel (a) compares the overall distribution of RQ in the complete data with distribution one period before marriage and one period before dissolution. As anticipated, the distribution preceding marriage is shifted to the right, indicating that individuals report higher RQ values before marriage across the entire distribution. Conversely, the distribution before dissolution is significantly shifted to the left. Individuals report lower RQ before dissolution across the entire distribution. Interestingly, the pre-dissolution distribution exhibits a more substantial deviation from the overall distribution than the pre-marriage distribution. This indicates that negative deviations in RQ have a more pronounced influence on marital decisions than positive deviations.

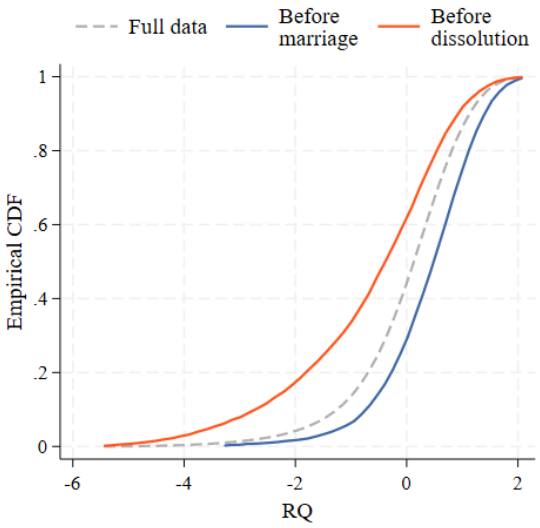
Figure A.2 (b) compares the distribution of RQ within the analysis sample of individuals becoming parents for the first time. We compare the general distribution of RQ in this sample to the distribution of the measure one period before the birth of the first child, at the time of conception. This distribution is slightly shifted to the right in comparison to the benchmark.

To formally test the differences between these distributions, we compute a two-sample Kolmogorov-Smirnov equality-of-distributions test. This test is designed to verify whether two samples are drawn from the same population and, thus, follow the same distribution. Table A.2 presents the D-statistics and p-values derived from this test for the samples considered. We find that the pre-divorce and pre-marital samples exhibit significantly smaller and significantly larger values than the full sample, respectively. Additionally, the pre-child sample contains significantly larger values than the complete sample of first-time parents. In aggregate, the test outcomes indicate that all three samples stem from distributions distinct from the benchmarks.

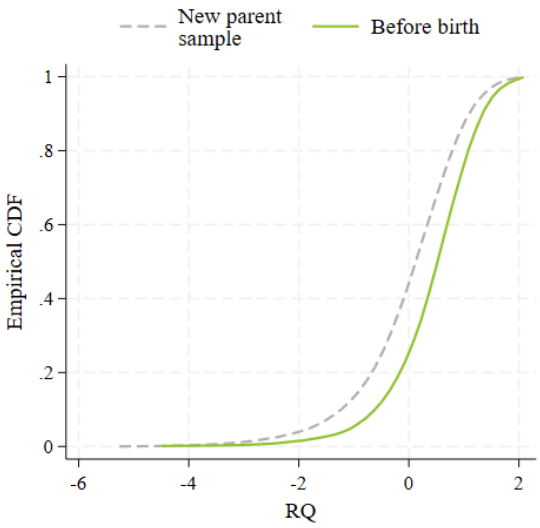
We delve deeper in the correlation between RQ and couple dissolution in the regression analysis in Table A.3. We regress a binary variable that equals one on the period before dissolution on RQ, and control for age, sex, college education, employment status, log monthly income, presence of children, relationship tenure, marital status, area of residence and period. We carry this analysis out separately in the full data and in the sample of parents used for the main analysis. All else equal, a standard deviation increase in RQ is associated with approximately a 0.8 percentage point higher probability of separation.

Figure A.2: Informativeness: behavior prediction.

(a) Marital transitions



(b) First child birth



Notes: This figure displays the empirical CDF of RQ for different samples. Panel (a) presents the distribution in the complete data, one period before marriage and one period before dissolution. Panel (b) displays the distribution in the complete analysis sample of individuals becoming parents for the first time and the distribution one period before birth.

Table A.2: Two-sample Kolmogorov-Smirnov test.

	d_0 = Full sample		d_0 = New parent sample
	d_1 = Before separation	d_1 = Before marriage	d_1 = Before first child
$d_0 > d_1$	0.000	0.1741	0.2491
	(1.000)	(0.000)	(0.000)
$d_0 < d_1$	-0.2192	0.0000	-0.0001
	(0.000)	(1.000)	(1.000)
Combined	0.2192	0.1741	0.2491
	(0.000)	(0.000)	(0.000)

Notes: This table displays the results of two-sample Kolmogorov-Smirnov tests on different samples. The reported coefficients are the resulting D-statistics and p-values (in parentheses).

The periods preceding marital transitions and fertility decisions are characterized by significant deviations from the average RQ. We conclude that RQ provides valuable information about couple behaviour, which is largely dictated by the quality of the relationship. This finding is consistent with the theoretical assumption of measurability.

Interpersonal comparability. Second, there should exist some degree of commonality in the concept of RQ across individuals. We test this by assessing the level of correlation of RQ between the members of a couple. Table A.4 displays the descriptive results from regressing women's RQ on their partners' RQ. Column (1) presents the partial correlation between partners' RQ and Column (2) introduces a relevant set of individual characteristics. Man RQ is a

Table A.3: Correlation between RQ and couple dissolution.

	Full data		Parent sample	
	(1)	(2)	(3)	(4)
Lagged RQ	-0.00931***	-0.00862***	-0.00833***	-0.00698*
	(0.001)	(0.001)	(0.002)	(0.003)
Controls	✓	✓	✓	✓
Individual FE		✓		✓
R-squared	0.037	0.024	0.054	0.050
Observations	106826	106826	15555	15555

Notes: This table displays the descriptive results from regressing a binary variable that equals one on the period before dissolution on RQ. The controls used are age, sex, college education, employment status, log monthly income, presence of children, relationship tenure, marital status, area of residence and period. Columns 1-2 do so using all available data on RQ, and columns 3-4 use the parent sample constructed for the main analysis. Standard errors clustered at the couple level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

highly significant predictor of woman RQ. In fact, it is the largest in magnitude, being almost five times larger the second largest: being married.

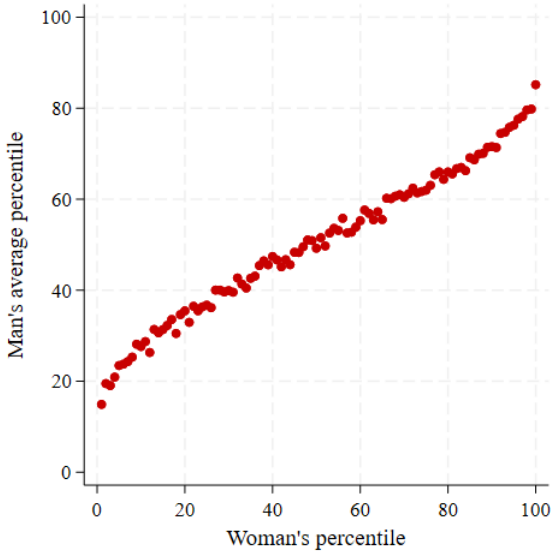
Table A.4: Regression of woman RQ on man RQ.

	Wife RQ		
	(1)	(2)	(3)
Husband RQ	0.613***	0.603***	0.594***
	(0.007)	(0.007)	(0.008)
Age × Tenure × Wave		✓	✓
Controls			✓
R-squared	0.320	0.330	0.334
Observations	42889	42889	39525

Notes: This table displays the descriptive results from regressing women's RQ on their (male) partners' RQ. The controls used are age, sex, college education, employment status, log monthly income, presence of children, relationship tenure, marital status, area of residence and period. Standard errors clustered at the couple level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We look at the non-linear relation between the RQ of both couple members through a rank-rank plot. Figure A.3 displays the average RQ percentile rank of men per woman's percentile rank. Although there is no perfect correlation between the two, there is a clear positive relation. Perfect correlation would result in a 45 degree line. The slope is steepest for the top and bottom percentiles, being of around one point. It flattens out at the center of the distribution by almost half. This indicates that extreme assessments of the quality of the relationship are shared much more intensely than intermediate ones.

Figure A.3: Rank-rank correlation of RQ across couple members.



Notes: This figure plots the average husband RQ percentile per wife RQ percentile.

Smooth evolution. We study the evolution of RQ over time by looking at this measure progresses with age and relation tenure. We estimate the following regression:

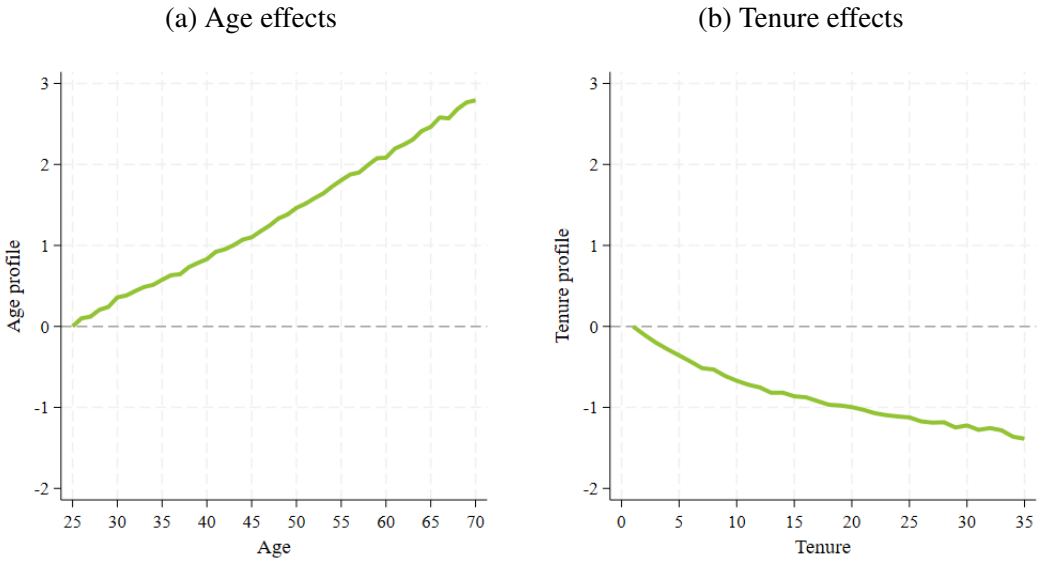
$$y_{i,t} = \alpha_i + \mu_t + \sum_a \mathbb{1}\{a = \text{age}_{i,t}\}\gamma_a + \sum_d \mathbb{1}\{d = \text{tenure}_{i,t}\}\lambda_d + u_{i,t}$$

where $y_{i,t}$ denotes RQ of individual i at time t. We include full sets of age and relationship tenure dummies. We use a two-way fixed effects approach to eliminate unobservable individual heterogeneity, which contain cohort effects, as well as period effects. Doing so, we preserve only the variation that can be attributed to an additional year of age or tenure. Since we include both variables non-parametrically, the estimated coefficients provide the age and tenure profiles of RQ.

Figure A.4 (a) plots the age profile of RQ, in comparison to the baseline of 25 years. Whilst this is clearly observational, aging has a positive effect on RQ. Additional years of age induce increasingly larger levels of RQ. These increments are highly smooth and almost linear. Figure A.4 (b) does the same for tenure, taking new relationships as a baseline. RQ steeply decreases with tenure during the first ten to fifteen years. It stabilizes for sufficiently long relationships. As with age, additional years of tenure reduce RQ smoothly, without significant jumps.

Observable characteristics. Past literature has linked match quality to observable characteristics of individuals and couples (Weiss and Willis, 1997; Eckstein et al., 2019; Low, 2024). To

Figure A.4: Age and tenure effects on RQ.



Notes: This figure plots the age and relationship tenure profiles of RQ. These are obtained estimating a non-parametric regression of age and tenure on RQ through fixed effects. Panel (a) takes 25 as the baseline age and panel (b) takes 1 as the baseline tenure.

investigate how RQ relates to these traits, we estimate the following model:

$$y_{i,t} = \mathbf{X}_{i,t}\beta + \mu_t + \sum_a \mathbb{1}\{a = \text{age}_{i,t}\}\gamma_a + \sum_d \mathbb{1}\{d = \text{tenure}_{i,t}\}\lambda_d + u_{i,t}$$

Here, $y_{i,t}$ denotes RQ of individual i at wave t, $\mathbb{1}\{a = \text{age}_{i,t}\}$ and $\mathbb{1}\{d = \text{tenure}_{i,t}\}$ represent age and tenure dummies, respectively, and μ_t denotes period dummies. $\mathbf{X}_{i,t}$ is a vector of individual and couple characteristics. Individual characteristics include age, sex, college education, employment status, and log personal monthly income of both the individual and their partner. Couple characteristics include relationship tenure, marital status, children, urban residence, and the female shares of both the labor market and housework, as defined in Subsection 2.2. Standard errors are clustered at the couple level.

Table A.5 columns (1) and (3) present the results from this estimation, both for the full sample and the parent sample used in the main analysis of this paper. The magnitudes and signs of the coefficients are maintained across samples. Women consistently report lower RQ levels than men. College education, for both the individual and their partner, is positively correlated with RQ. Notably, couple characteristics show the strongest associations: married individuals report higher RQ, while couples with children report lower RQ, on average, than their unmarried and childless counterparts. Additionally, higher female shares of housework are significantly associated with lower RQ.

This estimation provides level differences in RQ across individuals and couples with varying characteristics. However, unobserved individual heterogeneity may correlate with both the

observable characteristics and RQ. To account for this, we re-estimate the model including individual fixed effects, which removes this unobserved heterogeneity and offers within-individual estimates. The estimates from the fixed effects estimation reflect how changes in characteristics affect RQ for the same individual, rather than comparing different groups. The results are displayed in Table A.5, columns (2) and (4). The coefficients are smaller overall, and individual characteristics lose their significance, suggesting that compositional differences are driving the initial correlations with RQ. The coefficients for children and marriage remain significant, although marriage is only marginally so, indicating that RQ rises after marriage and declines with the presence of children.

Table A.5: Regression of RQ on individual, couple and partner characteristics.

	Full data		Parent sample	
	(1)	(2)	(3)	(4)
Female	-0.100***		-0.104***	
	(0.009)		(0.023)	
College	0.068***	-0.017	0.070*	-0.097
	(0.014)	(0.043)	(0.030)	(0.065)
Employed	0.038*	0.024	0.100	0.014
	(0.018)	(0.014)	(0.051)	(0.040)
Log Personal Income	0.017**	-0.002	0.014	0.006
	(0.006)	(0.004)	(0.013)	(0.010)
Married	0.256***	0.087*	0.333***	0.133*
	(0.029)	(0.035)	(0.060)	(0.057)
At least one child	-0.271***	-0.055**	-0.265***	-0.100**
	(0.021)	(0.019)	(0.043)	(0.034)
Urban	-0.057**	-0.009	-0.078*	-0.060
	(0.018)	(0.032)	(0.040)	(0.062)
Female share: labor market	-0.044	-0.036	-0.068	-0.029
	(0.024)	(0.026)	(0.075)	(0.058)
Female share: housework	-0.105**	-0.012	-0.183*	0.004
	(0.032)	(0.025)	(0.072)	(0.061)
Partner college	0.083***	0.050	0.071*	0.042
	(0.014)	(0.042)	(0.029)	(0.065)
Partner employed	0.007	-0.005	0.035	-0.070
	(0.016)	(0.013)	(0.048)	(0.042)
Partner income	0.017**	0.000	0.027*	-0.004
	(0.006)	(0.004)	(0.014)	(0.009)
Constant	0.106	-0.262	0.100	-0.243
	(0.118)	(0.541)	(0.263)	(0.776)
Age × Tenure × Wave	✓	✓	✓	✓
Individual FE		✓		✓
R-squared	0.047	0.053	0.078	0.110
Observations	54011	54011	10494	10494

Notes: This table displays the results from regressing a set of individual, couple and partner characteristics on RQ. Columns 1-2 present the results of this estimation in the full sample and columns 3-4 in the new parent sample used in the main analysis of this paper. Individual and partner characteristics include sex, college education, employment and log monthly personal income. Couple characteristics contain marital status, children, urban residence, and women's share in labor market work and in housework, as defined in Subsection 2.2. We control for age, relationship tenure and wave non-parametrically. Columns 1 and 3 present the results from a pooled OLS estimation, and columns 2 and 4 display the fixed effects estimation results. Standard errors clustered at the couple level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B Empirical Strategy

We take a dynamic difference-in-differences (DiD) approach to study the causal impact of children on RQ. This methodology exploits sharp changes in individual outcomes after the birth of their first child and allows for treatment heterogeneity with time relative to birth. In essence, it allows us to estimate the impact of having a child at each period relative to birth. Furthermore, it exploits differences in the timing of child birth to dynamically construct a control group.

We use a sample of individuals who become parents to study the impact of the birth of the first child (treatment) on RQ (outcome). This has two important implications. First, the treatment is staggered because different individuals have their first child in different periods. This divides the sample into different treatment adoption cohorts, depending on the calendar year when they become parents. Second, having a sample of new parents implies that everyone in the sample is treated at some point. This means that, in each period, we compare individuals who just had a child with individuals who have not become parents yet and with individuals who have been parents for more than one period.

We estimate the dynamic two-way fixed effects (TWFE) regression in Equation 1. The econometrics literature has pointed out several problems derived from estimating such specifications through OLS in contexts of staggered treatment adoption (see Goodman-Bacon, 2021, for a review). There are two main issues that threaten the causal interpretation of the obtained estimates. First, it requires assuming homogeneity of treatment effects across treatment cohorts. The violation of this assumption induces negative weights when computing the estimates of the average treatment effects, resulting in biased estimates. Second, it carries out *forbidden* comparisons between individuals changing status from control to treated and cohorts that have already been treated for more than one period. As we explain below, we take an alternative approach in our estimation.

B.1 Callaway and Sant'Anna (2021) method

The method proposed by Callaway and Sant'Anna (2021) overcomes the issues related to TWFE estimation by clearly separating identification, estimation and aggregation. This method is applicable in settings where a panel dataset is available and the binary and absorbing treatment is adopted in a staggered manner. We feel comfortable assuming that having a child is a binary (having a child or not) and absorbing treatment (having a child forever). The only situation where this treatment would not be absorbing is in the case of child death, which is a very rare occurrence in the data. Additionally, the treatment is adopted in a staggered way since individuals have children in different periods, i.e., there are different treatment adoption cohorts.

Identification. The building block of this method are the average treatment effects on the treated (ATT) for each treatment adoption cohort g and at each period t, denoted as $ATT(g,t)$. Identifying these requires two assumptions:

A1. Limited anticipation. If a unit is untreated in period t, its outcome in that period does not depend on when it will be treated in the future. In our context, we need to assume that changes in RQ before the birth of the first child do not predict when individuals have their first child. Note that "limited" anticipation implies that the method allows for some preceding reaction, but an assumption needs to be made on the amount of periods before treatment.

A2. (Conditional) parallel trends based on "not-yet treated". All treatment cohorts would have evolved in parallel in absence of treatment. In our context, individuals' RQ would have evolved in parallel regardless of the period when they had their first child. Importantly, the method allows for the looser assumption of parallel trends conditional on covariates.

Those two assumptions allow to construct the counterfactual for each $ATT(g,t)$ using (i) the period before treatment as the baseline period, and (ii) all treatment cohorts that have not been treated by t as controls. Therefore, the control group at each t for the same cohort g varies because at each subsequent period new cohorts enter treatment status.

Estimation. Each $ATT(g,t)$ is estimated as a 2×2 difference-in-differences coefficient using the baseline period and the control group described. The estimation can be done in three ways, which use information from different parts of the data generating process: using outcome regression, inverse probability weighting, or doubly robust estimands. Refer to Callaway and Sant'Anna (2021) for detailed information on the estimation methods. We use the last method, since it constitutes the combination of the other two.

Aggregation. The $ATT(g,t)$ estimates are the building blocks used to summarize the treatment effects across treatment adoption cohorts, periods or time relative to treatment. In each case, aggregation involves carefully chosen and estimable weights for the different $ATT(g,t)$. Our main aggregation involves obtaining the ATT at each event-time. Crucially, since the control group varies at each $ATT(g,t)$, the aggregated figure will internalize compositional changes. This may prevent the causal interpretation of the aggregated coefficients, unless homogeneity across treatment adoption cohorts is assumed. In this scenario, we need to assume that the impact of the birth of the first child is the same regardless of the year of birth.

B.2 Identification

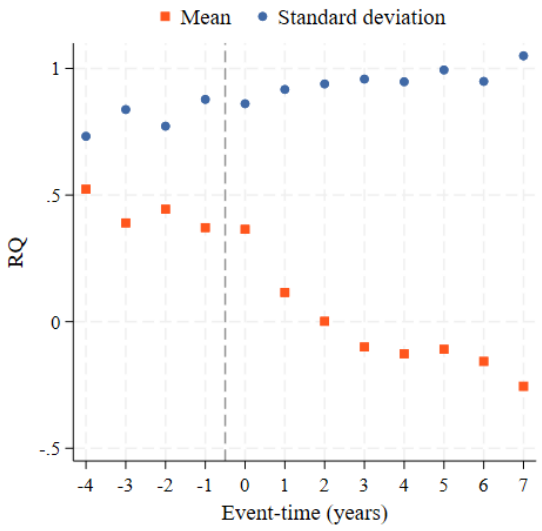
Although we cannot directly test whether the assumptions above are satisfied, we can provide evidence in favor of their plausibility.

The **limited anticipation** assumption would be violated if individuals decided to have children in response to either negative or positive shocks to RQ. Figure 2 indicates that, on average, there is no evidence that the birth of the first child is preceded by deviations in RQ, since all the point estimates before the event are not statistically different from zero. However, this could be masking heterogeneity in the evolution of RQ before first child birth, averaging individuals that decide to have children after a positive or a negative RQ deviation. If that were the case, the standard deviation of RQ would be larger over the periods before the event.

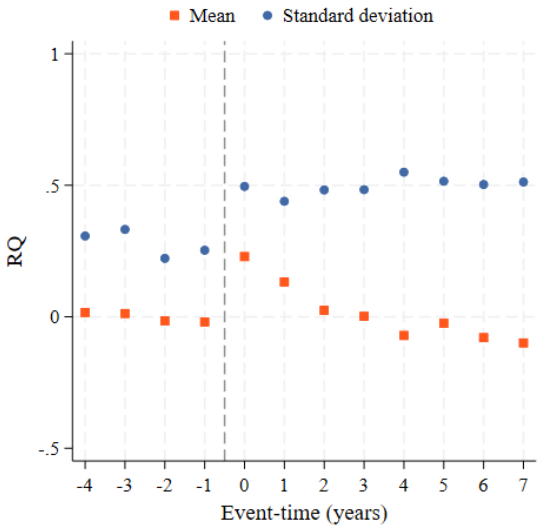
Figure B.1 (a) plots the mean and the standard deviation of the level of RQ at each event time. Consistent with Figure 2, average RQ is virtually flat before birth and steeply decreasing after. The standard deviation is smallest before the birth of the first child, supporting the evidence against anticipation, and it steadily increases after. Figure B.1 (b) does the same for deviations from the individual-specific average of RQ, computed separately before and after first child birth. Prior to childbirth, these deviations are nearly zero on average, suggesting that individuals do not experience significant RQ changes leading up to childbirth. This is accompanied by a low standard deviation, supporting the idea that this is not masking heterogeneity.

Figure B.1: Mean and standard deviation of RQ around first child birth.

(a) RQ levels



(b) Deviations from individual mean



Notes: This graph plots the sample average and the sample standard deviation of (a) the level of RQ at each period around first birth, and of (b) deviations from the individual-specific mean, computed separately before and after birth.

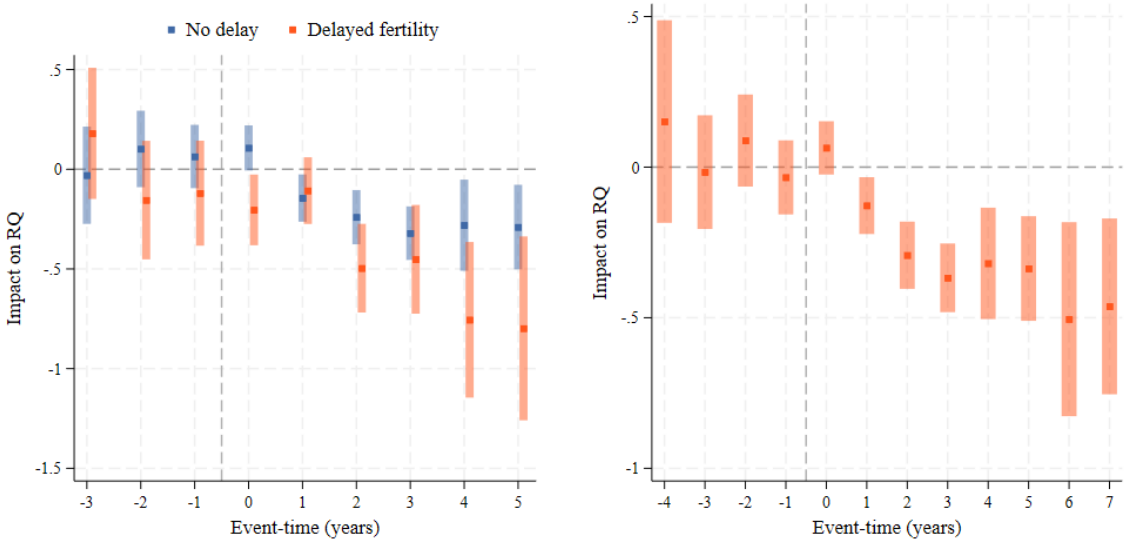
Conditional parallel trends cannot be tested, since it involves conjectures about unrealized scenarios. The flat pre-trends displayed in Figure 2 are also suggestive evidence in favor of this

assumption. We compute a formal test that the coefficients corresponding to all pre-treatment periods are statistically equal to zero and we cannot reject this hypothesis.

A related potential limitation to this approach is that we cannot observe intended fertility, but only the realized. We consider individuals who experience a delay in the birth of the first child separately, by looking at those reporting to have had a miscarriage or to have used fertility treatments.[30] Unfortunately, we cannot pin down exactly when these individuals first wanted to have a child, just that they experienced a delay.[31] Figure B.2 (a) displays the estimation results separately. The impact only really differs in the period of birth, where parents without a delay experience a slight increase in RQ, while delayed parents display a decrease. The point estimates are similar, but the precision is quite low in the estimates for parents who experience delayed fertility. Importantly, prior to birth, there are no significant differences in RQ between those who experienced delays, regardless of when they eventually had a child.

Figure B.2: Evidence on parallel trends assumption.

(a) Delayed pregnancy. (b) Similar baseline RQ.



Notes: This graph provides evidence in favor of the parallel trends assumption. (a) displays the impact of first child birth on RQ separately for individuals who experienced a delay and individuals who did not. We identify individuals who report having used fertility treatments or having had miscarriages before birth as subject to delayed fertility. (b) repeats the main analysis restricting comparisons to individuals with similar RQ levels before birth. 95% confidence intervals plotted.

Another way to verify the plausibility of the parallel trends assumption is to replicate our main analysis restricting comparisons to individuals with similar RQ levels before the birth of their first child, but who had children in different periods. The assumption is more plausible in this specification. The results, presented in Figure B.2 (b), are very similar to the main findings,

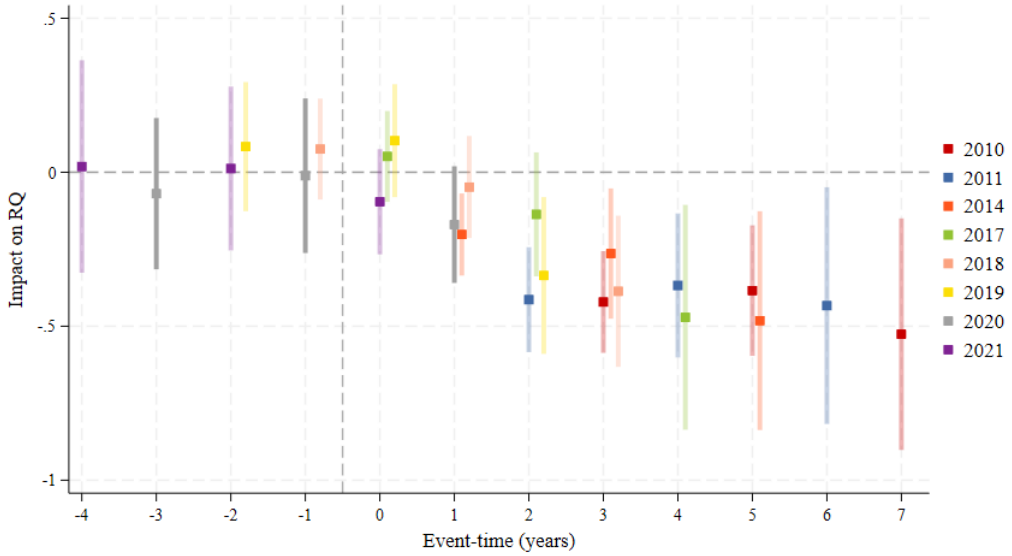
[30] Information on fertility treatments is only available if the treatment was successful and a pregnancy occurred since the previous interview. About 9% of each treatment cohort experiences a delay in fertility in our sample.

[31] On average, couples attempt to conceive for a year before using assisted reproduction technologies (Bögl et al., 2024).

though the decrease in RQ is slightly smaller, 0.4 instead of 0.5. This similarity reinforces the plausibility of the parallel trends assumption.

A final condition needed for the time aggregation displayed as our main result is **homogeneity of treatment effects across cohorts**. In Figure B.3 we plot the dynamic impact of the birth of the first child separately per treatment cohort. Given our sample size, the estimates are quite noisy at this level of disaggregation. The point estimates are quite similar across cohorts.

Figure B.3: Impact by treatment cohort.

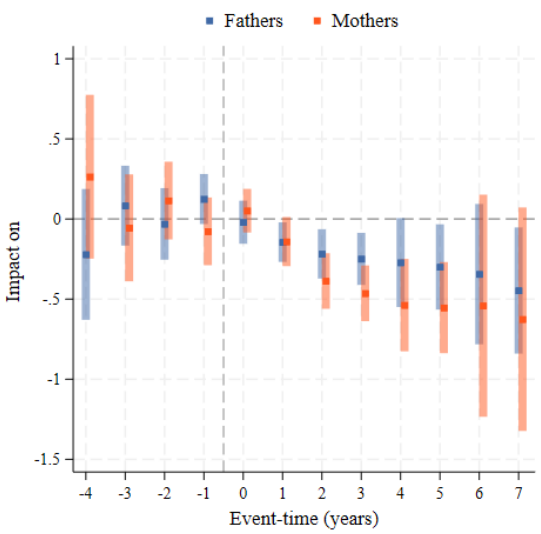


Notes: This graph displays the dynamic impact of the birth of the first child by treatment cohort, depending on the calendar year of birth. 95% confidence intervals plotted.

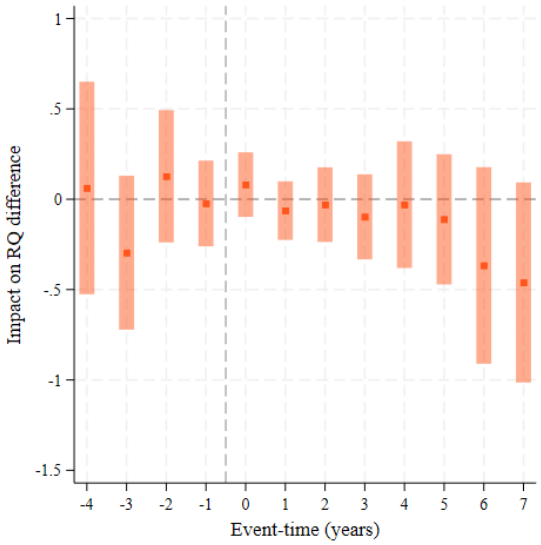
C Related Results

Figure C.1: Impact of first child birth on RQ for mothers and fathers.

(a) Separately

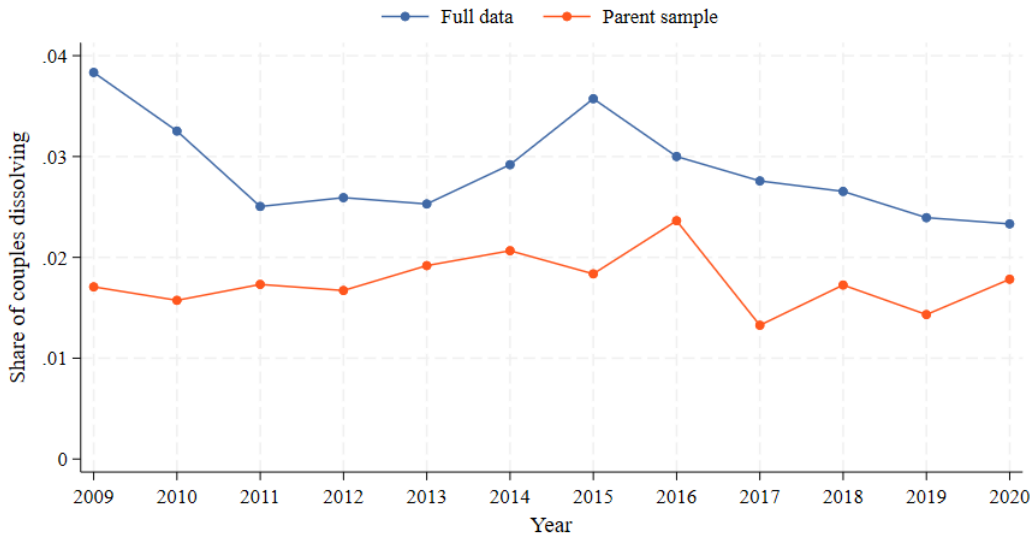


(b) Difference



Notes: This figure plots the results from estimating Equation 1 through the Callaway and Sant'Anna (2021) method (a) separately for mothers and fathers, and (b) using the RQ difference between wife and husband as the outcome. 95% confidence intervals plotted.

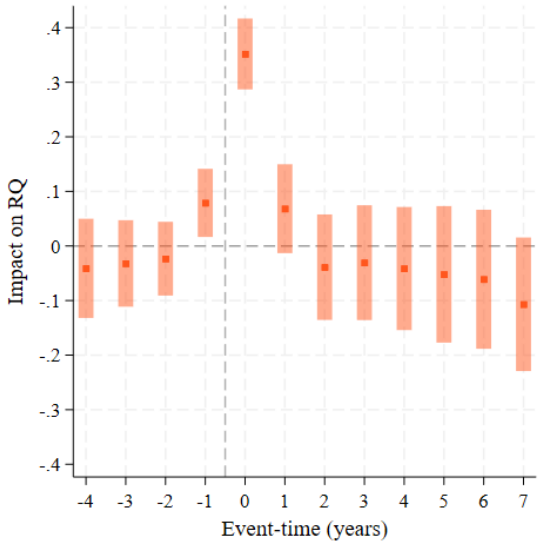
Figure C.2: Yearly couple dissolution rate.



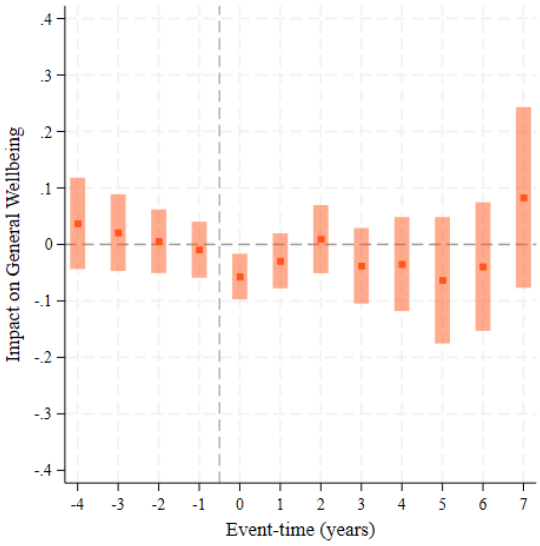
Notes: This figure plots the share of couples which dissolve out of all the couples observed each year in the complete data and in the new parent sample. According to Office for National Statistics (2022), on average 1.44% of the married couples in fertility ages (20-45 year-olds) living in England and Wales divorce every year, during the period 2009-2021.

Figure C.3: Impact of children on general happiness and subjective well-being.

(a) General happiness



(b) Subjective well-being



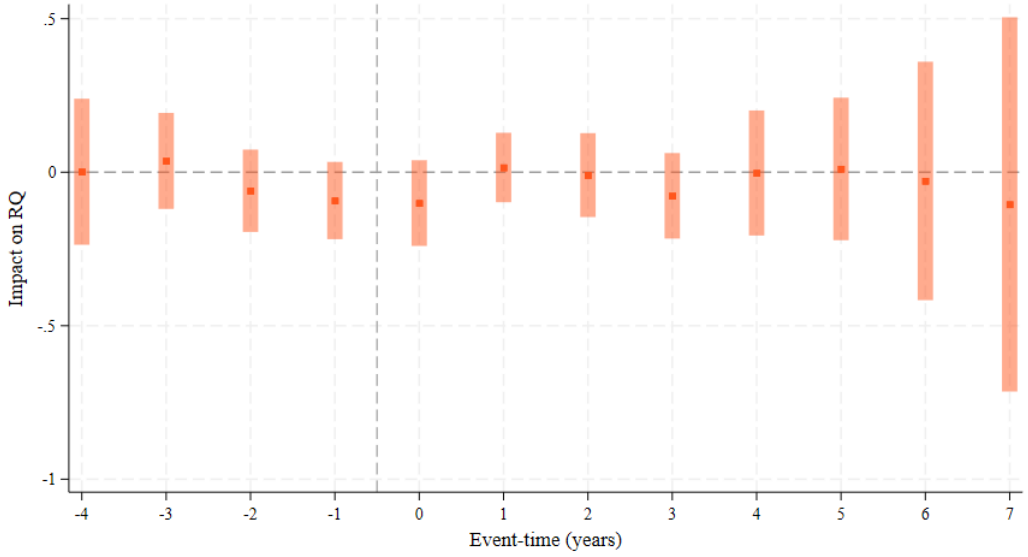
Notes: This figure plots the results from estimating Equation 1 through the Callaway and Sant'Anna (2021) method using (a) general happiness and (b) subjective well-being as outcomes. The question used for general happiness is "Have you recently been feeling reasonably happy, all things considered?". The responses are given in a four-point Likert scale. Subjective well-being is constructed by Understanding Society combining the answers to the 12 questions of the General Health Questionnaire. We standardize both measures, so the results are given in standard deviations. 95% confidence intervals plotted.

Table C.1: Regression of general happiness on RQ in the sample of never parents.

	General Happiness		
	(1)	(2)	(3)
RQ	0.222***	0.206***	0.185***
	(0.011)	(0.012)	(0.019)
Controls		✓	✓
Individual FE			✓
R-squared	0.047	0.069	0.033
Observations	18231	14953	14953

Notes: This table presents the results from regressing general happiness on RQ in the sample of individuals who never become parents. We identify never parents as those individuals not observed having children by the age of 50, the end of the fertility cycle. Controls include age, tenure, wave, sex, college education, employment status, marital status, and urban residence.

Figure C.4: Impact of unemployment on RQ.



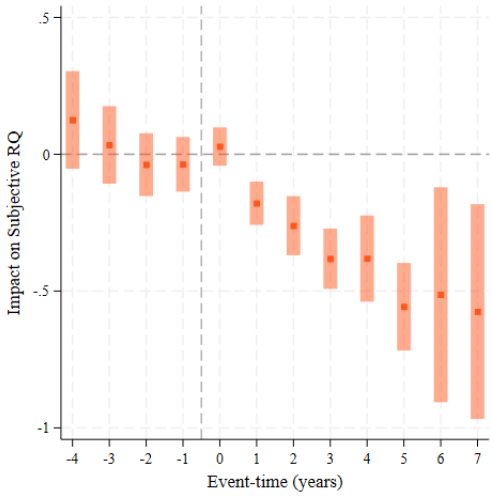
Notes: This figure plots the results from estimating Equation 1 through the Callaway and Sant'Anna (2021) around unemployment events. 95% confidence intervals plotted.

D Robustness

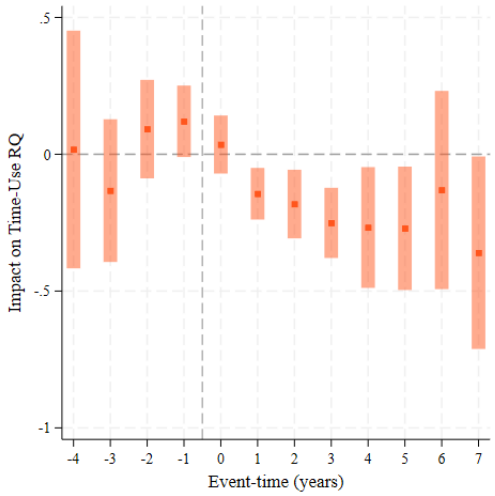
D.1 Time invariance of RQ

Figure D.1: RQ by item block.

(a) Subjective assessment

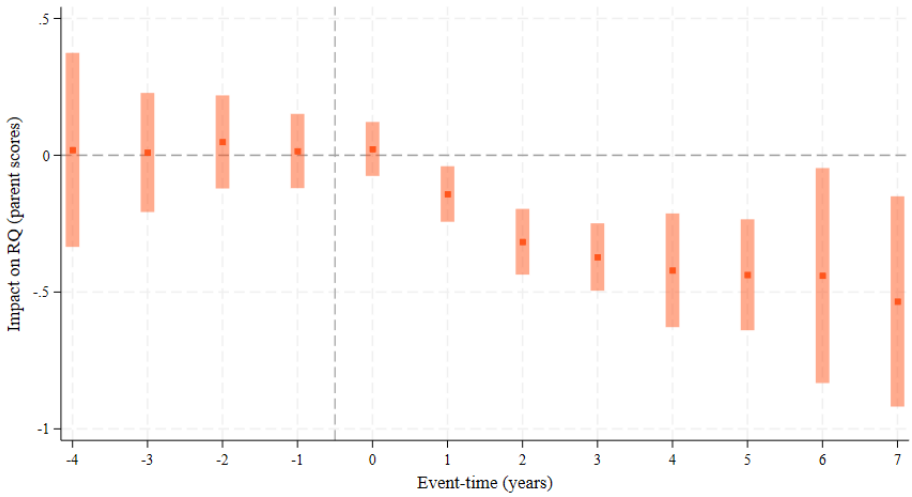


(b) Couple time use



Notes: This figure plots the results from estimating Equation 1 through the Callaway and Sant'Anna (2021) method. The outcomes are the resulting measures obtained from carrying out a factor analysis on (a) the subjective assessment and (b) the couple time use items in Table 1 separately. 95% confidence intervals plotted.

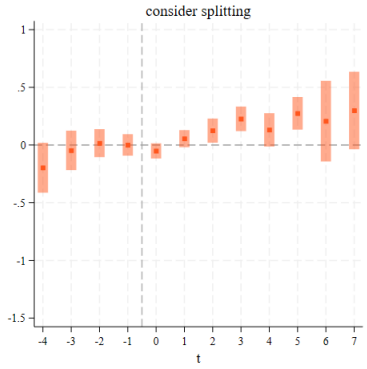
Figure D.2: Impact of children on RQ using factor scores after birth.



Notes: This figure plots the results from estimating Equation 1 through the Callaway and Sant'Anna (2021) method. The outcome is the resulting measure obtained from carrying out a factor analysis on the items in Table 1 using only observations corresponding to individuals who already became parents. 95% confidence intervals plotted.

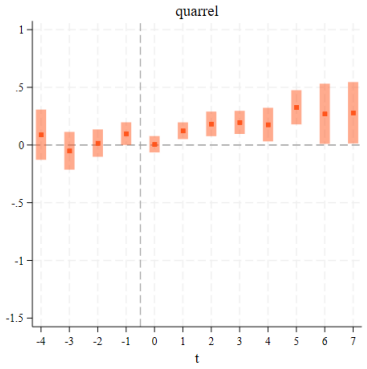
Figure D.3: Impact on subjective assessment items.

(a) Consider splitting

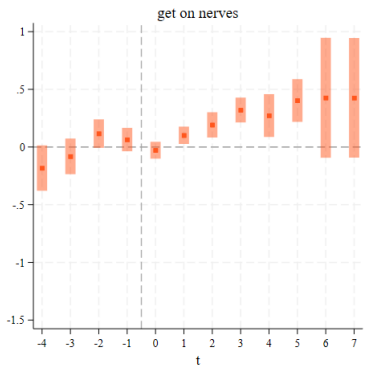


(b) Regret marriage

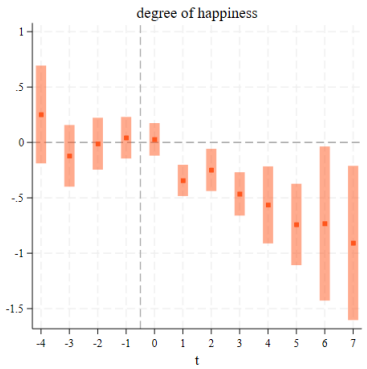
(c) Quarrel



(d) Get on nerves



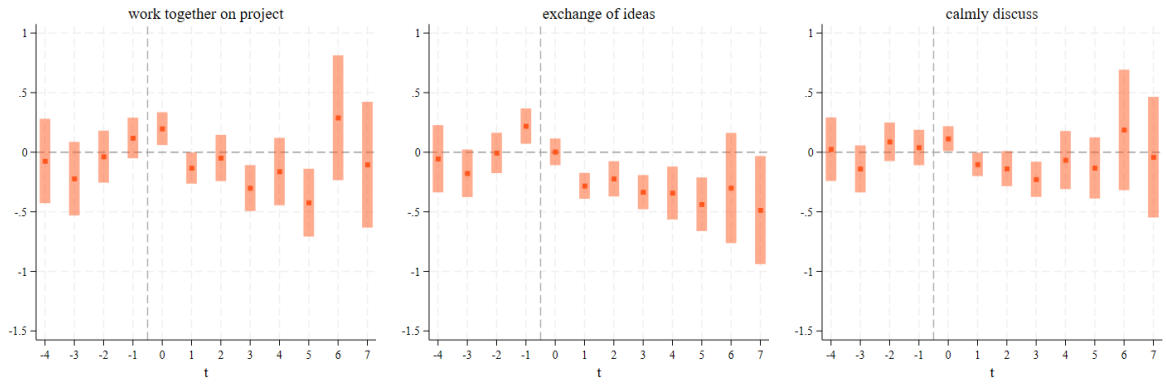
(e) Degree of happiness



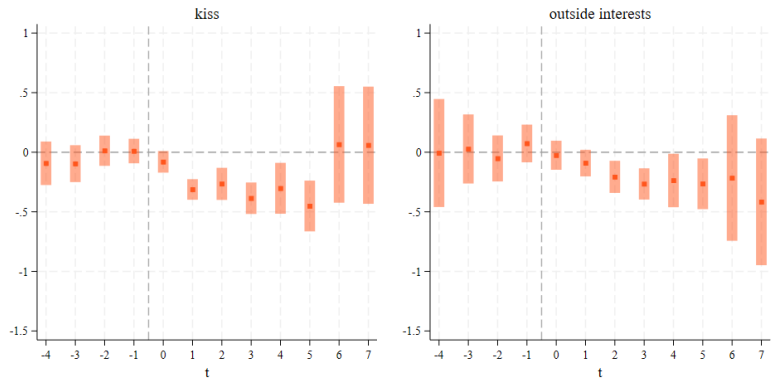
Notes: This figure presents the results from estimating the impact of first child birth on the subjective assessment items in Table 1 (a), using the Callaway and Sant'Anna (2021) method. 95% confidence intervals plotted.

Figure D.4: Impact on couple time use items.

(a) Work together (b) Exchange ideas (c) Calmly discuss



(d) Kiss (e) Outside interests

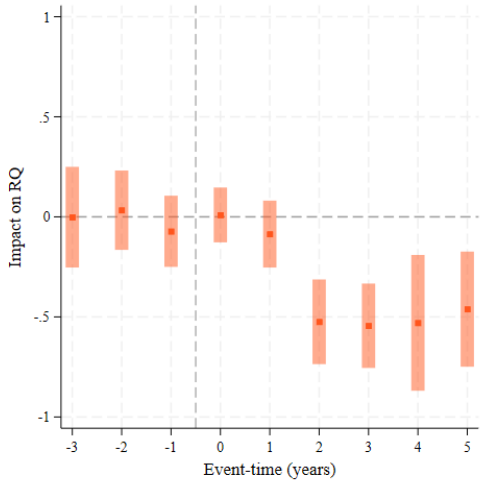


Notes: This figure presents the results from estimating the impact of first child birth on the couple time use items in Table 1 (b), using the Callaway and Sant'Anna (2021) method. 95% confidence intervals plotted.

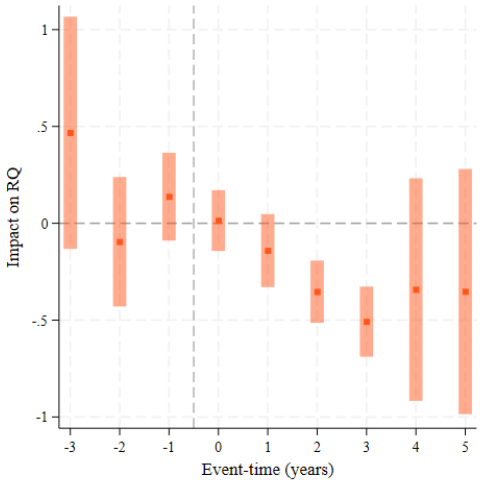
D.2 Total realized fertility

Figure D.5: Impact by final number of children.

(a) One child



(b) Two children



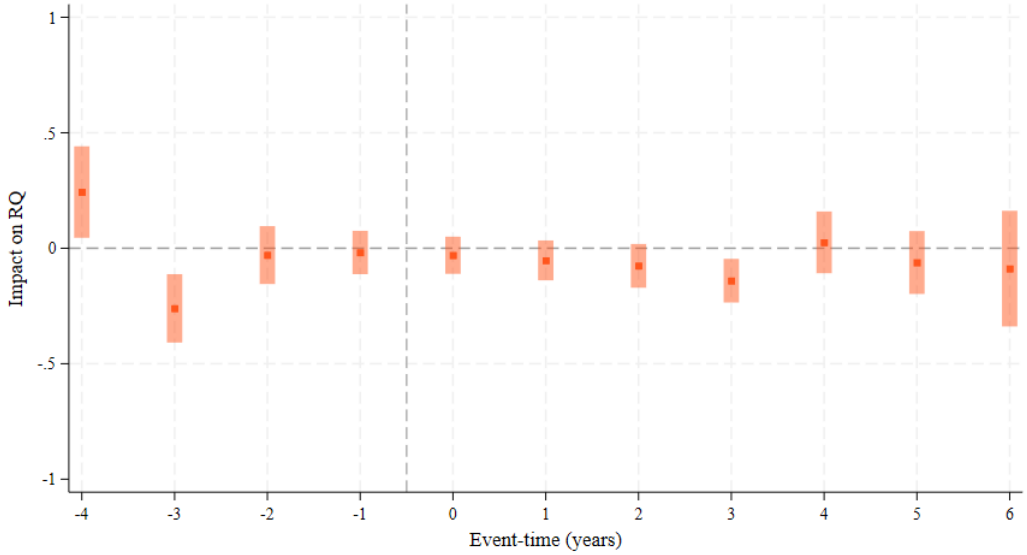
Notes: This figure plots the results from estimating Equation 1 through the Callaway and Sant'Anna (2021) method separately for (a) individuals who only have one child and (b) individuals having two or more children by the end of the observation period. 95% confidence intervals plotted.

Table D.1: Impact by total realized fertility.

	(1) One child	(2) More than one
ATT	-0.360***	-0.212*
	(0.102)	(0.101)
Observations	693	1041

Notes: This table displays the static difference-in-differences estimates obtained through the Callaway and Sant'Anna (2021) method separately for (a) individuals who only have one child and (b) individuals having two or more children by the time they are 40 or older. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure D.6: Impact of second child birth on RQ.

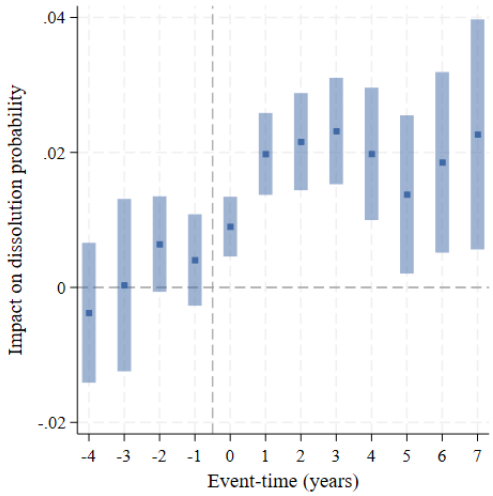


Notes: This figure plots the results from estimating Equation 1 through the Callaway and Sant'Anna (2021) method around the birth of the second child, for the subsample of individuals who become parents for a second time. 95% confidence intervals plotted.

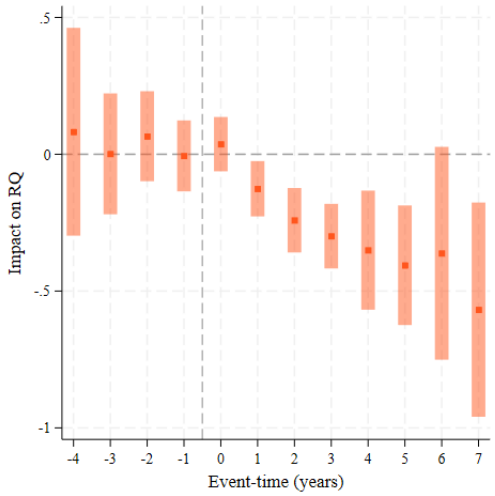
D.3 Selected sample

Figure D.7: Sample selection on separating couples.

(a) Impact on couple divorce.



(b) Impact on non-separating couples.

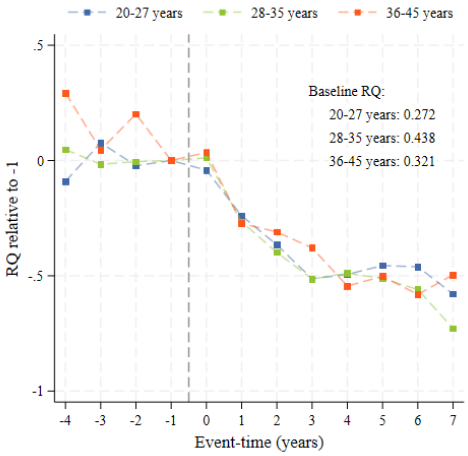


Notes: This figure plots the results from estimating Equation 1 through the Callaway and Sant'Anna (2021) method. The outcomes are (a) separation probabilities and (b) RQ on the subset of couples who do not separate during the observation period. Note that, by construction of the sample, individuals cannot separate before the time when the child is conceived. 95% confidence intervals plotted.

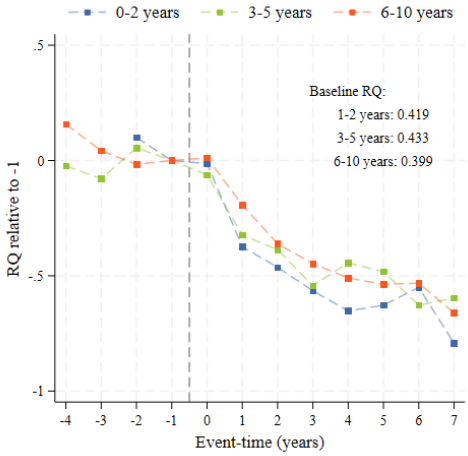
D.4 Age and tenure at birth

Figure D.8: Average RQ per event-time period.

(a) Age bins

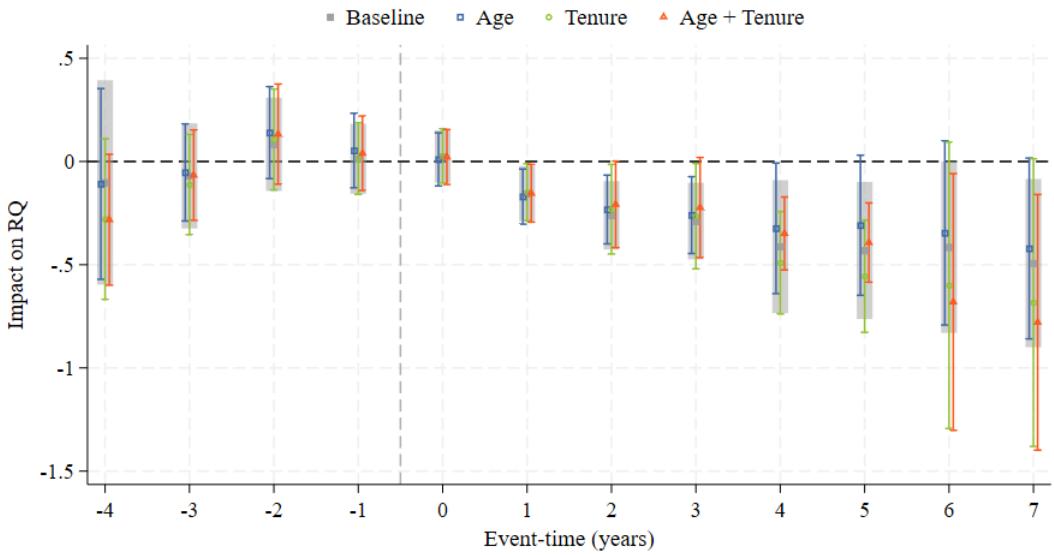


(b) Tenure bins



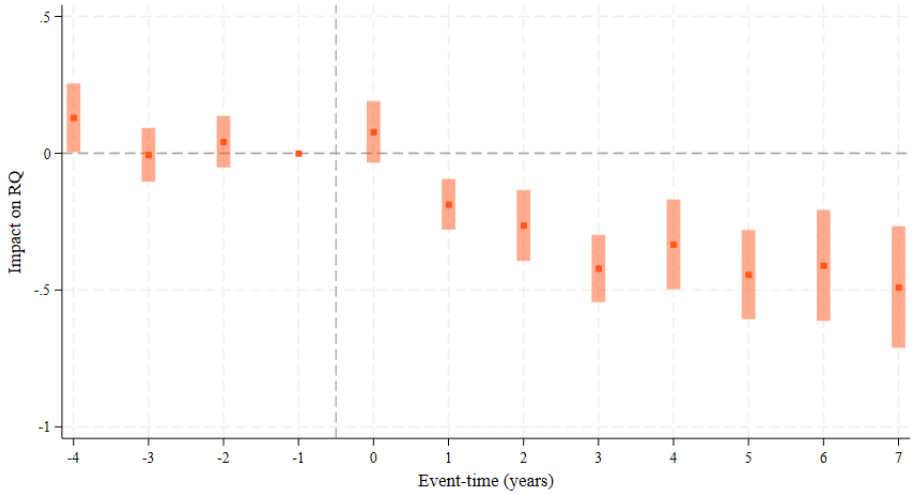
Notes: This figure plots the average RQ at each event-time period, normalizing the value on the period before the birth of the first child (-1) to zero, by (a) age and (b) tenure bin.

Figure D.9: Impact on RQ using different sets of controls.



Notes: This figure plots the results from estimating Equation 1 through the Callaway and Sant'Anna (2021) method for different sets of controls. This method only allows to account for the values of the covariates on the periods before the event. 95% confidence intervals plotted.

Figure D.10: Impact of first child birth on RQ using Kleven et al. (2019b).

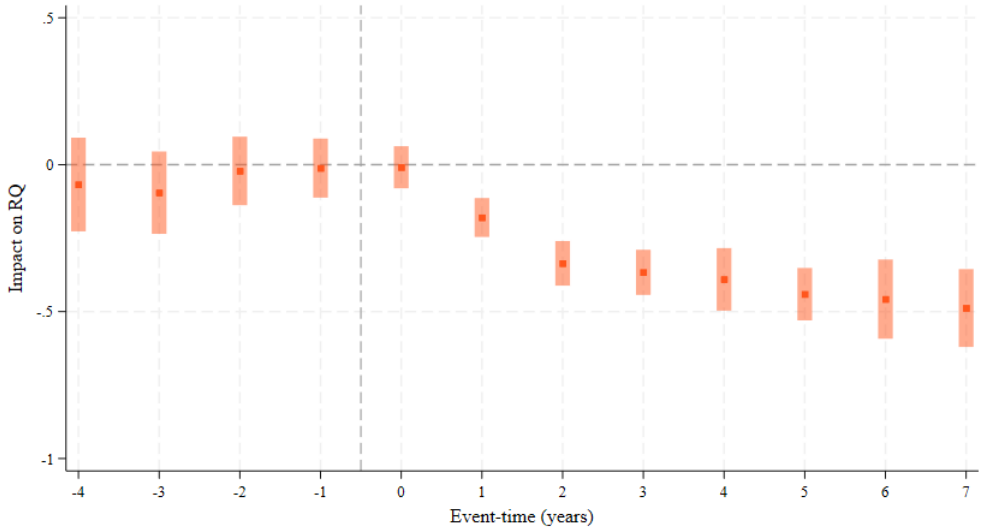


Notes: This figure plots the coefficients (and 95% confidence intervals) corresponding to the event-time dummies resulting from estimating an event-study specification as popularized by Kleven et al. (2019b), to account for the dynamic effects of age and tenure:

$$y_{i,t} = \mu_t + \sum_{j \neq -1} \mathbb{1}\{j = t - G_i\}\delta_j + \sum_a \mathbb{1}\{a = \text{age}_{i,t}\}\alpha_a + \sum_d \mathbb{1}\{d = \text{tenure}_{i,t}\}\gamma_d + u_{i,t}$$

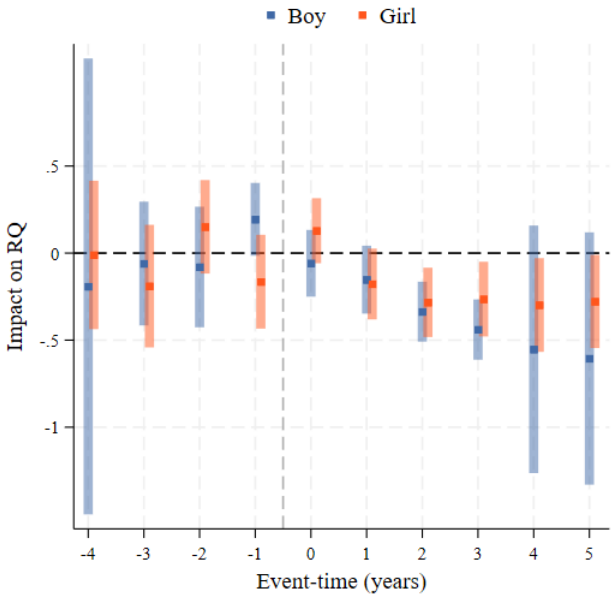
D.5 Additional checks

Figure D.11: Impact of first child birth on RQ using never parents as control.



Notes: This figure plots the results from estimating Equation 1 through the Callaway and Sant'Anna (2021) method and using never parents as the control group. Individuals are classified as never parents if they are not observed to become parents by age 50, the end of the fertility cycle. 95% confidence intervals plotted.

Figure D.12: Impact of having a first-born boy or girl.



Notes: This figure plots the results from estimating Equation 1 through the Callaway and Sant'Anna (2021) method separately for first-born boys and girls. 95% confidence intervals plotted.

E Household Specialization

Table E.1: Summary statistics before the birth of the first child, by couple type.

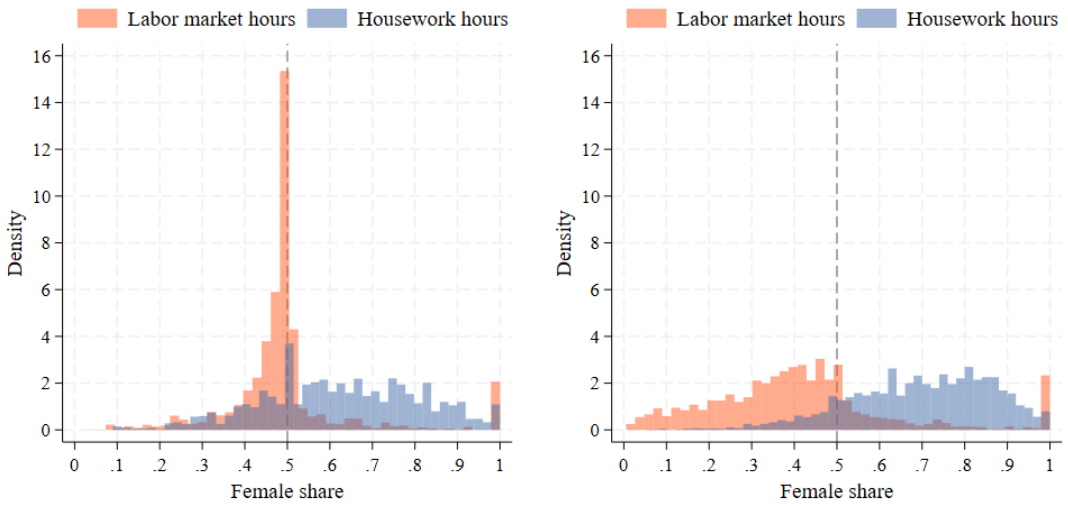
	Traditional		Unbalanced		Egalitarian		Counter-tradit.	
	Fathers (1)	Mothers (2)	Fathers (3)	Mothers (4)	Fathers (5)	Mothers (6)	Fathers (7)	Mothers (8)
(a) Individual characteristics								
Age	31.71	28.89	32.37	30.31	32.33	30.29	32.80	29.87
	(6.021)	(5.088)	(5.378)	(4.672)	(5.211)	(4.273)	(5.717)	(4.587)
College educated (%)	31.72	33.77	37.50	44.51	41.26	51.05	40.15	51.46
	(46.64)	(47.40)	(48.46)	(49.74)	(49.32)	(50.08)	(49.11)	(50.07)
Active in labor mkt (%)	98.70	85.31	92.82	98.49	99.65	100	96.74	92.73
	(11.37)	(35.33)	(25.85)	(12.22)	(5.913)	(0)	(17.79)	(26.02)
Employed (%)	97.39	77.85	90.36	98.11	99.65	98.25	94.93	90.91
	(15.97)	(41.48)	(29.54)	(13.64)	(5.913)	(13.13)	(21.98)	(28.80)
Weekly work hours	40.22	23.48	33.76	36.25	38.50	37.41	37.67	31.51
	(10.71)	(15.69)	(12.82)	(7.035)	(3.530)	(4.359)	(12.17)	(12.59)
Gross monthly income	2432.3	1180.7	2236.7	1914.5	2510.1	2065.8	2459.4	1975.4
	(1663.3)	(1040.2)	(1475.5)	(1066.0)	(1232.8)	(1199.2)	(1524.5)	(1251.1)
Weekly housework hours	3.305	12.26	3.833	9.713	6.941	7.122	8.210	5.803
	(2.542)	(6.343)	(2.959)	(4.966)	(3.435)	(3.442)	(4.108)	(3.935)
RQ	0.111	0.415	0.162	0.168	0.535	0.506	0.355	0.426
	(1.013)	(0.992)	(0.779)	(0.702)	(0.587)	(0.588)	(0.922)	(0.786)
Observations	230	228	529	529	286	286	276	275
(b) Couple characteristics								
Tenure	4.539		4.824		4.679		4.749	
	(3.267)		(3.085)		(2.816)		(2.995)	
Married (%)	65.73		70.36		65.68		68.23	
	(46.99)		(45.09)		(47.28)		(46.25)	
In urban areas (%)	80.21		74.80		78.71		75	
	(40.05)		(43.50)		(41.07)		(43.42)	
Female share of paid work	0.319		0.547		0.492		0.449	
	(0.189)		(0.159)		(0.0318)		(0.187)	
Monthly household income	3866.9		4220.8		4631.9		4500.1	
	(2290.6)		(2253.4)		(2266.8)		(2425.8)	
Female share of housework	0.788		0.721		0.507		0.400	
	(0.133)		(0.146)		(0.102)		(0.168)	
Observations	232		533		287		277	

Notes: This table presents mean values of a set of individual and couple characteristics in the sample the period before the birth of the first child, separately for each couple type. Standard deviations in parentheses.

Figure E.1: Distribution of the female share of labor market and housework hours.

(a) Before birth

(b) After birth

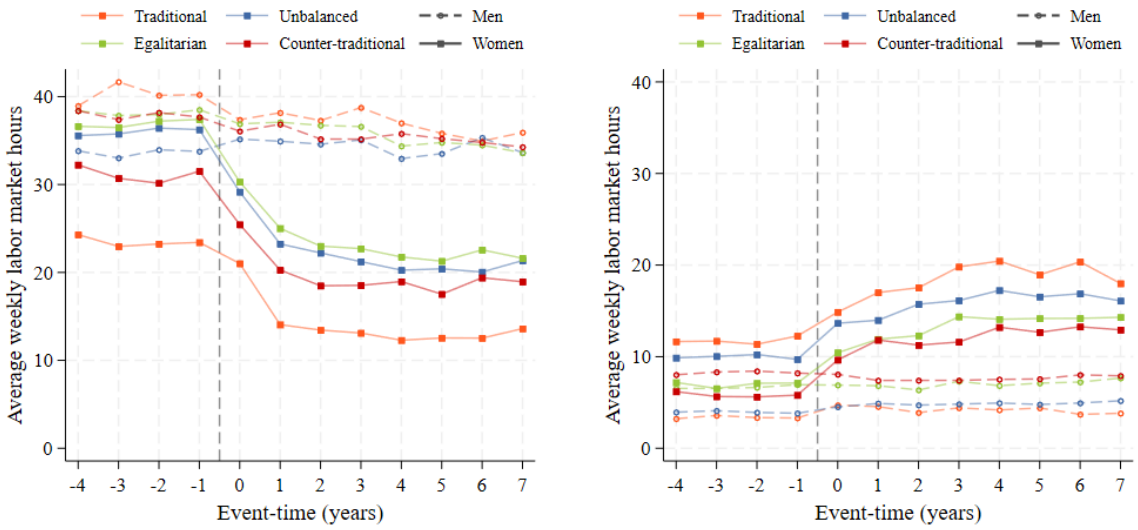


Notes: These graphs plot the distribution of the share of the household total housework and labor market hours carried out by women (a) before first child birth and (b) after.

Figure E.2: Average hours worked around first child birth.

(a) Paid work share

(b) Housework share



Notes: These graphs plot the average weekly hours that men and women spend in (a) labor market work and (b) housework at each time around the birth of the first child by couple type.