Master Thesis, Economics

Childbirth, household labor division and gender equality

Evidence from ten European countries

Francesco Cilia and Anders Halland

Supervisors:
Fanny Landaud and Vincent Somville

Norwegian School of Economics
Bergen, Fall 2019

Abstract

As couples transition into parenthood, they can face adjustments in terms of the labor division of standard household tasks. In this thesis, we analyze a selection of ten European countries utilizing a Regression Discontinuity-like approach and assess whether households experience any significant shift in relative standard housework shares. We exploit data obtained by the Generations and Gender Program to compare heterosexual couples that have just given birth to their first child to those who are about to do so and look at work sharing on seven different household tasks. We find mixed evidence: when we consider a very narrow time window around the event of birth, our estimates suggest that couples share tasks more equally in the time after birth. On the other hand, when switching to a broader bandwidth of 48 months, birth effects fade out as there are no significant estimates and, if anything, trends in the data suggest a gradual deepening of the gendered housework labor division as time passes. We relate our findings to both measured labor market outcomes and descriptive statistics on childcare and expand the analysis by running a heterogeneity check of our estimates by first splitting the sample according to geographical region and subsequently according to the relative education levels between the partners. We fail to find heterogeneous effects for the latter, while the regional analysis shows that the results are sensitive to which countries we consider. Furthermore, mothers reduce their employment and hours supplied to paid work relative to childless women. Lastly, fathers show an increased probability of labor force participation in the period after childbirth.

Acknowledgements

We would like to thank our supervisors Fanny Landaud and Vincent Somville, for their most appreciated guidance along the way, from the topic decisions to the final refinements in the days before the deadline. We are truly grateful for the suggestion of this interesting and creative topic, which was crucial for the success of our work.

Further, we would like to thank Nora Sværen for the introduction to LaTex, and her help in phase of setting up the document. We would directly like to thank Martine Kverne, Knut Kjosaas, Janice Husvaeg, Silje Halland, Reidun Halland, Umberto Maria Tomasini, Eugenia Castellazzi, Sonia Messori for proofreading and inputs on our work. Lastly, we also wish to thank friends and family for support and care along the way in the time writing our master thesis.

Contents

List of Figures

List of Tables

1 Introduction

The transition to parenthood entails challenges, choices and the need for adaptation to a new routine where the couple has to decide how to best deal with the responsibilities, advantages and drawbacks of giving birth to and raising a child. Among these, couples are necessarily faced with the dilemma of distributing household labor between partners, namely all those unpaid tasks that need to be executed in the household (from routine maintenance activities to childcare). What is easy and straightforward to understand is that, with rare exceptions, total household labor inevitably increases due to the extra time spent on childcare. However, it is less trivial to assess whether transitioning into parenthood also determines significant shifts in responsibilities over basic housework activities. Ultimately, measuring the effects of parenthood and investigating the mechanisms through which these unfold on household labor division is of economic relevance as it can produce fruitful insights for policymakers concerning the relationship between household labor division and labor force participation. Even though working women, on average, spend shorter weeks in paid employment than men do, it does not necessarily imply that they enjoy more leisure time. If housework is also considered, the seemingly short workweek becomes only a facet of a more complex situation, where women spend over twice the amount of hours, compared to their male counterparts, on various household tasks. By considering not only housework, but all unpaid work, working women face in total a longer working week than men. (Eurofound, 2007; Eurofound, 2017).[1]

Although contemporaneously relevant, the gendered division of (household) labor neither originates in nor is unique to the 21st century. Labor division, where the genders specialize in different kinds of production, can be traced back to pre-industrial times. As an example, societies originally practicing plough cultivation, an agricultural technology deemed more physically demanding, typically developed a more gendered division of labor where men are the breadwinners and women spend most of their time taking care of the household.[2] Alesina, Giuliano, & Nunn, 2013 show that the type of agricultural technology adopted by the society's ancestors can explain societal variations in female labor force participation and gender role attitudes observed today. As the observed division of labor can be partly connected to culturally persistent beliefs about the appropriate role of women in a society, it is not clear how gender role attitudes shape the effects of parenthood on the distribution of housework between partners.

[1] Unpaid work typically includes both housework and caring for children and adults (Eurofound, 2007). Notice that the survey is representative solely of men and women in the labor force and does not consider outcomes for those who are not employed.

[2] As opposed to societies that adopted shifting cultivation, a more labor-intensive practice that entailed the exploitation of handheld tools like the hoe or digging stick and that, as a consequence, was not as physically challenging. Societies of this type are more likely nowadays to have an equal labor division.

On another note, mothers spend twice as many hours in unpaid work compared to also coupled, but childless women (Eurofound, 2012). This naturally raises the question of whether childcare is the only responsible for the increased time spent on housework for women or if shifts in the division of basic tasks (that might originate from the transition into parenthood) also contribute to this phenomenon. The existing literature has already given several answers to these questions over the years but clear consensus has not been reached. On the basis of these motivations our thesis aims at answering the following research question: *Does household labor specialization evolve as a consequence of childbirth? and if so, how?*

Blaming the sharp increase in unpaid working hours for women after parenthood to child responsibilities, can be a bold extrapolation, as couples, in most cases, select into the role of parenthood. As preferences regarding having children can differ, there are strong reasons to believe that these couples are also unique in many other aspects determining household labor division. This calls for a setting to compare couples, similar in all other aspects except the parenthood status. As randomizing couples into parenthood seems neither feasible nor ethical, we exploit variation in the timing of parenthood to implement a Regression Discontinuity Design (RDD) inspired method. Considering a cross-section of households from the Generations and Gender Survey, we compare couples that have just given birth, to couples that are just about to.[3]

The thesis is organized in the following manner. Section 2 goes over the relevant existing literature and introduces the theoretical framework of the analysis. Section 3 follows up by describing the data at hand and the main restrictions that lead us to the sample we ultimately work with. Section 4 describes the empirical approach we base our results on. Findings are presented in Section 5 together with robustness checks and tests for the validity of our analysis. In Section 6 we discuss results, validity and limitations of our analysis and further expand the basic model. Section 7 concludes.

2 Literature Review

In this section, main findings on the effects of childbirth on household labor division are presented, together with the former theories that are usually called upon when testing mechanisms and assumptions.

The literature relating parenthood to shifts in household labor division, generally points

[3]We only utilize the panel dimension of the dataset in order to identify households that are about to give birth.

towards an increase in the time devoted by mothers, while finding non-significant (Sanchez & Thomson, 1997; Campolo & Rizzi, 2016; Kühhirt, 2012) or even adverse effects for fathers (Pollmann-Schult, 2017; Yavorsky, Kamp Dush, & Schoppe-Sullivan, 2015). In total, these simultaneous shifts result in a more gendered division of household tasks after the start of parenthood. Importantly, household labor division of regular chores is intertwined with childcare responsibilities for parenting couples. In this sense, Baxter, Hewitt, & Haynes, 2008 represent the above-mentioned body of literature nicely: the authors' primary purpose is that of investigating, from a life course perspective, how gendered household labor division of basic traditional household tasks is affected by both the transition into marriage and parenthood. Through random effect estimation procedures and by relying on exchange bargaining and gender perspective theories (described at the end of this section), the authors find that men's housework patterns are generally stable and insensitive to both marriage and parenthood, while women are profoundly affected by childbirth and significantly increase the number of hours spent on standard housework. On the other hand, other studies like that of Gjerdingen & Center, 2005 maintain that there is no significant increase in the number of hours worked on basic household tasks. This is the case for the mentioned study that finds no changes in the partners' shares of housework with the only exception of an increase in time spent on cleaning up after meals for mothers. The main reason why housework increases according to the authors is found in the new childcare tasks and responsibilities that are mainly taken up by mothers, rather than a shift in shares of standard housework. Dribe & Stanfors, 2009 assess how the presence of a young child affect the time allocation of mothers and fathers. The authors consider cross-sectional data and compare how their results change from the early 90s to the 2000s. For couples observed in the 90s, the authors find parenthood to strengthen the traditional division of household labor. Comparably, the effects seem to differ quite substantially after the ten years, as fathers in this cross-section now change their time use in housework and childcare in a similar way as mothers. As parenthood in the early 2000s, could not to the same extent as before be blamed for the gendered division in household labor, this study further motivates the relevance of our thesis, as the effects of childbirth on household labor are arguably not static phenomenon, requiring research as time goes and societies change.

When interpreting results, most of the mentioned papers also look at labor market outcomes in an attempt to describe the relationship between paid and unpaid work. Consistently with the statistics mentioned in Section 1, these papers describe an inverse relationship between market work and household labor, after the birth of a child, were more specifically the mothers reduce the number of hours they spend in paid work, while fathers generally face no significant effects on their work supply after childbirth (Sanchez & Thomson, 1997;

Campolo & Rizzi, 2016; Kühhirt, 2012; Pollmann-Schult, 2017). In the literature relating the earning gap between men and women to the introduction of parenthood, Kleven, Landais, & Søgaard, 2019 find reductions in working hours to be one of three driving channels of the long-run "child penalty".[4] The goal of their paper is not to estimate unequal pay for equal work, but rather consider possible firm and occupation choices as a consequence of childbirth. This in contrast with more traditional decomposition studies where the gender wage gap is decomposed into what can be explained by education, occupation, and firm choices, and what part of the earnings gap is left unexplained (Blau & Kahn, 2017). Building on the same event study methodology Kleven, Landais, Posch, Steinhauer, & Zweimüller, 2019 find a "child penalty" ranging between 20% and 60% for the six European countries considered in the study. The authors point to gender norms as a possible explanation for the observed heterogeneity, as they find a high correlation between the magnitude of the penalty coefficients and the specific country's gender norms regarding working mothers. Where these papers strongly suggest that women are penalized in their careers due to taking up the primary responsibility of providing care for their children, they do not proceed further into assessing how this burden plays out within the household. Given that these papers are closely related to the objective of our analysis in terms of focus and methodology, our main contribution pushes into the direction of expanding on their findings by investigating whether uneven burdens on genders in terms of household labor division are caused by the birth of a first child, in the light of gender unequal labor market outcomes.

When it comes to the theoretical frameworks which authors resort to for identifying potentially generally valid mechanisms, a practical distinction can be made between gender-neutral economic theories and perspectives that instead rely on the gender ideology. A gender-neutral theory is introduced by Becker, 1981. In his treatise, Becker dedicates the second chapter to the description of household labor specialization. The model builds on the concepts of comparative advantage, human capital investment, biological differences between the sexes and their commitment to childbearing. The theory predicts that household labor division occurs on the basis of an optimization criterion: ideally, household members allocate time on the market or the household to maximize the output (or utility). Following this idea, and based on the comparative advantage that they might have, household members would also invest in household or market capital. The biological differences in the commitment that both sexes necessarily experience towards childbearing are accounted for as Becker states that, even considering similar investments in human capital, the time of women is not considered a perfect substitute for that of

[4]Child penalty here refers to the long-run earning gap between men and women that can be caused by children.

men when it comes to childcare, considered an integral part of the household activities. The allocation of labor is then the result of economic and biological factors: historically persistent differences in preferences and needs experienced by men and women, as well as the inclination to develop a different set of skills, might explain the predominant gendered division of labor in the household. In the most recent context of increased gender equality, some of these assumptions might sound anachronistic. Nevertheless, the concept of specialization, optimization, and that of comparative advantage are still compelling theoretical tools with which empirical results can be assessed and validated.

Another theoretical framework according to which relationships are purely economic in nature is that of exchange bargaining. Partners in the household bargain and exploit the resources they have to strike the best deal (Pollmann-Schult, 2017). The deal usually results in an equilibrium where the primary breadwinner provides economic support to the dependent partner who, in turn, exchanges household production, *irrespective of their sex category*.[5] Brines, 1994 develops different specifications of the model and takes into account the fact that socio-institutional arrangements are likely to result in 'limited structural opportunities' for women compared to men and that this might result in women being more prone to occupy the dependent position in this theoretical framework. According to this theory, the more a dependent partner relies on the provider of economic support, the more they will engage in household labor production, and more so when the woman is the dependent partner in the couple. Several studies focus on this theory, like Sanchez & Thomson, 1997, who considering a panel of couples from the U.S. and the relative income of the partners, find that women economically dependent on their partners before transitioning to parenthood, have stronger positive increases in the time allocated to housework. Campolo & Rizzi, 2016 find similar results by considering a panel of Italian dual-earner households. The post-birth effects for men are here negligible both for paid and unpaid work as opposed to women, where weekly paid work decreases by 17 hours, and domestic labor increases by 20 hours. This theory is compelling as the main arguments used to explain household labor division are intuitive and economically rational. However, there is a competing theory that provides different reasoning as to how the household labor specialization dynamics unfolds. Brines refers to this theory as that of "gender display".

Gender display is a concept mainly defined and discussed in the field of sociology (West & Zimmerman, 1987; Goffman, 1976). It has nevertheless been exploited in economic papers (Dribe & Stanfors, 2009; Kühhirt, 2012; Kim & Cheung, 2019) to assess results related to household labor division from the point of view where gender is at the heart

[5]As described in West & Zimmerman, 1987, sex category is the category, identified through socially-agreed-upon biological criteria, one claims membership in.

of the interpretation. Brines, 1994 characterize this perspective as one where household division has a dual purpose: in addition to the production of home goods and services, it contributes to producing gender. More specifically, the author states that by (not) taking part in the household labor production, partners display features of the gender they identify themselves with.[6] According to this reasoning then, and conditional on the societal belief that providing for the family is the main work of a man while taking care of the household that of a woman, one would expect men (women) to substitute from (towards) household labor to (from) paid work, in an attempt to display their masculinity (femininity). More broadly speaking, gender ideology is arguably a relevant factor in shaping the household labor division process: S. N. Davis & Greenstein, 2009 propose a research review on the origins and consequences of gender ideology and mention several studies that identified effects on household labor division. There seems to be a consensus, for example, on less traditional men taking up a more significant share of the household tasks.

When considering both the exchange bargaining and the gender display perspectives, there is empirical evidence that rather than being mutually exclusive, they might in reality complement each other by suiting different situations. The perspectives would produce aligned expectations whenever a household with a male breadwinner and a female dependent partner are considered. Indeed, any imbalance in household labor division could be explained by means of economic dependency or gender display, just as well. However, when considering unconventional unions, namely dual-earner or female-breadwinner couples, things are different. In these cases, the exchange bargaining perspective would predict, given its gender-neutral design, that men would take up a more significant share of the household labor, at least to some extent (depending on the socio-institutional settings that the households face). On the other hand, the gender display perspective would predict men and women to act in a way that better helps them displaying their gender. In other words, men would still be incentivized to avoid any involvement in the household production to correctly signal their masculinity (even more so, considering that their gender identity would be threatened by the non-standard financial structure of the household, where women earn more than men); women, on the other hand, would increase household labor production to compensate their "unladylike" behavior. Many studies make use of both types of theories. For example, Bertrand, Kamenica, & Pan, 2015 look at relative income in the household and argue for it to be, together with gender norms, a determinant factor that can contribute to explaining the different outcomes in terms of labor production, labor force participation and the probability of marriage and marriage preferences. They analyze the relationship between relative income and social norms

[6] Grunow, Schulz, & Blossfeld, 2012; Berk, 1985.

and hypothesize what could be the interplay that characterizes the two factors. Given the belief that "if a woman earns more money than her husband, it is almost certain to cause problems", the results hint at the social norms penalizing women whose potential income is higher than that of the male counterpart in the household. Indeed, women who earn more than their partners or that could potentially do so, spend a significantly higher number of hours in unpaid housework than women living in households where the pattern is reversed. The authors believe that women do this in an attempt to fulfill their traditional gender roles and abate the perceived undesired dissatisfaction that comes with the female partner having the highest relative income in the couple. A previous contribution to these same views was brought by Kühhirt, 2012. In his study, Kühhirt investigates, by means of the fixed effect methodology, how outsourcing of the household production, bargaining power in the couple and gender norms affect time allocation in West Germany. The main findings point again at the gender norms being the main factor in which time allocation decisions orbit around. The reduction in hours of paid work in dual-income or female-breadwinner households is faster than that experienced by women in male breadwinner families, but there seems not to be a substantial impact of relative income on the observed outcomes, as is the case instead for Bertrand et al., 2015. The study concludes by maintaining the position that none of the three factors (relative income, bargaining power, outsourcing of household production) can counteract the effects of social norms and alter the typical gender roles and behaviors that these entail. Although most studies on childbirth and labor division draw on evidence from either Europe and the US, Kim & Cheung, 2019 analyze a panel of Southern-Korean husbands and wives in order to gather evidence from Asian societies and produce results in line with theories regarding gender roles, household specialization, and exchange bargaining. By resorting to a fixed-effects methodology, they show that the pre-birth gendered division is further magnified as wives spend far more time on household labor than their husbands, leading to an even more pronounced inequality in household labor division. The authors further investigate whether the spouses' employment status can moderate the results. By conditioning the birth effect on the employment status, the authors find that employed wives dampen the effect slightly, while no such impacts are in place for men.

3 Data Description

We will in this section present the data we consider in our analysis. In order to answer to our research question, we need basic information on couples and their parental history, household characteristics, socio-economic background. Having this information allows us to properly assess the effect of birth on household labor division. We need to work

on either a sample of couples observed over time, while they transition into parenthood, or a sample composed by similar couples that only differ in their parental status. A panel data structure can best fit this requirement as, by definition, it gathers data on the same observations at different points in time. The number of observations needs to be high enough to make sure that even after applying potential restriction, the remaining sample has a size sufficient enough for us to produce significant and precise estimates and consequently credible causal claims.

3.1 The Generations and Gender Survey

For the purpose of this thesis, access to The Generations and Gender Survey (GGS) data was obtained.[7] The survey serves as the core data source for the Generations & Gender Programme, a project developed within the work of the United Nations Economic Commission for Europe (UNECE), that aims at providing insights and information for research on family dynamics and relationships (Generations & Programme, n.d.). The GGS data is well suited for the scope of our analysis as it contains around 170,000 observations on individuals with age ranging from 18 to 79 years old. The survey is implemented multiple times, in different waves: so far, only two waves of data are available for most countries, while the launch of a new round of data collection is scheduled for 2020. In each country, a new wave is carried out at a three-year distance from the previous one. It allows for cross-national comparability as the same survey was carried out in all participating countries according to the same general guidelines. Moreover, the dataset has a longitudinal design, which enables researchers to follow households through time and potentially investigate several household dynamics. It comprises a multitude of variables that allow for the analysis of both macro- and micro-level information, related to fertility, pregnancy, paid and unpaid labor, childcare, preferences, opinions and values (Generations & Programme, 2019). Lastly, and importantly, the survey asks every household who is the main responsible for specific household tasks. Seven of them are considered: preparing daily meals, doing the dishes, shopping for food, cleaning the house, small repairs in and around the house, paying bills and keeping financial records, organizing social activities. Respondents can indicate who is the main responsible for the above mentioned tasks, choosing from the following options: always the respondent, usually the respondent, the respondent and their partner equally, usually the partner, always the partner, always or usually other people in or outside the household. Noticeably, the relative way in which information is provided in terms of labor allocation might constitute a significant advantage

[7]This paper uses data from the GGS Waves 1, and 2 (DOIs: 10.17026/dans-z5z-xn8g, 10.17026/dans-xm6-a262), see Gauthier, A. H. et al. (2018) or visit the GGP website (https://www.ggp-i.org/) for methodological details.

for our analysis compared to other studies where the amount of work is expressed in absolute terms (for example number of hours). The latter usually encounter problems related to the impossibility of controlling for multitasking and the incapacity to grasp those nuances concerning the dynamics that characterize the labor allocation between partners or its delegation to third parties (Gjerdingen & Center, 2005; Kühhirt, 2012; Balbo & Arpino, 2016). Furthermore, having information on the execution of these different tasks in relative terms enables us to implement an analysis at the household level rather than having the individual set as the focus of the research. A potential drawback entailed by this type of measurement is, however, that the relative data is provided by only one of the household member and the true shares of housework might be biased by the particular perspective of the respondent. The relative measure of household labor division is set in terms of a Respondent-Partner frame: we will further discuss how switching to a Man-Woman framework will ease our analysis in the following section, as this relates strongly to our identification strategy.

3.1.1 The sample of interest

We consider the data gathered through the Generations and Gender Survey from, in total, 10 European countries, selected based on their repeated participation to both waves, and whether they have obtained the required data on births and children necessary to execute the analysis. The sample is then restricted to those households that are surveyed in both consecutive waves.[8] It includes individuals for whom data is available with regards to the year and month in which they give birth (if they do so), are surveyed, and move in together. We consider couples where both the partners are at least 21 years old, and the woman is below 45 years. We keep households with couples that are surveyed within the same time window for each country, before or after their first birth. This time window is determined by the maximum country-specific intra-wave distance, (Poland and Austria having the widest intra-wave time gap of 51 months). To produce meaningful interpretations of the outcome variables, we only keep couples who have lived together (irrespective of marital status) for a period of time, before childbirth, at least as long as the minimum country-specific intra-wave time gap (Bulgaria having the shortest one of 29 months). [9] The Regression Discontinuity (RD) analysis carried out in Section 5, is focused on the birth of the first child. Thus, as previously mentioned, we consider households that get their firstborn between waves, or not earlier than four years before the first wave. We consider biological, adopted and foster children. We require them to reside at the

[8] We eliminate households that only have a single observation from either the first or the second wave.

[9] For further explanation on why we carry out this sample restriction procedure consult the Appendix, section A.1.1.

household at the time of the survey. There are, however, no restrictions on the presence of other relatives of the couple, like parents and siblings, in the household.

4 Empirical approach

In this part of the thesis, arguments for the choice of methodology are provided and, subsequently, the empirical approach is explained in terms of main equation, key variables and assumptions.

4.0.1 Choice of methodology

When it comes to our specific research question and the scope of this analysis, we would ideally aim at obtaining results characterized by internal validity as well as broad external validity. A satisfactory outcome would be that of finding significant results and identifying mechanisms able to explain the dynamics and patterns that we might find by analyzing the data at hand, consistently with the main theories related to household labor division. At the same time, we would like these mechanisms to be valid even when taken out of the specific context where they have been identified. As mentioned in the introduction, a researcher would typically resort to RCTs (randomized control trials) to efficiently control for all factors that cannot be accounted for in a natural experiment or more generally in a real-life context. However, RCTs are often faced with ethical limits as well as budget and time constraints: in our specific case, randomizing birth would easily cross all of these boundaries and consequently is not a viable option. Alternative methods need to be considered. The existing literature has made extensive use of the Fixed Effect methodology in an attempt to exploit national and cross-country panel data (Kühhirt, 2012; Bertrand et al., 2015; Schober, 2013; Pollmann-Schult, 2017; Kim & Cheung, 2019; Yavorsky et al., 2015), while other methods are also exploited such as Random Effects (Baxter et al., 2008), Seemingly Unrelated Regressions (Sanchez & Thomson, 1997; Campolo & Rizzi, 2016) and Instrumental Variables (Bertrand et al., 2015). Given the nature of our dataset and the fact that the majority of the literature has been focusing on fixed effects, we decide to implement a different approach, drawing on the quasi-experimental approach called Regression Discontinuity Design. This method can be related to the way Kleven, Landais, & Søgaard, 2019 utilize the Event Study Methodology in their paper: even though they do not focus on household labor division, they still investigate other outcome effects related to the introduction of parenthood.[10] At first sight, the event study methodology would

[10]This methodology relies on the assumption that unobserved determinants, other than the variable of interest should evolve smoothly over time.

be compelling, considering that one can follow the same individuals over time. However, a panel of more than two waves would be preferable, since one needs to observe both pre-birth and post-birth outcomes for the observed couples. By instead turning to the Regression Discontinuity Design, we can exploit couples that have already given birth at the time of the first wave and their post-birth outcomes information, without requiring to have observed them before this event occurred. Indeed, we can compare these same couples with those that transition into parenthood between waves and for which, as a consequence, both pre- and post-birth outcome information is available. This comparison builds on the assumption that couples that experience birth at different, but close, points in time should be similar and comparable. We believe that this different approach will, in the first place, allow us to effectively exploit the short panel that we have (composed of only two waves of observations). Secondly, it will contribute to the research on the topics of childbirth and household labor division by assessing whether findings resulting from the application of this methodology are consistent with those gathered by the existing body of literature.

4.0.2 Regression Discontinuity Design

The empirical approach we mainly refer to is that of Regression Discontinuity Design (RDD), which was first introduced by Thistlethwaite and Campbell (1960). The method is based on outcome comparisons of individuals just above or just below a certain cutoff point, determined by a continuous assignment variable (also called running variable). Typically, individuals to the right of the cutoff receive treatment whilst individuals to the left do not. Examples of treatment can be getting access to a top school or a welfare program, becoming of age, or getting married. The methodology is appealing for our research question in that we want to compare individuals that have recently given birth to their first child whit individuals that are about to do so. Observing these two groups should ensure comparability since we can think that preferences and other characteristics of parents and their soon-to-be counterparts are somehow aligned as opposed to a comparison of childbearing couples with childless ones. Before specifying the model equation, it is worth mentioning that our settings are somehow different from those of a traditional RD setup as we resort to this methodology due to the limited availability of waves of interviews and gathered data. Indeed, in typical RD studies, the cutoff originates from an objective rule that defines a fixed value in time or other dimensions (the running variable) that is valid and invariable for all individuals, as well as exogenous to the running variable (Jacob, Zhu, Somers, & Bloom, 2012). In our case, on the contrary, we identify childbirth as the life-changing event that makes couple transition into parenthood but that does not define

a specific point in one's lifetime, valid and invariable for everyone: partners can become mothers and fathers at different ages and different points of their life. Yet we are able to exploit childbirth as our cutoff because of the relative terms in which the running variable is defined (distance of the Wave 1 survey from birth). Indeed, it is as if we are stacking up all childbirths and figuratively think of having this event as the reference point for our analysis. When the distance is negative, couples still have to give birth at the time of the Wave 1 survey, on the other hand, when the distance is positive, the interview takes place after parents give birth to their first child. In this sense then, the analysis we implement should be considered to be an RD-like study rather than a proper RDD. On a last note, we emphasize that our goal is to capture the effects of the transition into parenthood: this necessarily implies that we want to observe couples that give birth to their first child. We believe that major changes in the labor division of standard household tasks are more likely to be experienced with the birth of a first child: when investigating the potential impact of a second child, Kim & Cheung, 2019 observe no further change in time allocation, and argue that it is only the birth of the first child that determines household labor division alterations. Regardless, this is arguably a reasonable restriction of the scope of our analysis given that if we were to consider couples with more than one child, we would not be able to disentangle the effect of the first birth to, say, that of the second or the third. This allows for a more precise interpretation of our results at the cost of a reduction in sample size, which, in turn, will result in a loss of precision of our estimates. We now define the model equation, while the correspondent results are presented in Section 5.

4.1 The model equation

In our analysis, the variable indicating treatment is $Birth_i$, which equals one if a couple gives birth before the Wave 1 interview (treatment group) and zero if the couple gives birth in the period after the Wave 1 survey and before the Wave 2 survey (control group).

The assignment variable (or running variable) is $Distance_i$. This is the difference between the month of the first survey wave W_i^1 and the birth month of the first child B_i.

$Distance_i = W_i^1 - B_i$.

The cutoff point is the particular value of the running variable $Distance_i$ at which the probability of receiving treatment drastically jumps. In particular, we implement a *Sharp* RD where the probability jumps from zero to one at the cutoff. In our case, the cutoff value is zero: this virtually corresponds to the situation where a couple gives birth to its first child in the same month of the Wave 1 survey (null distance). Nevertheless, we

exclude observations that have a zero $Distance_i$ value: since we do not have daily data, we cannot determine whether the birth happened before or after the day of the survey.

$$Birth_i = \begin{cases} 1 & \text{if } Distance_i \geq 1 \\ 0 & \text{if } Distance_i < 0 \end{cases}$$

Our coefficient of interest ρ measure the variation in the different outcome variables as a consequences of childbirth.

$$Y_i = \alpha + \rho Birth_i + \gamma_1 Distance_i + \gamma_2 (Distance_i * Birth_i) + \beta_i X_i^{'} + \varepsilon_i \tag{1}$$

The dependent variables that we include in Equation 1 are different in nature. As mentioned in the previous section, we have data on seven tasks and on who is mainly taking care of them, stored in seven variables. We construct new variables so that there are four different dummies, for each one of the tasks, that turn to 1 when either the man, the woman, them equally or someone else (which includes both third parties living in the household as well as external individuals) is the main responsible, respectively. The labor division dummies show the probability of, for example, the man in the household always or usually doing the dishes, as opposed to the woman, the partners equally or someone else. We extrapolate information from these new variables and reorganize them in task counts for the four alternatives. In Section 5 we provide tables and graphs that show how these task count variables are affected by childbirth. We also provide results for the individual household tasks and labor market outcome in the discussion (Section 6).[11] As we mentioned in the data description, these variables collect relative housework shares from the sole perspective of the respondent: we draw on the findings of Yavorsky et al., 2015 that compare the parenthood effects on paid, unpaid and childcare labor assessed by means of the analysis of both time diaries and survey response measurements, to critically handle results related to these self-reported and possibly gender-biased outcomes. Indeed, when considering time diaries, the gendered division of housework evolves such that women perform more housework than men. By considering highly educated dual-income couples, the authors find that men decrease the time they spend in housework, possibly in order to make up for the engagement in new time-consuming activities of physical childcare. Women are burdened with the majority of childcare responsibilities, but this does not affect their housework engagement. Comparably, the authors do not

[11] Given that, for some household, we have information for some of the tasks and not for others, but we are still interested in single tasks, we do not drop these households and will only consider observations with a complete set of information when estimating task count outcomes.

find evidence for gender inequalities in outcomes when they consider the survey response results, possibly suggesting that the parents do not perceive their labor division as unequal. This, importantly, serves as a reason for being cautious when implementing research that relies on such measurement methods, as the identification of any gender unequal result might depend on the individual perception of the subjects studied.

The interaction term between the treatment dummy and the running variable, $Distance_i * Birth_i$, allows for different slopes on either side of the cutoff: this is recognized as being one of the standard specifications employed in many RD studies. Compared to the specification where the slope coefficients are imposed to be the same on either side of the cutoff, our choice avoids this restriction thus preventing observations from the control to affect estimates for the treatment and vice versa. (Imbens & Lemieux, 2008; Angrist & Pischke, 2015)

Lastly, X_i represent a set of control variables. Theory tells us that one should not expect the covariates to be of any relevance: as in randomized control trials, these should not influence the magnitude of the local treatment effect if this was truly as good as randomly assigned in the RDD. Nevertheless, it is common practice in many RCTs and applications of the RD methodology to account for covariates as this might result in improved precision of the estimates (Lee & Lemieux, 2010; Calonico, Cattaneo, Farrell, & Titiunik, 2019; Imbens & Lemieux, 2008). Furthermore, by controlling for covariates we can run a sensitivity check of the main results: consistent estimates that do not show important changes, once the covariates are added to the model, are considered to be good evidence of the suitability of the identification strategy to the research question (Lee & Lemieux, 2010). Also, in some cases, covariates might be controlled for in an attempt to restore the validity of the analysis, should those covariates exhibit an imbalance around the cutoff (Calonico et al., 2019; Imbens & Lemieux, 2008). In Section 5, we first present results for the regression without controls but switch to a specification that considers covariates right afterward. We include controls at the individual and household levels. For both sexes, we add a variable measuring age and a dummy that equals 1 when the individual has achieved tertiary education (and 0 otherwise). We include a dummy variable equal to one for western countries and zero otherwise: we validate the grouping procedure by both looking at how the countries in our sample are traditionally split according to their spatial location as well as how egalitarian, on average, their populations are. Given the question *To what extent do you agree with the following statements? A pre-school child is likely to suffer if his/her mother works*, we sort countries into egalitarian and traditional based on the country share of people that self-reported to be in agreement with the statement, when considering the unrestricted full sample. We observe a quite neat grouping: Austria, Czech Republic, Germany and France have the smallest shares - never higher than 51 percent -

thus being the egalitarian countries, while Bulgaria, Georgia, Hungary, Lithuania, Poland and Russia have at least 61 percent of their population agreeing.This grouping procedure is exploited later on in the analysis in an attempt to capture heterogeneous effects of childbirth. In addition, two dummies control respectively for sex of the respondent (and thus any gender-driven bias in survey answers) and subjective gender role attitudes, while their interaction captures gendered variation in social-norm and gender-role views in the household. In particular, the sex dummy turns to 1 when a woman answers the survey, to 0 when a man answers. Noticeably, the dummy controlling for gender role preferences builds on the basis of the same survey question that we use to sort countries into groups, yet it directly focuses on the individual answer given by the respondent without referring to a country average.

We make use of a triangular kernel and construct weights for the entire sample: this is a common procedure to localize the regression fit by attributing more weights to the observations that are closer to the cutoff. (Calonico et al., 2019)

4.2 Assumptions

The Regression Discontinuity Design hinges on two main assumptions. In this subsection, we will discuss them and argue whether these are likely to hold or not. The formal testing of the assumptions will be carried out in Section 5.4.3

4.2.1 No precise manipulation

The first assumption is that individuals should not be able to precisely manipulate the running variable. Mathematically speaking, this is equivalent to say that for every different type of individual, with a specific combination of observable and unobservable baseline characteristics, there is an equal probability of getting a running variable draw to either side of the cutoff. (McCrary, 2008) As described above, the running variable is determined by the time of birth B_i and the interview time of the first survey W_i^1. Manipulation will then depend on the individual's ability to manipulate either of the two determinants and also the relation between the two.

It is known that contraceptives enable women to manipulate the timing of pregnancy better. The question, however, specifies to: will the time of the survey interview affect the couple's decisions on when to have a child? The respondent, in most cases, does not know about the upcoming survey interview. It seems, therefore, very unlikely that the couple will consider the survey time when they decide to have a child. On the other

hand, and given that $Distance_i$ is defined considering two reference points (time of birth and time of the survey), it is reasonable to wonder whether the date of the interview can be postponed or the interview itself canceled (this would correspond to selecting out of the sample). When it comes to this second concern, we deem logical to think that women that have just given birth might be more prone to either postpone or directly avoid the interview due to the recent event. The money compensation for the survey (if any) is extremely small. There is reason to think that people only look at the survey as a time consuming activity. In that sense, a household with a pregnant woman might have different preferences when it comes to either having the survey before or after the birth of the baby. The survey procedure vary across the different countries. In some countries, the respondent has influence over the survey time, while in others there were no possibility to manipulate the interview. Whether the assumption of imperfect manipulation holds is something we will come back to in Section 5.4.3.

4.2.2 Continuity of all other determinants

The second assumption is that all other factors determining the outcome variable are continuous over the threshold $Birth_i$. (Lee & Lemieux, 2010). This means that ideally all other variables, which determine household labor division, should evolve smoothly around the cutoff point. If also other variables jump at the threshold, then the coefficient of interest ρ, can be biased. In particular, the uptake of parental leave will be a factor possibly determining household labor division, which is not continuous over the cutoff. As mentioned in Angrist, 2009, controls that are pre-treatment variables should not jump at the cutoff, precisely because of their nature (variables whose value is pre-determined, irrespective of the treatment). To investigate whether this assumption holds, in Section 5.4, we will test if baseline covariates are continuous as we move over the cutoff.

4.3 Choice of bandwidth

In RD studies, a core step of the whole procedure is that of choosing the bandwidth. The RDD has limitations when it comes to external validity: at best, they can only provide an average effect for the subpopulation at the cutoff (Imbens & Lemieux, 2008), as the assumptions on the treatment being as good as random can only be valid in this area. In this sense it is then necessary to identify the correct bandwidth for which meaningful results can be obtained. When choosing the bandwidth, it must be considered that there will be a trade-off in terms of bias and precision: when we adopt a bigger bandwidth we will increase the precision of the estimates as we will include more observations, at

the same time, though, we are considering units that are further away from the cutoff and for these the assumptions might not hold anymore (in other words, units this far from the cutoff on one side of the threshold might be systematically different from those on the other side at the same relative distance). We rely on the methodology developed by Calonico, Cattaneo, Titiunik and Farrell to identify the optimal bandwidth for the different outcome variables. We both consider the task count variables as well as the single outcome variables for the seven chores and consider specifications with and without controls. We average out the suggested bandwidth for the two groups of variables and for both we approximate to the closest integer, which happens to be 13 months, so about one year around the cutoff (birth).

5 Results

Here, we first present descriptive statistics relevant to the discussion of our findings. Consequently, we provide table and graphical results from the analysis, followed by a series of robustness checks, pertinent to the assumptions of the method we base our analysis on.

5.1 Descriptive statistics

Averages, decomposed by the parenthood status of the couples are presented in Table 1. Panel A display average outcomes for the number of tasks each potential household member execute always or usually. There are seven potential tasks that the household can execute in total.[12] Firstly, across the two samples, women generally tend to be taking on a higher number of tasks compared to men. Sharing the tasks equally is also frequent as opposed to someone else taking on the tasks, which is the least likely scenario occurring. Secondly, there are some differences between the samples, mainly by the fact that parent couples share fewer tasks equally, accompanied by women and men executing a higher amount of tasks on their own. Lastly, the increase in the amount of tasks for women is markedly higher than that of men.

Panel B show relevant characteristics describing the households and their members. Across the sample, we can see that there are more female than male survey respondents. The women are about two years younger than their male counterparts and also more likely to have obtained tertiary education. Over half of the respondents hold traditional views

[12]Table 16 in Section 6.2.2 present a full overview of the seven household tasks in the analysis.

about working mothers and live mostly in eastern European countries. Across the panels, childless couples are more likely to have a female respondent as opposed to the parents. The parent couples are generally a few years older which is in line with the intuition that older couples are more likely to have children. Lastly, among the parents men are more likely to have tertiary education. The parent couples are more likely to reside in western Europe as opposed to the non-parent couples.

Table 1: Descriptive statistics

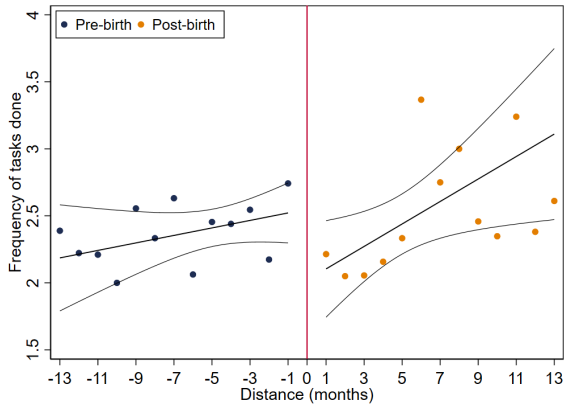
	Non-Parent sample		Parent sample	
Panel A: Average task count				
Woman	2.22	(0.06)	2.84	(0.04)
Man	1.39	(0.04)	1.46	(0.03)
Equally	2.89	(0.07)	2.26	(0.05)
Someone else	0.50	(0.05)	0.44	(0.03)
Observations	636		1171	
Panel B: Average characteristics				
Female R	0.58	(0.02)	0.56	(0.01)
Age - Woman	27.16	(0.15)	29.01	(0.13)
Age - Man	29.89	(0.19)	31.61	(0.15)
Tertiary education - Woman	0.46	(0.02)	0.38	(0.01)
Tertiary education - Man	0.35	(0.02)	0.32	(0.01)
Western Europe	0.35	(0.02)	0.34	(0.01)
Traditional	0.56	(0.02)	0.53	(0.01)
Observations	652		1214	

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Panel A present average outcomes for the amount of tasks a certain person within the household (man, woman, man and woman equally, someone else) is always or usually the main responsible for. The table is split into averages presented for parent and non parent couples separately. Panel B show average characteristics where Female R is the fraction of households where there are a female respondent to the survey. Tertiary education is the fraction that has obtained at least tertiary education. Western Europe show the fraction of households living in a Western European country and traditional is the share of respondents agreeing with a statement regarding working mothers.

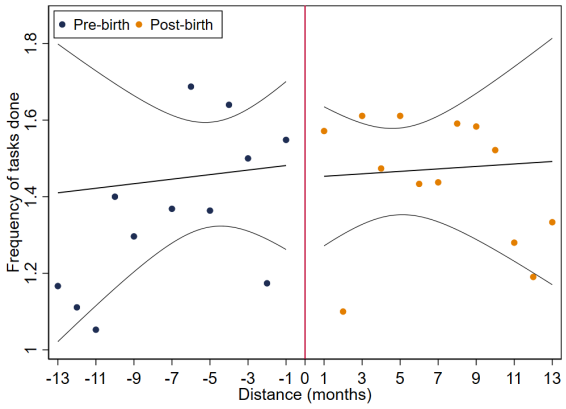
5.2 Graphical results

Figure 1 shows the RD graphs for the number of household tasks taken on by each household member around the event of birth. Both Figure 1a and 1c show shifts around the threshold, but like the other panels, they are not significant at a conventional level, as the confidence bands overlap. Both of these panels have strong trends, Figure 1a shows a upward sloping trend, as opposed to Figure 1c, which has a downward sloping trend.

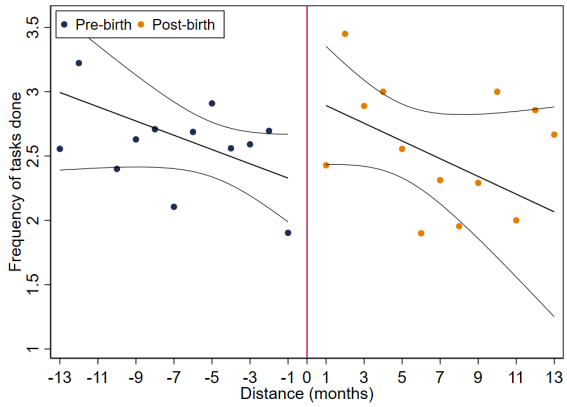
Figure 1: RD estimates. Frequency of household tasks



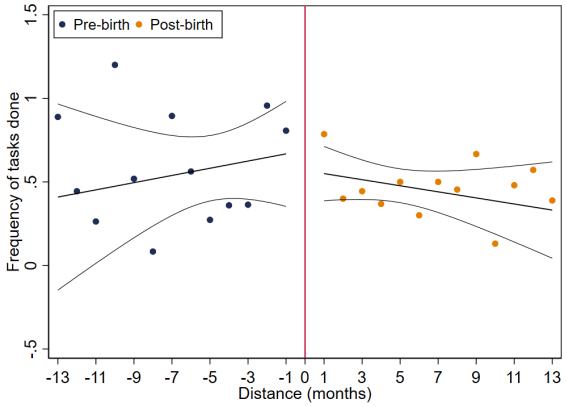
(a) Woman always or usually



(b) Man always or usually



(c) Partners equally



(d) Someone else always or usually

5.3 Regression results

Table 2 display each potential household member in different columns. The content show estimates for the effect of birth on the number of household tasks usually executed. The estimate in column 1 is significant at the 5% level: the woman executes, on average, half a task less after birth compared to women that still have to give birth. The positive slope

both seen in Figure 1a and by the positive ($Birth * Distance$) coefficient in the table indicate a gradual increase in outcomes. Secondly, the coefficient in column 3, significant at the 1% level, indicates that couples introduced to parenthood share approximately 0.8 more tasks equally between themselves as opposed to couples that still have not had their first child. Similarly, as in Figure 1c the mean gradually decreases seen by the negative post-birth slope. The regression results presented in the table is expanded with individual and household characteristic controls in the following subsection.

Table 2: RD estimates. Frequency of household tasks. No controls

	(1)	(2)	(3)	(4)
	Woman	Man	Equally	Someone else
Birth	-0.536**	-0.062	0.797***	-0.200
	(0.254)	(0.176)	(0.269)	(0.211)
Distance	0.028	0.005	-0.066**	0.032
	(0.030)	(0.021)	(0.032)	(0.025)
(Birth x Distance)	0.058	0.000	-0.013	-0.046
	(0.044)	(0.031)	(0.047)	(0.037)
Observations	547	547	547	547

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The outcome variables tell for how many tasks a certain person within the household (man, woman, man and woman equally, someone else) is always or usually the main responsible for. The regression makes use of a triangular kernel. The control variables $Distance_i$ and ($Distance_i * Birth$) are included in the regression consistent with the RD model. Additional control variables are not included in this specification.

5.4 Robustness checks and internal validity

We now test whether the results are robust to different model specifications as well as whether the fundamental assumptions of the methodology are likely to hold. We do this according to the procedures suggested in Imbens & Lemieux, 2008. More specifically, we first observe how results change when adding covariates and test whether these are continuous or jumpy at the threshold. Further discontinuity tests for the density function of the running variable are implemented. We check how results behave when changing the bandwidth, and when switching to a rectangular kernel.

5.4.1 Adding covariates

Motivated to validate our results further, we include a set of control variables in line with our main specification introduced in Section 4.1. Immediately evident is the loss of significance of the birth coefficient for women in column 1. By controlling for relevant characteristics, both the effect for women and the couples equally is now moderated in terms of magnitude. The inclusion of control variables does not improve our estimates in terms of precision, and does if anything increase the standard errors slightly.[13] The effect of parenthood on the number of tasks the partners share equally still shows a downward trend in the months around the birth event. There are, however, still a discontinuous jump at the threshold, showing that couples just about to give birth share significantly fewer tasks than their parent counterparts that have just given birth. In Table 14 in Appendix A.4 we further looking into the specific tasks that can drive the results. We find that more equal sharing of grocery shopping, vacuum cleaning and record keeping activities are identified as the main drivers for these aggregate short term effects.

[13]The only estimate which improves in terms of precision is the potential outcome of someone else taking on the task.

Table 3: RD Estimates. Frequency of household tasks. With controls

	(1) Woman	(2) Man	(3) Equally	(4) Someone else
Birth	-0.482*	-0.306*	0.746***	0.042
	(0.261)	(0.176)	(0.270)	(0.208)
Distance	0.020	0.028	-0.057*	0.009
	(0.031)	(0.021)	(0.032)	(0.024)
(Birth x Dist)	0.065	-0.013	-0.035	-0.017
	(0.044)	(0.030)	(0.046)	(0.035)
Age - Woman	0.033	0.001	-0.009	-0.024
	(0.023)	(0.016)	(0.024)	(0.018)
Age - Man	-0.022	0.034**	0.030	-0.042***
	(0.020)	(0.013)	(0.020)	(0.016)
Tertiary education - Woman	-0.334**	-0.038	0.378**	-0.007
	(0.160)	(0.108)	(0.165)	(0.128)
Tertiary education - Man	-0.091	0.337***	-0.044	-0.201
	(0.169)	(0.114)	(0.175)	(0.135)
Western Europe	-0.204	0.007	0.745***	-0.549***
	(0.151)	(0.101)	(0.156)	(0.120)
Female R	0.246	-0.437***	-0.077	0.268
	(0.212)	(0.142)	(0.218)	(0.168)
Observations	547	547	547	547

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. As described in Section 4 we examine the effect of birth on a set of variables that describe how household tasks are split among the household members. In particular, for every household, these outcome variables tell for how many tasks a certain person (man, woman, man and woman equally, someone else) is always or usually the main responsible. Coefficients for all the covariates are included. The regression makes use of a triangular kernel. We include controls for the male and female household members age, whether they have obtained tertiary education or not and whether the sex of the respondent. Controls for whether the respondent reside in a western European country. A binary variable describing whether the respondent agrees with a traditional gender role views and a interaction this variable and the sex of the respondent is also added, but not included in the table. Birth is the coefficient of interest, that should capture the effect of giving birth on the number of tasks the specific household member is responsible for. The control variables $Distance_i$ and $(Distance_i * Birth)$ are included in the regression consistent with the model presented in Section 4.1.

5.4.2 Discontinuity test for the covariates

Contextually to the exhibition of the new regression results from the expanded model, we inspect whether the covariates representing baseline characteristics of the observed households shift discontinuously at the threshold. If this is the case, the second RD assumption mentioned in Section 4.2.2 might be violated. We provide both regression estimates and visual results to assess the continuity of covariates at the threshold.

As a first falsification check, we regress each covariate on the basic model introduced in Section 4. If the covariates are balanced and smooth at the threshold, we should not find significant results for the $Birth_i$ coefficient. Table 4 presents the estimates resulting from the above-mentioned regressions. While coefficients are not statistically different from zero for female and male age, female education, country grouping, and gender role attitudes, both sex of the respondent and male education are significantly affected by birth. The coefficients for these two covariates are significant at the 1% level, giving birth to a child would decrease the probability of having a household with a female respondent by 22.4 percentage points and would increase the probability of men in the household having achieved tertiary education by 24.4 percentage points. These last two variables might actually jump at the threshold and the households on different sides of the cutoff would not be entirely comparable. These results are virtually unchanged when considering alternative specifications for some of the covariates (particularly when we switch the dummy for traditional gender role attitudes with an index that consider an additional set of questions through which views on social norms and gender roles can be assessed).

Table 4: Test for continuity in baseline covariates

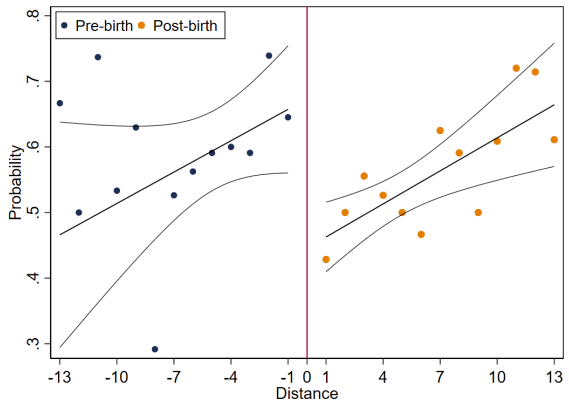
	Household level			
	Female R	Traditional	(Female R x Traditional)	Western
Birth	-0.224***	-0.063	-0.100	-0.016
	(0.079)	(0.077)	(0.072)	(0.073)
	Woman		Man	
	Age	Tertiary	Age	Tertiary
Birth	0.676	-0.080	0.159	0.244***
	(0.489)	(0.070)	(0.578)	(0.065)
Observations	547	547	547	547

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Female R" refers to the dummy for the sex of the household member responding to the survey (equal to 1 for female respondents). "Traditional" refers to the dummy used as a proxy for gender role attitudes. It switches to 1 when the respondent agrees with the following statement: *"A pre-school child is likely to suffer if his/her mother works"*, and zero otherwise. The third estimate relates to the interaction term between the two dummies mentioned just now. "Western" refers to the dummy describing whether the household resides in a western country. For both men and women the estimates for age and tertiary education are presented. Estimates for the effect of birth on the age and education covariates are included for both men and women.

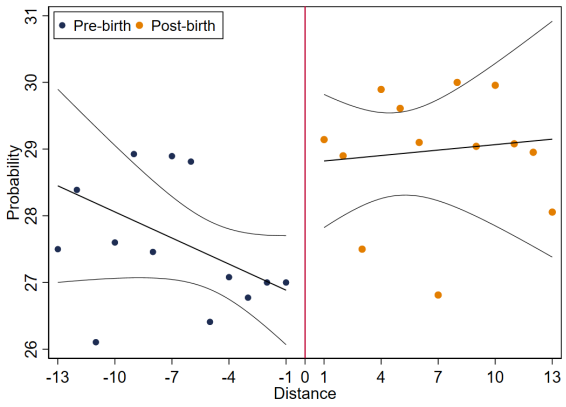
To further investigate these findings, we plot the covariates against the running variable and assess their behavior at the cutoff. Consistently with the findings presented in Table 4, there is no proof of discontinuity at the cutoff for male age, female education, the regional dummy, the dummy for traditional gender role attitudes and the interaction between this last one and sex of the respondent, while sex of the respondent itself is still jumping. On the other hand, when plotting the graph for the male tertiary education dummy, we cannot observe a significant jump at the threshold as the confidence intervals are partially overlapping. Conversely, when plotting the age of women on the running variable, the graph shows a significant jump that is not matched by the regression results. We provide here the plots for the covariates for which we either find significant coefficients in the regressions or a jump in the graphs. Figures 2a, 2b and 2c show the jump in probability of having a female respondent, female age, and the seemingly continuous distribution of the tertiary education dummy for males, respectively.[14]

[14]Graphs for the other variables included in the main regression can be found in the Appendix, A.3.1.

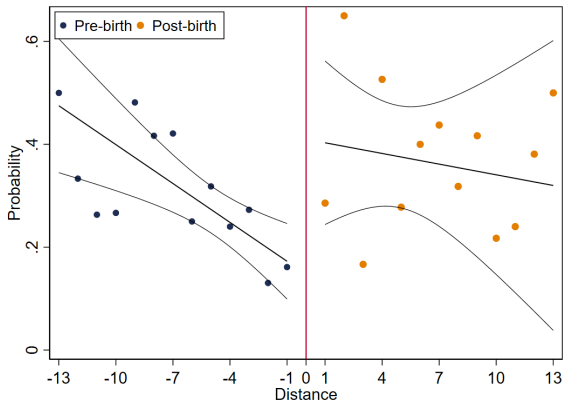
Figure 2: Discontinuity test: Female respondent, female age, tertiary education (men)



(a) Female respondent



(b) Age, women



(c) Tertiary education, men

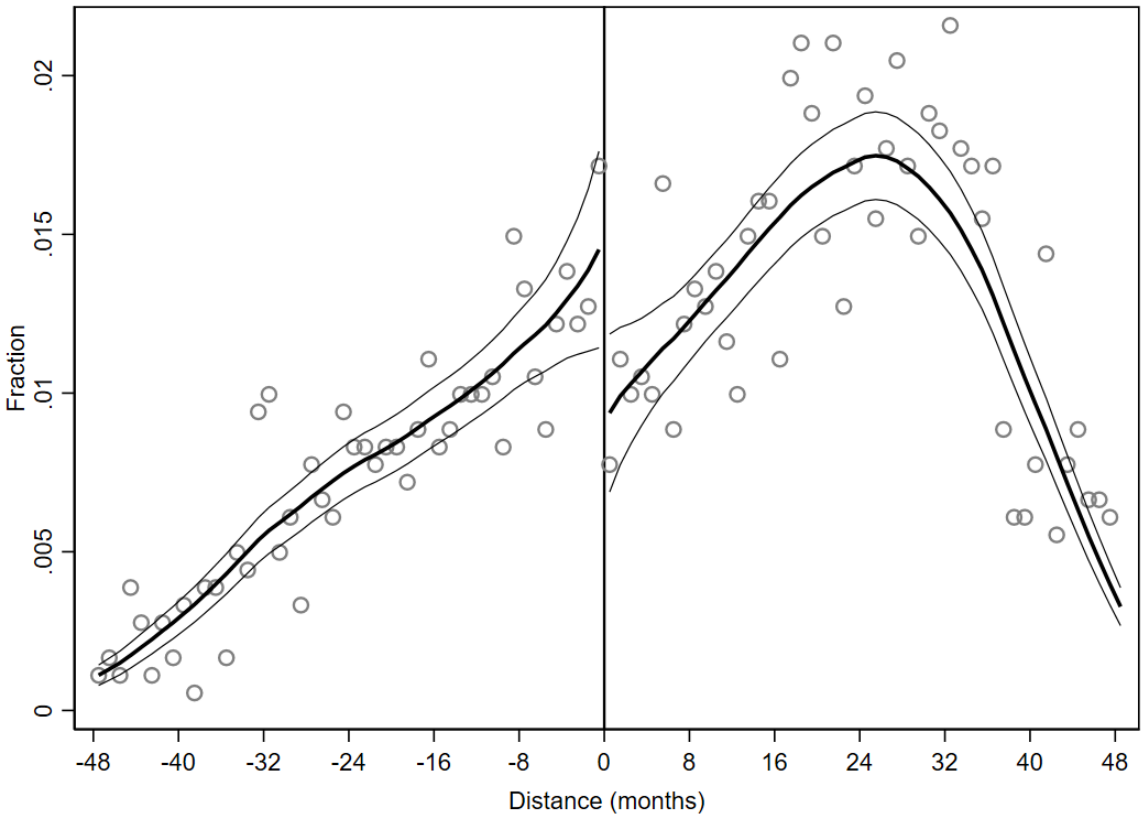
We reserve the thorough discussion of these findings and their implications on our analysis for the next section. Nevertheless, since we observe changes in the magnitude and significance of our estimates, from here on out, we consider this specification with covariates as our reference point for any additional robustness and validity check.

5.4.3 Discontinuity test for the density function of the running variable

As mentioned in Section 4, one of the identifying assumptions of the RDD is that individuals (or households in our case) should not be able to precisely manipulate the running variable at the cutoff. We have argued why this is not likely to be the case and will here formally test whether this assumption is likely to hold or not. As previously said, the local randomization assumption hinges on observables and unobservables being equally distributed on either side of the cutoff. Given that, by definition, unobservable characteristics cannot be assessed, it is not possible to directly test the assumption.

McCrary, 2008 does, however, propose a partial test that considers the density function of the running variable. The main intuition which the McCrary test builds on is that the continuity in this density function is implied by the continuity of the density function of the running variable *conditional on observables*. According to Lee & Lemieux, 2010, this last condition alone is not sufficient to ensure that the first assumption of the RD methodology is met. However, evidence of the above mentioned continuity can be utilized to state a failure to reject the local randomization assumption. Conversely, if the function is discontinuous at the cutoff this would be a sufficient indication of "bunching" of households to either side of the threshold and the assumption would be rejected. Thus, we run the test and provide results in Figure 3. As for when the main graphs are plotted, $Distance_i$ is not binned and an average is computed for every value of the running variable. We plot the graph for the entire support of the data but the test considers an actual bandwidth of 13 months on either side of the cutoff, consistently with the arguments made in Section 4. In this figure, there is no significant jump at the threshold (the confidence intervals are overlapping). The test result implies that we fail to reject the null hypothesis of continuity of the running variable.

Figure 3: Discontinuity test for the density function of the running variable



Notes: We only consider observations within our preferred bandwidth window of 13 months in the test. Even though more observations are displayed they do not affect the results of the test.

5.4.4 Sensitivity to the choice of bandwidth

One of the most common robustness checks implemented when dealing with the RD methodology is to test how the results change in relation to the choice of alternative bandwidths. We follow the suggestions produced by Imbens & Lemieux, 2008 and consider bandwidths that are half and twice the size of the original one, respectively. More specifically, we compute estimates using the same model (with controls for covariates, a 13-month bandwidth and triangular weights), considering households that are in a 6-month range and consequently in a 24-month range.

When moving to a bandwidth of 6 months the coefficient on "man and woman equally" remains significant at the 1% level: even if the standard error increase (from 0.270 to 0.442), the magnitude of the effect is now 64,3% higher compared to the initial estimate (1.226 as opposed to 0.746). The coefficient for "man" and "someone else" are still not significantly different from zero (the coefficient for "someone else" changes sign but the effect is extremely close to zero). Conversely, the coefficient for "woman" is still marginally significant: while still being consistent in sign, the magnitude increases considerably, as well as the standard error (from -0.482 to -0.797 and from 0.261 to 0.421 respectively). The coefficient for "men" is not statistically significant anymore (both the magnitude and the standard error increase but the latter disproportionately more so, compared to the former).

When considering a bandwidth of 24 months the coefficient for "man and woman equally" loses one significance level, mainly because the drop in magnitude (almost 50% of the initial value) is bigger than the improvement in precision due to the inclusion of observations that are further away from the cutoff (standard error diminishes by 32.2%). The coefficient for "woman" becomes insignificant for a similar reason, while the effect on "man" is still marginally significant at the 10% level (the intensity of the effect is reduced of approximately one third while the standard error slightly more than that). The coefficient for "someone else" is still insignificant and negative.

Overall, it seems that, by changing the bandwidth, the only result that remains consistently significant (at least at the 5% level) is that related to the probability of the task count where man and woman are equally responsible. The magnitude of this effect is however negatively correlated with the bandwidth: greater bandwidths are associated with a smaller coefficient for this task count outcome. A similar pattern can be observed for the "woman" task count, yet the estimates are generally less significant (at best, only at the 10% level). These findings could be interpreted as the results being rather sensitive to the choice of bandwidth. When it comes to the "man" task count, evidence from the

robustness check seems to be more at odds and, if anything, shows that the results follow no clear pattern with reference to the bandwidth.

5.4.5 Rectangular kernel

When adopting a different weighting system (rectangular kernel), insignificant estimates from the regression with covariates are virtually unchanged. On the other hand, the coefficient for "man and woman equally" face a drastic drop, from being significant at the 1% level to not being significant at all. Noticeably, this is partly explained by a loss of precision (as the standard error increase from 0.270 to 0.306) but also by a drop in the magnitude of this effect (which falls from 0.746 to 0.437). The coefficients for "woman" and "man" are not significant anymore.

6 Discussion

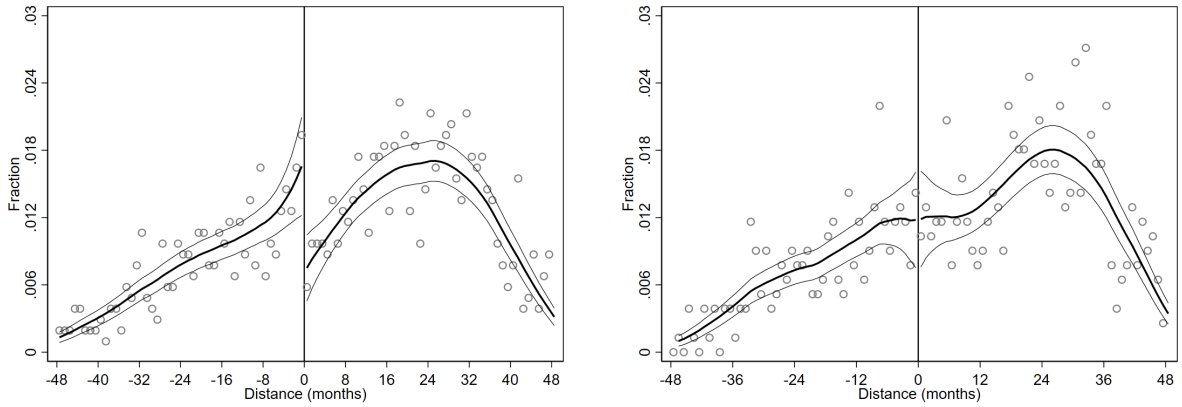
Main results and validity checks are now discussed. Furthermore, we propose expansions of the main analysis that relate to the theories introduced in the literature review.

6.1 Validity of main results

The validity check we carried out in Section 5 brought evidence indicating that some of the covariates are not continuous at the cutoff: it's the case for the binary variables measuring the probability of the respondent being a woman, that of male partners having achieved tertiary education, and female age. We first discuss the dummy for the respondent being a woman and we subsequently relate this to the other variables.

As an alternative way to assess the potential discontinuity of the respondent's sex dummy, we once again exploit the McCrary test and plot the density function of the running variable for both the subsamples of household with female (Figure 4a) and male (Figure 4b) respondents.

Figure 4: Density function of the running variable by sex of the respondent



(a) Subsample of female respondents

(b) Subsample of male respondents

Notes: even though more observations are displayed (up to 48 months on either side of the cutoff), the test only considers those contained in our preferred bandwidth of 13 months.

What immediately stands out is that the confidence intervals of the distribution are not overlapping in the subsample of households with a female respondent, while they are almost perfectly matching in that of male respondents. This could constitute a threat to our identification strategy: the estimates might be biased as household with female respondents are selecting out at the cutoff. As we mentioned in Section 4, we believe that this observed phenomenon is strictly related to the event of birth itself: women are likely to be less available postpartum as they are the member of the household that takes most responsibility on childcare related tasks.[15] This easily result in the impossibility - due to lack of time - or unwillingness to spend time on the survey, an activity that would cause greater disutility to a childbearing woman. Given that the subsample of households with a male respondent shows no jump in the density distribution of the running variable, we believe that estimates restricted to this subset of observations should be more reliable and more likely to be unbiased, conditional on controlling for the other covariates. We run the analysis on both subsamples and assess how it relates to the main findings we presented in Section 5.4.1. Table 5 shows the coefficients of interest on the same four task-count outcome variables for households with female respondents (Panel A) and those with male respondents (Panel B): some interesting patterns emerge.

[15]See to Table 8.

Table 5: RD Estimates. Frequency of tasks, by sex of the respondent

	(1) Woman	(2) Man	(3) Equally	(4) Someone else
Panel A:				
Female Respondent	-0.112	-0.443**	0.485	0.070
	(0.333)	(0.206)	(0.350)	(0.297)
Observations	316	316	316	316
Panel B:				
Male Respondent	-0.863**	-0.237	1.030**	0.070
	(0.428)	(0.308)	(0.430)	(0.286)
Observations	231	231	231	231

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. These outcome variables tell how many tasks a certain person within the household (man, woman, man and woman equally, someone else) always or usually takes on the responsibility for. The regression makes use of a triangular kernel weighting system. The table show coefficients for the variable of interest, "Birth", which should capture the effect of giving birth on the amount of tasks the specific household member is responsible for. We include controls for the male and female household members' age, whether they have obtained tertiary education or not, as well as control for the sex of the respondent. Controls for whether the respondent reside in a western European country, has traditional gender role views, conditional on the sex of the respondent, are also added.

In Panel A, we see that the effects for this subsample are not in line with those observed in Section 5.4.1, in fact, the birth coefficient is not significant anymore for the "man and woman equally" outcome variable, there is no significant reduction in the share of tasks done always or usually by women, while there seems to be a bigger and more significant drop in the share of tasks for which the man is the main responsible (-0.433, significant at the 5% level as opposed to -0.306, only marginally significant). This comparison is compromised by the questionable validity of the results from the panel, yet by observing the differences with the main model we can formulate some considerations. As we have already mentioned, for this particular subsample with female respondents, $Distance_i$ has a discontinuous density function: households might be selecting out at the cutoff. As a result, individuals on either side of the cutoff would not be comparable with those on the other side as they might show systematic differences in their baseline characteristics. When regressing the other covariates on birth, for this particular subsample, we observe a positive and significant effect on the probability of men achieving tertiary education: this somehow validates the reasoning that has just been made.[16] It seems that the drop

[16]Tables for a discontinuity check of the covariates in the subsamples of male and female respondents can be found in A.

in the number of households with female respondents at the cutoff is associated with an increase in the probability of men achieving tertiary education. We speculate that the households that select out of the survey are those where neither the mother nor the father are willing/can answer the survey and, at the same time, where men have, on average, an education equal or lower to that of their counterparts on the other side of the cutoff. Assuming that this was true, we would be left with a higher concentration of households of more educated men on the right side of the cutoff. Analysing this from the exchange bargaining perspective, men in the households to the right of the cutoff have greater bargaining power (since higher education can be considered as a proxy for higher potential income) and, as a consequence, this could determine a more gendered division of household labor. In contrast, one could think that more educated men might have more egalitarian views[17] when it comes to social norms or gender role attitudes and so, households to the right of the cutoff would be more prone to implement an equal split of the chores. The regression results on the subsample of female respondents seem to be more in line with the behavior described by the exchange bargaining perspective and the estimates from the full sample might constitute a lower bound to the true effect of the transition into parenthood on the probability of sharing household tasks more equally. If we consider the sub sample of males, instead, the increase in the share of tasks done equally by man and woman is still significant (yet at the 5% level now) , and if anything, stronger in magnitude (from 0.746 to 1.030). The drop in significance is most likely determined by the reduced precision of the estimates as only 231 observations are considered in this case as opposed to the 547 of the full sample. This subsample does not comply perfectly with the assumptions of the RD methodology either: the regression results from Table (13) show a significant effect (at the 1% level) of birth on female age. Given the loss of precision induced by the sample restriction, there is reason to believe that a significant result her is proof of a true imbalance around the cutoff. It seems then that for the subsample of household with male respondents, birth is associated with women being older. Nevertheless, we think this is not surprising. The effect of birth on female age is of 1.579, meaning that a woman that has already given birth is on average one and a half year older than a woman who is about to give birth to her first child. We argue that this is unlikely to have a huge impact on the estimates of our sample as the difference in age is not particularly high and, regardless, it is reasonable to think that the probability of experiencing childbirth is positively correlated with age. Valid results for this subsample are then interpretable on the basis of the theoretical frameworks introduced in Section 2: when in a close window around birth, the transition into parenthood generally induces an increase in the probability of men and women sharing more equally the routine chores in

[17] S. N. Davis & Greenstein, 2009; Cassidy & Warren, 1996; N. J. Davis & Robinson, 1991; Rhodebeck, 1996.

the household as women decrease the share of tasks for which they take full responsibility. Against the intuition of the gender display perspective, the man helps carrying out tasks more frequently, even if this coincide with adopting behaviors that would normally display femininity. On the other hand, the asymmetrical nature of the dependency described by the exchange bargaining perspective would imply smaller loss for the man (the provider) and bigger gains for the woman (the dependent partner) when a more equal division of tasks is implemented.[18] When considering the full sample, it needs to be mentioned that the covariate representing gender role attitudes is consistently continuous at the cutoff, no matter the different specification of the model. We interpret this as proof for individual preferences being aligned around the cutoff and for the fact that individuals end up to either side of the threshold as good as randomly. If we believe that, conditional on controlling for the covariates, the true effect of birth on the task-count variable is correctly estimated for the male subsample, then we can think that these results somehow confirm those found when using the full sample.

6.1.1 Reverting effects

As previously said, our findings, presented in Table 3 in Section 5.4.1, show significant coefficients for the amount of tasks the partners share equally between themselves. This is somehow in contrast with the findings of the existing literature, as studies usually find that female partner increase their time devoted to household labor paired with null or negative effects for the male partner (Sanchez & Thomson, 1997; Campolo & Rizzi, 2016; Kühhirt, 2012; Baxter et al., 2008; Pollmann-Schult, 2017; Yavorsky et al., 2015). When considering the graphs, though, it appears that these effects are not destined to persist: even with a small bandwidth of 13 months, the slope of the fitted line to the right of the cutoff in both Figure 1a and 1c oppose to the sign of the effects we find, suggesting that these effects might only be temporary. More precisely, the sign of the coefficient on the task count for women is negative, which is not in line with what we would expect from simply assessing the descriptive statistics in Table 1, where mothers tend to perform, on average, more household tasks than childless women. The explanation can be found in the positive slope represented by $(Birth * Distance)$, that causes the mean differences to increase sharply as time goes by. This is evident both in the table as well as for the post-birth trend in Figure 1a. In Figure 1c this jump, seem also to revert back, as the post-birth trend is negative and as well as bigger in magnitude than the pre-birth trend, causing the average task count to revert back to its pre-birth mean. Moreover,

[18] Brines, 1994 argues that, given macro-level socio-institutional settings in favor of a gendered division of labor, it will always be easier for the man to implement changes in his share of labor while the woman is likely to struggle more.

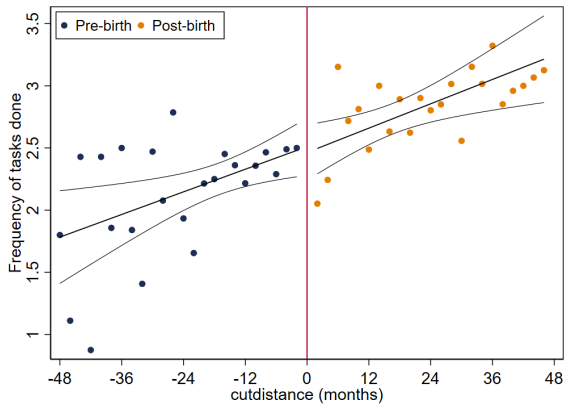
the sensitivity test to the choice of bandwidth reveals again that these significant effects are only found in the vicinity of the cutoff: when increasing the bandwidth to 24 lose statistical significance. A similar outcome is reached when considering the alternative specification with a rectangular kernel: when equal weights are given to all the observation, irrespective of their proximity to the cutoff, results become insignificant. It then appears that childbirth only produces, if anything, a relevant shifts in the division of standard housework only in a limited post-birth time window.

6.2 Expansion

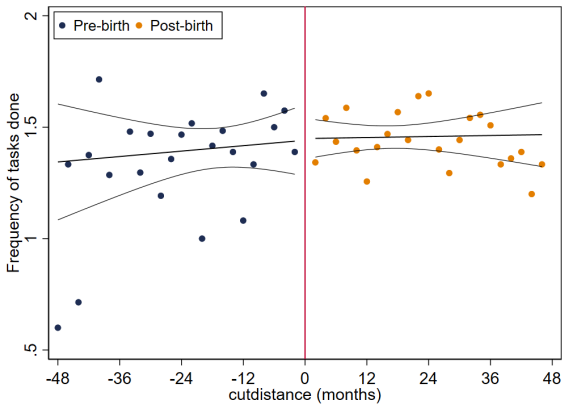
6.2.1 Long term effects

As an explanation for the sharp short term penalties in the Scandinavian countries, the Kleven, Landais, Posch, et al., 2019, point to differences in parental leave duration, as these countries happen to have longer and more generous leave periods. Parental leave might not only have consequences for labor market outcomes as Schober, 2013 show that longer maternal employment interruptions after childbirth, lead to a more traditional labor division within the household, the main channel being mothers contributing with more housework mirrored by fathers cutting their share. By design, the parental leave period starts at the same time as childbirth. This raise the concern of how to disentangle their possible joint effects on household labor division. With this consideration in mind and given that results in Section 5.3 and 5.4.1 seem to be very local and potentially temporary, we are motivated to expand the the bandwidth to assess more long term effects. This is especially appealing since a high fraction of women are on maternity leave in the period immediately after birth and it would be interesting to assess effects of household labor division when this period is over. In Table 5 we propose new graphs and estimates produced by applying the same model on a 48-month bandwidth: the minor shifts previously observed in Section 5.2 are no longer as clear, and seem to be less relevant compared to the steep trends in Figure 5a and 5c.

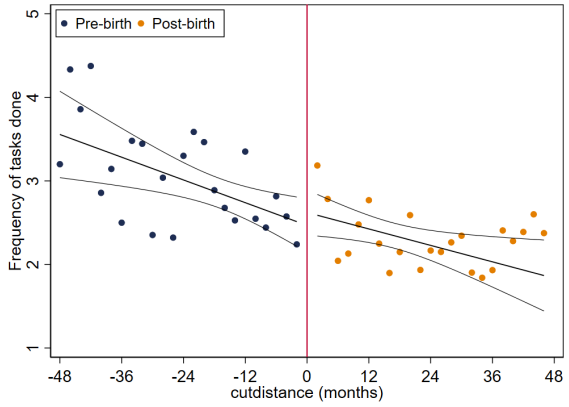
Figure 5: RD estimates. Frequency of household tasks



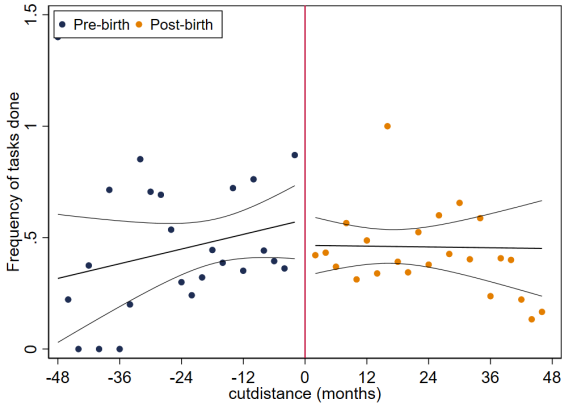
(a) Woman always or usually



(b) Man always or usually



(c) Partners equally



(d) Someone else always or usually

Table 6 show, consistently with the above figures, no significant effects for the outcomes at the threshold. Strikingly, the effect for women seem to converge towards zero as the bandwidth is increased, and a similar pattern applies to the effects on the number of tasks the couples share equally. The local effect from the main analysis (Table 3) die out.

Table 6: RD Estimates. Amount of household tasks

	(1) Woman	(2) Man	(3) Equally	(4) Someone else
Birth	-0.001	-0.072	0.050	0.023
	(0.131)	(0.087)	(0.138)	(0.104)
Distance	0.014**	0.006	-0.024***	0.005
	(0.006)	(0.004)	(0.006)	(0.005)
(Birth x Dist)	0.002	-0.004	0.009	-0.007
	(0.007)	(0.005)	(0.008)	(0.006)
Observations	1806	1806	1806	1806

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. As described in Section 4 we examine the effect of birth on a set of variables that describe how household tasks are split among the household members. In particular, for every household, these outcome variables tell for how many tasks a certain person (man, woman, man and woman equally, someone else) is always or usually the main responsible. Coefficients for all the covariates are included. The regression makes use of a triangular kernel. We include controls for the male and female household members age, whether they have obtained tertiary education or not and whether the sex of the respondent. Controls for whether the respondent reside in a western European country. A binary variable describing whether the respondent agree with a traditional gender role views and a interaction this variable and the sex of the respondent is also added, but not included in the table. Birth is the coefficient of interest, that should capture the effect of giving birth on the amount of tasks the specific household member is responsible for. The control variables $Distance_i$ and $(Distance_i * Birth)$ are included in the regression consistent with the model presented in Section 4.1.

Both the table and visual evidence show a loss of significance in the effects, as the bandwidth increases. The temporary nature of the findings can for example be explained by increased sharing and male participation in tasks, due to mothers health a few months after the birth. We consider the effects in the long term to be of more interest in terms of household labor division and employment situation in the years following childbirth. We will therefore in the following sections consistently examine the full sample of observations scattered over 48 months to either side of the threshold.

6.2.2 Household task decomposition

We expand on the previous findings by decomposing the outcomes over the 7 standard household tasks. Table 16 in Appendix A.4 show averages for the outcome variables decomposed into the seven different tasks measured by the survey. The different labor division outcomes are presented in rows, and the columns show averages for the seven household tasks. The sample is split into Panel A and B where the former is the parent sample and the latter the non-parent samples, which corresponds to the treatment and control groups mentioned in 4.1. First, we see a few patterns emerge across the whole

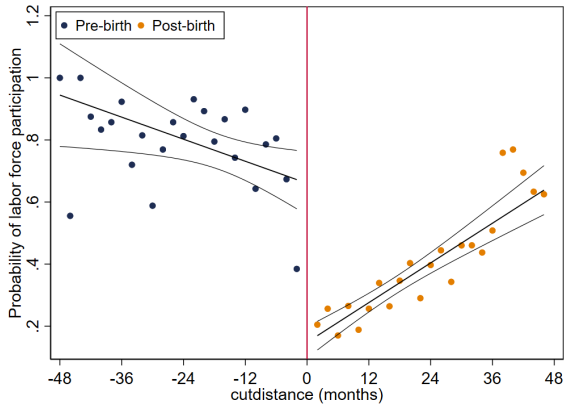
sample. As seen in column 1, 2 and 4, about half of the households report women always or usually preparing meals and doing cleaning related tasks. For the same tasks, we observe very low frequencies of males executing them. On the other hand, more than 75% of the households report that men always or usually do small repairs, in or around the house. Outcomes for this tasks are matched with a low share of women taking care of such repairs. Secondly, there seem to be some differences between samples as well. The parent sample reports women doing nearly all task more frequently relative to the non-parent sample. In addition, splitting the tasks equally or having someone else doing them, is less frequent for the parent sample. When it comes to men, they are more likely to take care of the financial records and small repairs, but typically are less likely to be doing the dishes.

As discussed in the previous section, Figure 5 indicates no sharp shifts in the total amount of tasks within the household. There can however be relative shifts hidden in the aggregate task count, regarding the type of task the members usually execute. Meaning that the household members could potentially change the way that they distribute tasks between themselves, but still remain at the same aggregate amount. This seem however not to be the case, as shown by the continuous RD graphs presented in Appendix A.4 (Figure 12 to 18). Table 18 in Appendix A.4 points in the same direction as it only suggest a marginally significant effect for men decreasing their participation in washing dishes. Both the tables and the graphical results show that there are no discontinuous effects after birth even when looking into the decomposed tasks.

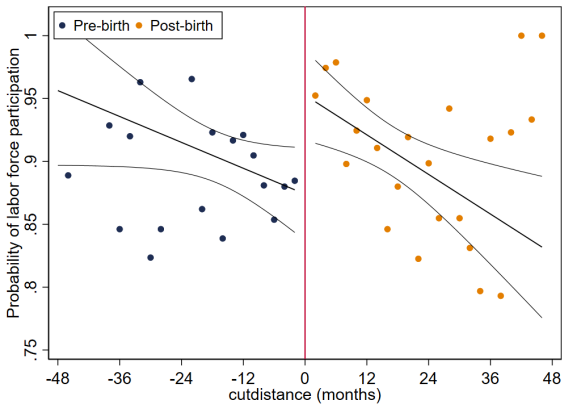
6.2.3 Labor market outcomes

Kleven, Landais, & Søgaard, 2019 discuss the three potential factors determining the earnings penalty from children: labor force participation, hours of work and the wage rate. We will here present RD graphs for hours worked and the probability of labor force participation, as we lack income data from most countries. As seen in Figure 6c and Figure 6a there are sharp and discontinuous shifts for women at the threshold, but not for men. Women work fewer hours and are less likely to participate in the labor force after the birth of a child, relative to the time before parenthood. The return of women to the labor force as time passes in the post-birth period can be seen by the positive sharp positive slope in the post-birth period, but even after a four year period the mean outcome has still not converged back to its pre-birth level.

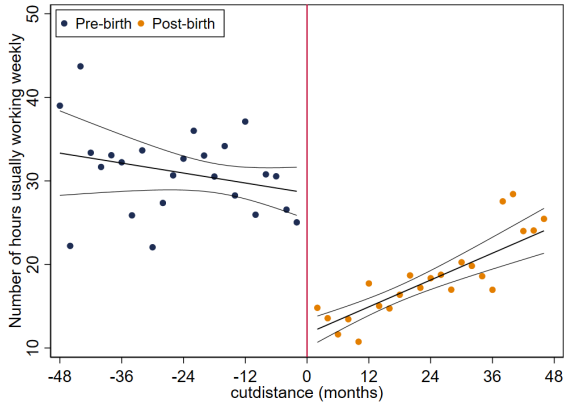
Figure 6: RD estimates. Labor market outcomes



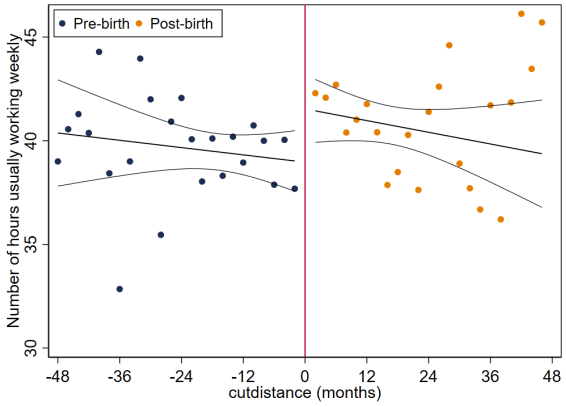
(a) Labor force participation: Women



(b) Labor force participation: Men



(c) Weekly working hours: Women



(d) Weekly working hours: Men

The results emerging from Panel A in Table 7 are consistent with Figure 6c, showing a 16,4 hour decrease in the weekly hours women spend in paid work, as opposed to men experiencing an increase of 2,8 hours on average. The coefficients for female hours are highly significant at the 1% level, while the effects for men are only marginally significant. Panel B show that women are 51% less likely to participate in the labor force after birth. This drop needs to be seen in the light of the positive trend in the post birth period and a maternity leave status being treated as non-participation in the labor force.

Table 7: Labor market outcomes

	(1) Women	(2) Men
Panel A:		
Birth	-16.463***	2.781*
	(1.586)	(1.430)
Distance	-0.119*	-0.038
	(0.071)	(0.064)
(Birth x Dist)	0.358***	-0.026
	(0.088)	(0.079)
Observations	1824	1816
Panel B:		
Birth	-0.510***	0.064**
	(0.036)	(0.025)
Distance	-0.008***	-0.002
	(0.002)	(0.001)
(Birth x Dist)	0.018***	-0.001
	(0.002)	(0.001)
Observations	1852	1856

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Amount of weekly hours in paid work, here to time normally worked per week in a job or business. Respondents with missing or non applicable survey answers, which had a activity status other then some type of employment were assumed to be working zero hours. Maternity leave status is being treated as non-participation in the labor force, and is given a zero value.

Consistent with the findings in the relevant literature, mothers in our sample reduce the time they allocate to paid employment, while the fathers face no significant changes in their work supply (Kleven, Landais, & Søgaard, 2019; Kleven, Landais, Posch, et al., 2019; Sanchez & Thomson, 1997; Campolo & Rizzi, 2016; Kühhirt, 2012; Pollmann-Schult, 2017). Additionally women are less likely to participate in the labor force when they become mothers. Even though we lack data on wages and earnings, we have identified effects for two of the three main drivers of the child penalty in earnings (Kleven, Landais, Posch, et al., 2019; Kleven, Landais, & Søgaard, 2019).

6.2.4 Childcare

Table 8 show averages for the share of household members executing various household tasks related to childcare. In column 1 to 3 we can see that over 60 percent of the women in the sample take up always or usually the responsibility of dressing the children, putting them to bed and taking care of them in case of illness. For all of the four tasks displayed in the table, men have very low probabilities, around 1 to 5% of fathers always or usually performing the relevant task. Even though the introduction to parenthood generally increase the amount of time devoted to childcare for both genders, our findings are consistent with the general literature showing that the mother in particular, face a significantly larger increase in time devoted to the newly imposed childcare tasks (Gjerdingen & Center, 2005; Yavorsky et al., 2015; Kühhirt, 2012).

Table 8: Mean outcomes. Who does household tasks with children

	(1) Dressing	(2) Putting to bed	(3) Illness	(4) Leisure activities
Woman	0.72	0.61	0.79	0.29
	(0.01)	(0.01)	(0.01)	(0.01)
Man	0.01	0.05	0.03	0.03
	(0.00)	(0.01)	(0.01)	(0.01)
Equally	0.25	0.33	0.15	0.65
	(0.01)	(0.01)	(0.01)	(0.01)
Someone else	0.01	0.01	0.03	0.03
	(0.00)	(0.00)	(0.00)	(0.00)
Kids	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	1209	1210	1193	1205

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Who does household tasks with children, always or usually?

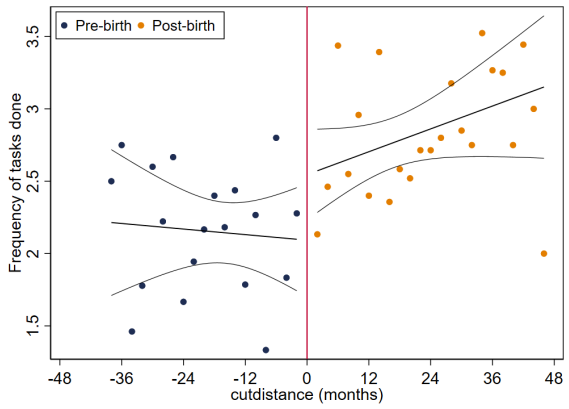
6.3 Heterogeneity analysis

In Section 2 we introduce three theories regarding determinants of household labor division. We start this subsection by assessing whether we can find evidence for the exchange bargaining and household specialization theories proposed in Section 2. We

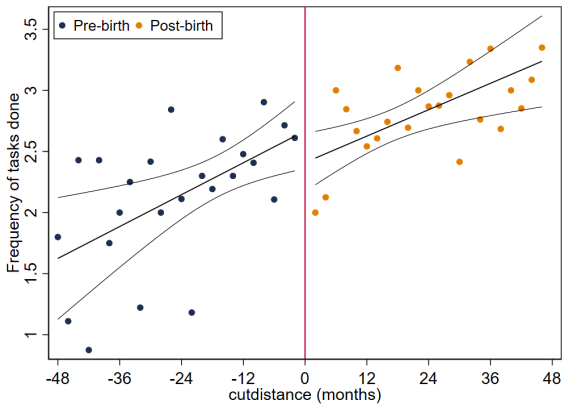
investigate whether the relative efficiency, income and resource levels between the partners can have impacts on the effects. In the absence of income data, we consider the relative education level between partners as a proxy for earnings potential and resources. We want to compare three groups of households: those where both partner have tertiary education, those where only the woman has it and lastly, those where only the man has it. We choose these three groupings since this is the way that give us most variation across groups and to each of them assign the highest number of observation given the distribution of tertiary education in the sample. We do not observe any significant effect no matter the group that we consider. Especially for the groups characterized by an education gap between partners we see that the standard errors are quite high: this might explain why we do not observe relevant effects.[19] Given the insignificance of our findings we cannot take a stand on the exchange bargaining theory. To further proceed in the heterogeneity check we move on to assessing whether there is any difference in how labor division unfolds according to gender norms. As discussed in Section 4.1, we split the countries in two groups, determined by the share of each county's population agreeing with a question regarding working mothers and the well being of their children. The resulting sorting corresponded quite conveniently with the geographical region of Eastern and Western Europe. In Figure 7a and 7b we consider the western and eastern European countries separately. Post-birth trends seem to be quite similar, but a difference comes from the pre-birth outcomes and the shifts at the threshold. The western European countries seem to have a rather flat trend in the years before birth, but do however see a shift and a more steep upward trend after the event. This shift is not significant in the figure but does however have a significant coefficient in Table 9. This is opposed to the eastern European countries which has an upward sloping trend and seemingly continuous outcomes across the threshold while results are not significant at a conventional level for neither the figure nor the table.

[19] Table 17 in Appendix A.4 show the regression results for this heterogeneity check.

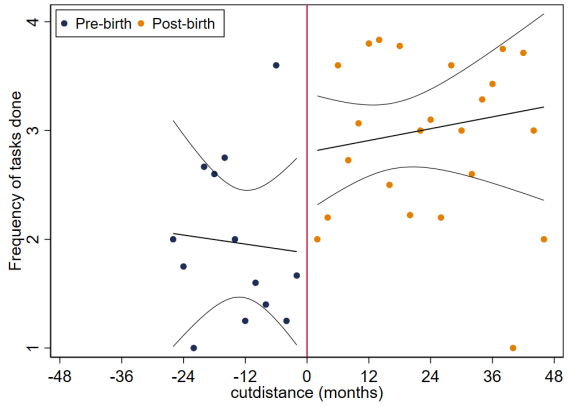
Figure 7: RD estimates. Cross-country effects



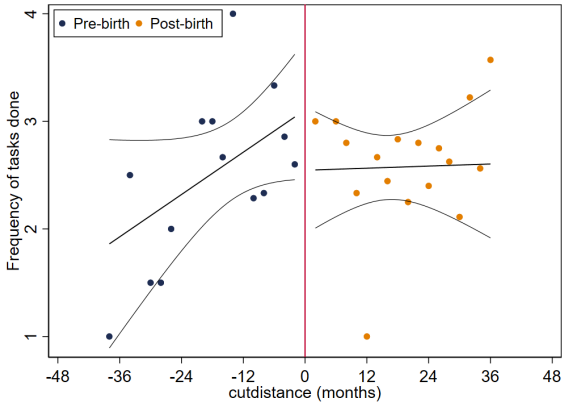
(a) Western European countries



(b) Eastern European countries



(c) Western European countries



(d) Eastern European countries

This finding is rather puzzling and we look at countries separately to identify the drivers of these observed behavior in the data. Among the western countries, Austria seems to be the main contributor, showing stronger shifts in both regression estimates and the corresponding plot. The coefficient for Austria in Table 9 is highly significant and suggest that women, on average perform one task more after childbirth. By looking at the eastern countries, Georgia seem to be a contributor to the aggregate effects with a stronger effect and significant regression coefficient. This coefficient suggest that Georgian mothers perform on average one less task than their childless counterparts.

Table 9: RD Estimates. Cross-country effects

	(1) Woman	(2) Man	(3) Equally	(4) Someone else
Western Europe				
Birth	0.522**	-0.280*	-0.221	-0.020
	(0.229)	(0.150)	(0.233)	(0.090)
Observations	620	620	620	620
Eastern Europe				
Birth	-0.291*	0.039	0.130	0.122
	(0.162)	(0.107)	(0.172)	(0.149)
Observations	1186	1186	1186	1186
Austria				
Birth	1.059***	-0.369	-0.701*	0.012
	(0.402)	(0.226)	(0.400)	(0.178)
Observations	237	237	237	237
Georgia				
Birth	-0.960**	-0.295	-0.110	1.365**
	(0.419)	(0.384)	(0.309)	(0.590)
Observations	163	163	163	163

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

It important to note that the sample sizes for each country vary markedly and finding significant effects for the above mentioned countries does not imply that the rest lack thereof.[20] It might as well be the case that we fail to find country specific effects in the other countries, due to lack of statistical power. Furthermore, the sign of the coefficients by splitting by western and eastern countries based on the observation that western countries have less traditional views on working mother. Based on this intuition we would initially expect these countries to have less gendered shifts in household labor division, meaning that we would expect countries with more modern gender role views to have weaker effects. This is however the opposite of what we observe, as the evidence suggest that parenthood increase the gendered division of labor by having the mother take on more tasks.

[20]We do find effects for Lithuania. These results are however not presented or discussed as we only have observations from 63 Lithuanian households.

All in all, the heterogeneity check that we run only help us partially in our effort to assess how well the theories introduced in Section 2 fit the data we had at hand. The splitting of the sample according to household education gap does not bring significant results, suggesting that the process of household labor division is insensitive to the household education gap (when considering couple where both partners have tertiary education and couples where only one of the two has it). As we assume education to be a proxy for potential income, we conclude that the dynamics described by the exchange bargaining power are not in place in the subsample that we consider. Conversely, the gender norm heterogeneity check brings up puzzling results as we observe that households belonging to a country with more egalitarian view, face an increase in the share of tasks for which women are mainly responsible, as opposed to household in more traditional country, where women reduce the share of tasks for which they are mainly responsible, as a consequence of birth. This is clearly in contradiction with what theories based on the gender ideology would predict: we would expect household living in more modern countries to have a more egalitarian view of gender roles and, as a result, a more equal share of the household tasks. It might be the case that the average behavior of a country's population is not necessarily aligned with those of the household that we consider in this subsample: considering a direct measurement of subjective gender role preferences would allow for a more precise identification of the relationship between gender norms and household labor division.

7 Conclusion

This thesis focuses on the effect of childbirth on household labor division. Our main objective is to answer the following research question: *does household labor specialization evolve as a consequence of childbirth? And if so, how?* By answering this question, we produce two contributions to the existing literature: first, we test whether the previous findings are matched with the application of this alternative methodology, which has not been applied yet in the literature we reviewed on this field of research. Secondly, we investigate whether relative education levels or exposure to different societal gender norms can produce heterogeneous results, in line with either what the theory of gender display or resource bargaining theory would predict. Cross-country data is used to compare couples that have just recently experienced the transition into parenthood with couples that are about to do so. We implement an RD-like methodology, assuming that individuals close to the cutoff have, on average, similar characteristics.

We find that in the very short run, couples share household labor more equally after the birth of their first child, as opposed to couples that are about to give birth. A more equal

sharing of grocery shopping, vacuum cleaning, and record keeping activities are identified as the main drivers for the aggregate short term effects. When testing the robustness of these results to different specifications of the model, we find that they are quite sensitive to the bandwidth choice and that they gradually lose significance the more we expand the bandwidth. When considering the largest bandwidth (our full sample), we can assess the more long-term implications of birth: the effects are not significant anymore and it appears that there is no difference in terms of household labor division of the seven tasks between the pre- and post-birth couples. Our descriptive statistics show a positive association between parenthood and a gendered labor division. This is consistent with the increasing long-run trend we observe in our graphical results, where the couples transition from sharing more equally, towards the woman taking on more tasks herself. Aligned with previous studies, our descriptive statistics show that mothers take on the main share of child-related responsibilities. When we, in turn, assess the effects of parenthood on labor market outcomes, we see that mothers significantly decrease their time spent in paid work with 16 hours weekly after birth, and are also 50% less likely to participate in the labor force compared to childless women. Men, on the other hand, are more likely to be participating in the labor force when they become fathers.

As an expansion, we run a heterogeneity analysis to address the theories related to our topic. We start by first looking at household education gaps as a proxy for partners' relative income to test the resource exchange bargaining theory. We do not find heterogeneous effects of birth on households with different education gaps and are thus unable to confirm that relative income and partners' resources moderate the effects of birth on household labor division. We then split the household according to geographical regions to test the theories based on gender ideology. We find that Western European countries, associated with a more modern view on working mothers, tend to have gendered effects, playing out in terms of mothers performing on average half a task more after childbirth. The effects are particularly strong for Austrian women, as they generally perform an additional task after the start of parenthood. While significant, these effects are somehow inconsistent with the theories we presented in the literature review.

When it comes to further expansions on this topic and our analysis, in particular, the most relevant note is that a third round of the Generations and Gender Survey is scheduled to launch in 2020. This will most likely bring new valuable data that could significantly improve the internal validity of the analysis. More generally, further research should attempt to identify and test new mechanisms capable of producing consensus on the dynamics of the childbirth effect on household labor division.

References

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: Women and the plough *. *The Quarterly Journal of Economics*, *128*(2), 469–530. doi:10.1093/qje/qjt005

Angrist, J. D. (2009). *Mostly harmless econometrics : An empiricist's companion*. Princeton: Princeton University Press.

Angrist, J. D. & Pischke, J. S. (2015). *Mastering 'metrics : The path from cause to effect*. Mastering metrics. Princeton, N.J: Princeton University Press.

Balbo, N. & Arpino, B. (2016). The role of family orientations in shaping the effect of fertility on subjective well-being: A propensity score matching approach. *Demography*, *53*(4), 955–978. doi:10.1007/s13524-016-0480-z

Baxter, J., Hewitt, B., & Haynes, M. (2008). Life course transitions and housework: Marriage, parenthood, and time on housework. *Journal of Marriage and Family*, *70*(2), 259–272. doi:10.1111/j.1741-3737.2008.00479.x

Becker, G. S. (1981). A treatise on the family. *Cambridge, MA: Harvard University Press*.

Berk, S. F. (1985). The gender factory: The apportionment of work in american households. *New York: Plenum*.

Bertrand, M., Kamenica, E., & Pan, J. (2015). Gender identity and relative income within households *. *The Quarterly Journal of Economics*, *130*(2), 571–614. doi:10.1093/qje/qjv001

Blau, F. D. & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations †. *Journal of Economic Literature*, *55*(3), 789–865. doi:10.1257/jel.20160995

Brines, J. (1994). Economic dependency, gender, and the division of labor at home. *American Journal of sociology*, *100*(3), 652–688.

Calonico, S., Cattaneo, M., Farrell, M., & Titiunik, R. (2019). Regression discontinuity designs using covariates. *The Review of Economics and Statistics*, *101*(3), 442. doi:10.1162/rest_a_00760

Campolo, M. & Rizzi, E. (2016). Parenthood and labour division in italian two-earner couples. *International Journal of Social Economics*, *43*(12), 1315–1333. doi:10.1108/IJSE-12-2014-0248

Cassidy, M. & Warren, B. (1996). Family employment status and gender role attitudes: A comparison of women and men college graduates. *Gender Society*, *10*(3), 312.

Davis, N. J. & Robinson, R. V. (1991). Men's and women's consciousness of gender inequality: Austria, west germany, great britain, and the united states. *American Sociological Review*, *56*(1), 72. doi:10.2307/2095674

Davis, S. N. & Greenstein, T. N. (2009). Gender ideology: Components, predictors, and consequences. *Annual review of Sociology*, *35*, 87–105.

Dribe, M. & Stanfors, M. (2009). Does parenthood strengthen a traditional household division of labor? evidence from sweden. *Journal of Marriage and Family*, *71*(1), 33–45. doi:10.1111/j.1741-3737.2008.00578.x

Eurofound. (2007). *Fourth european working conditions survey*. Office for Official Publications of the European Communities, Luxembourg.

Eurofound. (2012). *Fifth european working conditions survey*. Publications Office of the European Union, Luxembourg.

Eurofound. (2017). *Sixth european working conditions survey – overview report (2017 update)*. Publications Office of the European Union, Luxembourg.

Generations & Programme, G. (n.d.). The infrastructure is run by leading social scientists and national statistical offices. Web Page. Retrieved from https://www.ggp-i.org/about/

Generations & Programme, G. (2019). Generations and gender survey ggs (wave 1) [data file and code book version 4.3]. Dataset. doi:doi.org/10.17026/dans-z5z-xn8g

Gjerdingen, D. K. & Center, B. A. (2005). First-time parents' postpartum changes in employment, childcare, and housework responsibilities. *Social Science Research*, *34*(1), 103–116. doi:10.1016/j.ssresearch.2003.11.005

Goffman, E. (1976). Studies in the anthropology of visual communications. *Gender Adv*, *3*(2), 69–154.

Grunow, D., Schulz, F., & Blossfeld, H.-P. (2012). What determines change in the division of housework over the course of marriage? *International Sociology*, *27*(3), 289–307. doi:10.1177/0268580911423056

Imbens, G. W. & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635. doi:10.1016/j.jeconom.2007.05.001

Jacob, R., Zhu, P., Somers, M.-A., & Bloom, H. (2012). A practical guide to regression discontinuity. *MDRC*.

Kim, E. H.-W. & Cheung, A. K.-L. (2019). The gendered division of household labor over parenthood transitions: A longitudinal study in south korea. *Population Research and Policy Review*, *38*(4), 459–482.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., & Zweimüller, J. (2019). Child penalties across countries: Evidence and explanations. *AEA Papers and Proceedings*, *109*, 122–126an. doi:10.1257/pandp.20191078

Kleven, H., Landais, C., & Søgaard, J. E. (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics*, *11*(4), 181–209. doi:10.1257/app.20180010

Kühhirt, M. (2012). Childbirth and the long-term division of labour within couples: How do substitution, bargaining power, and norms affect parents' time allocation in west germany? *European Sociological Review*, *28*(5), 565–582. doi:10.1093/esr/jcr026

Lee, D. S. & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355. doi:10.1257/jel.48.2.281

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714. doi:10.1016/j.jeconom.2007.05.005

Pollmann-Schult, M. (2017). Sons, daughters, and the parental division of paid work and housework. *Journal of Family Issues*, *38*(1), 100–123.

Rhodebeck, L. A. (1996). The structure of men's and women's feminist orientations: Feminist identity and feminist opinion. *Gender Society*, *10*(4), 386–403.

Sanchez, L. & Thomson, E. (1997). Becoming mothers and fathers: Parenthood, gender, and the division of labor. *Gender Society*, *11*(6), 747–772. doi:10.1177/089124397011006003

Schober, P. S. (2013). Maternal labor market return and domestic work after childbirth in britain and germany. *Community, Work Family*, *16*(3), 307–326. doi:10.1080/13668803.2013.820096

West, C. & Zimmerman, D. H. (1987). Doing gender. *Gender society*, *1*(2), 125–151.

Yavorsky, J. E., Kamp Dush, C. M., & Schoppe-Sullivan, S. J. (2015). The production of inequality: The gender division of labor across the transition to parenthood. *Journal of Marriage and Family*, *77*(3), 662–679.

A Appendix

A.1 Data

A.1.1 Further explanation on the sample-restriction procedure

Our analysis builds on the comparison between households that already have a child at the time of the Wave 1 survey (observations that constitute the treatment group) and those couples that are together but childless at the time of the first survey and give birth to their first child between waves. This necessarily implies stricter limitations to the second group of observations (control group). Indeed, households from the control group are composed of couples whose relationship is at least as long as the intra-waves distance. In addition, these same households give birth at a certain distance from the Wave 1 survey which is, at most, equal to the few years between the first and second waves. Given that different countries implemented the Wave 1 and Wave 2 surveys with different time gaps, these additional constraint on the control group will vary by country. That is why, in an attempt to equalize the conditions under which we sort observations into treatment and control, we will drop households from the treatment group according to their country specific maximum and minimum intra-waves distances. This will ensure that the same time constraint in terms of the minimum duration of the relationship and the maximum distance from the Wave 1 survey to childbirth are applied to treatment and control group.

A.2 Empirical approach

Table 10: Share of population agreeing to statement

	A pre-school child is likely to suffer if his/her mother works
Bulgaria	0.65
Russia	0.72
Georgia	0.75
Germany	0.46
France	0.51
Hungary	0.86
Romania	0.45
Austria	0.41
Estonia	0.18
Belgium	0.40
Lithuania	0.61
Poland	0.64
CzechRepublic	0.44
Sweden	0.11
Total	0.56

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.3 Results

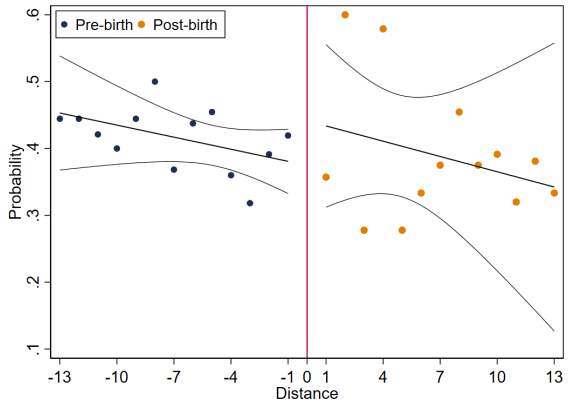
Table 11: Descriptive statistics

	Non-Parent sample		Parent sample	
Panel A: Average task count				
Woman	2.40	(0.10)	2.59	(0.10)
Man	1.41	(0.06)	1.44	(0.07)
Equally	2.62	(0.10)	2.53	(0.10)
Someone else	0.57	(0.08)	0.45	(0.07)
Observations	279		268	
Panel B: Average characteristics				
Female R	0.58	(0.03)	0.56	(0.03)
Age - Woman	27.55	(0.24)	28.96	(0.26)
Age - Man	30.27	(0.31)	31.07	(0.28)
Tertiary education - Woman	0.41	(0.03)	0.40	(0.03)
Tertiary education - Man	0.31	(0.03)	0.37	(0.03)
Western Europe	0.32	(0.03)	0.41	(0.03)
Traditional	0.58	(0.03)	0.55	(0.03)
Observations	288		283	

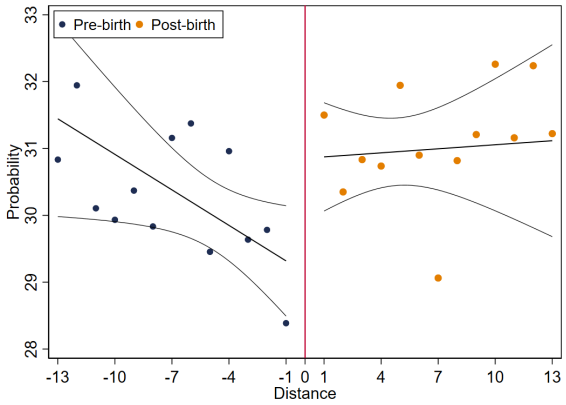
Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Panel A present average outcomes for the amount of tasks a certain person within the household (man, woman, man and woman equally, someone else) is always or usually the main responsible for. The table is split into averages presented for parent and non parent couples separately. Panel B show average characteristics where Female R is the fraction of households where there are a female respondent to the survey. Tertiary education is the fraction that has obtained at least tertiary education. Western Europe show the fraction of households living in a Western European country and traditional is the share of respondents agreeing with a statement regarding working mothers.

A.3.1 Additional graphs for the main covariates

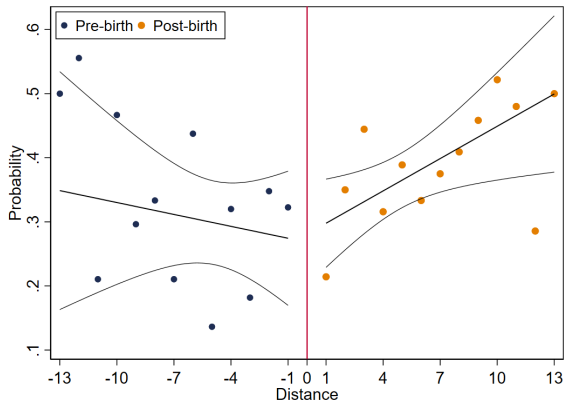
Figure 8: Discontinuity test for the other covariates, main model



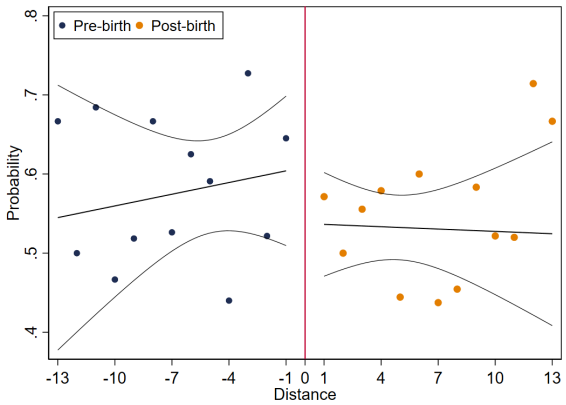
(a) Tertiary education, women



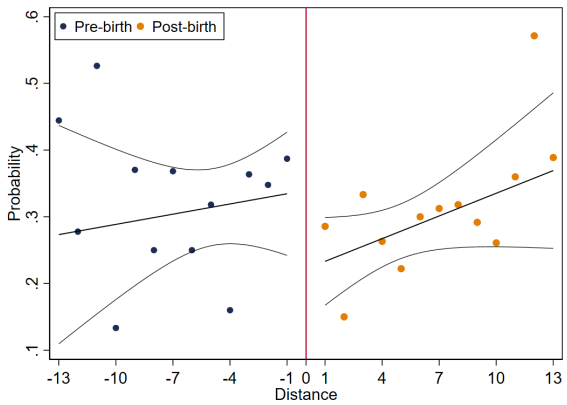
(b) Age, men



(c) Region



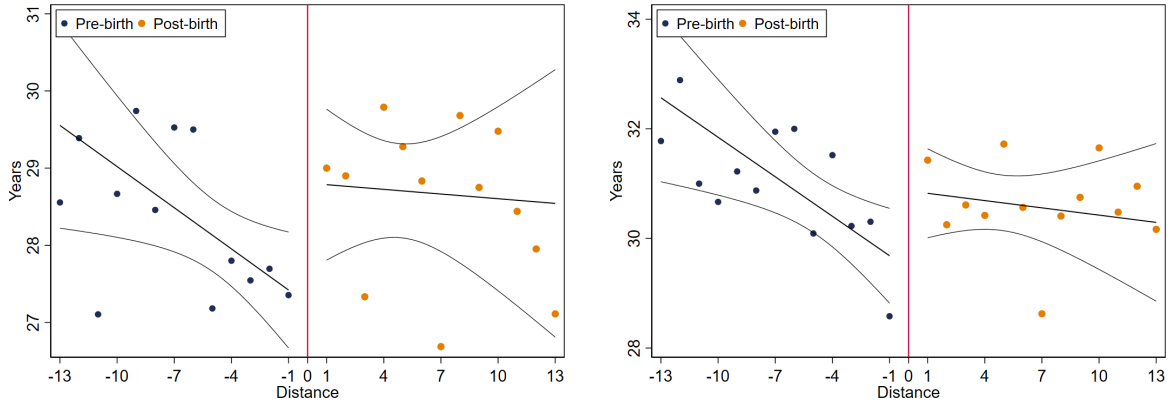
(d) Subjective gender role attitudes



(e) Interaction: sex of the respondent * subjective gender role attitudes

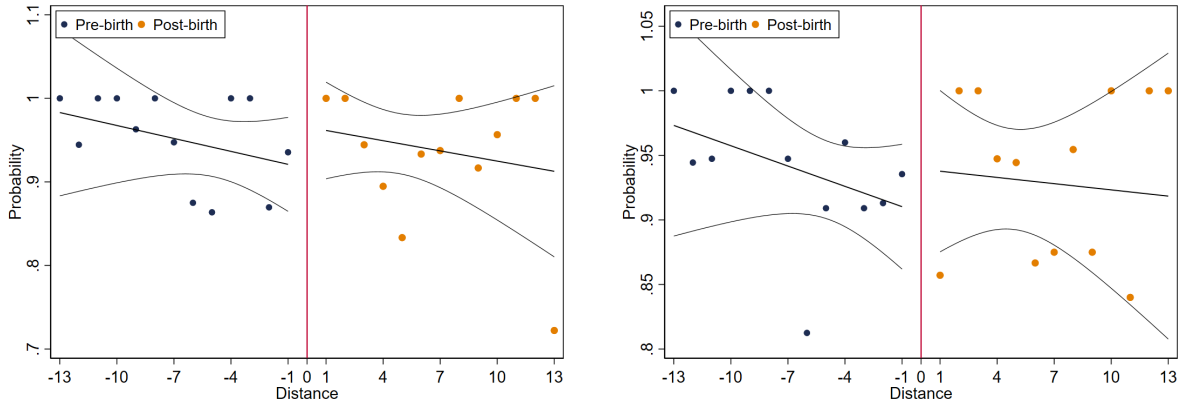
A.3.2 Alternative specifications

Figure 9: Discontinuity test: age at birth



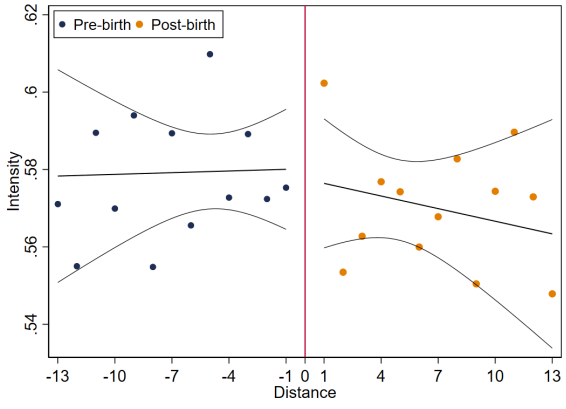
(a) Women (b) Men

Figure 10: Discontinuity test: upper secondary education



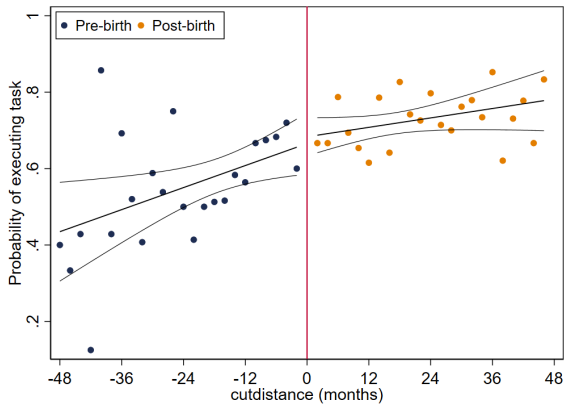
(a) Women (b) Men

Figure 11: Discontinuity test: social norms index

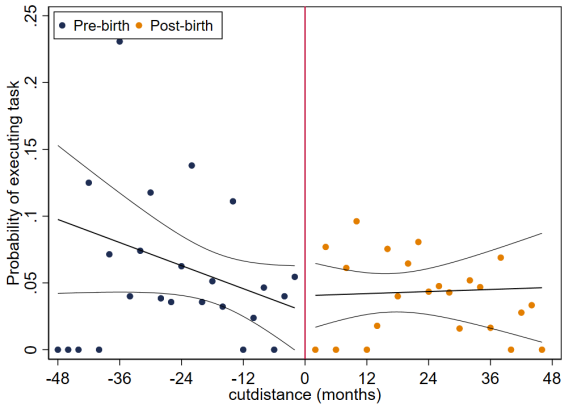


A.4 Discussion

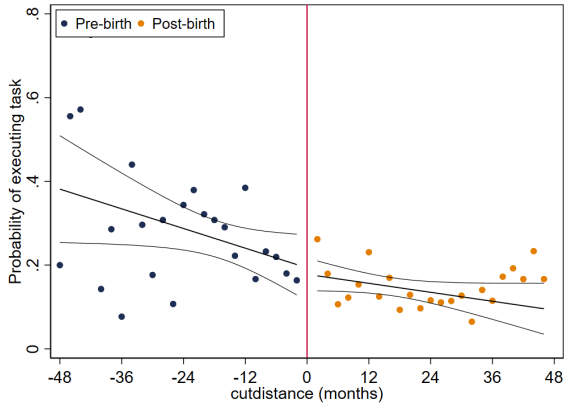
Figure 12: RD estimates. Probability of doing task: Preparing daily meals



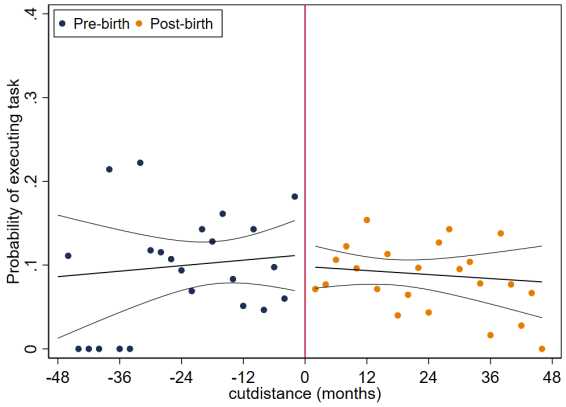
(a) Woman always or usually



(b) Man always or usually

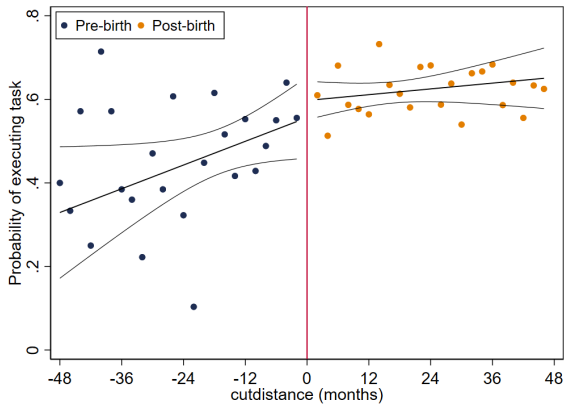


(c) Partners equally

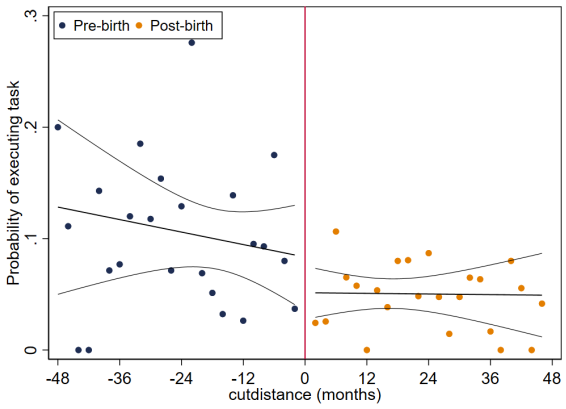


(d) Someone else always or usually

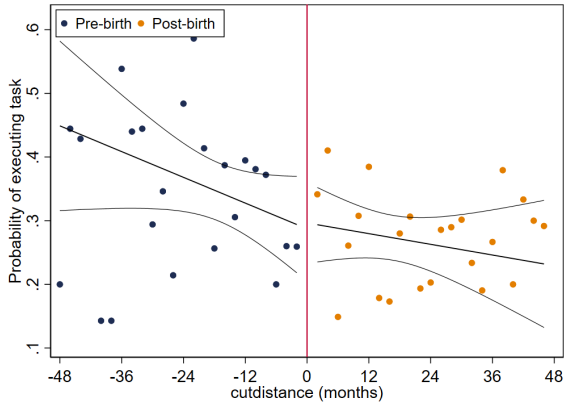
Figure 13: RD estimates. Probability of doing task: Doing the dishes



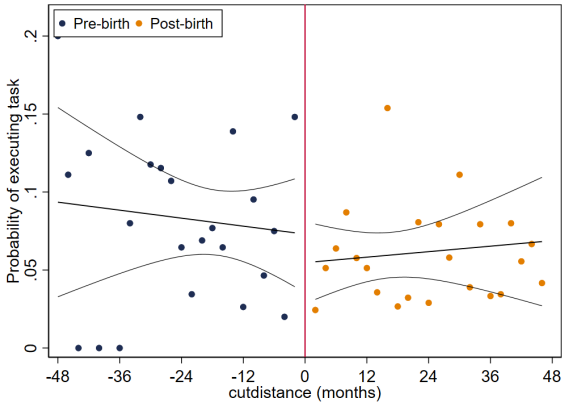
(a) Woman always or usually



(b) Man always or usually

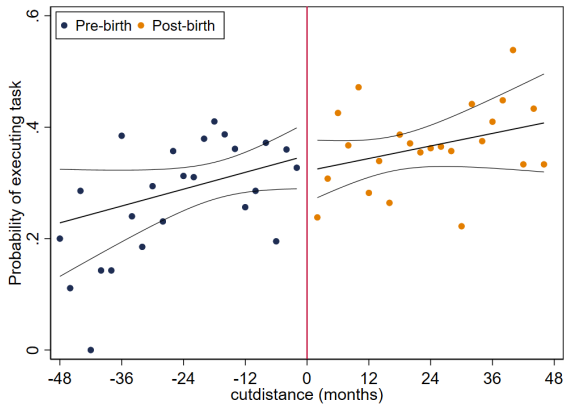


(c) Partners equally

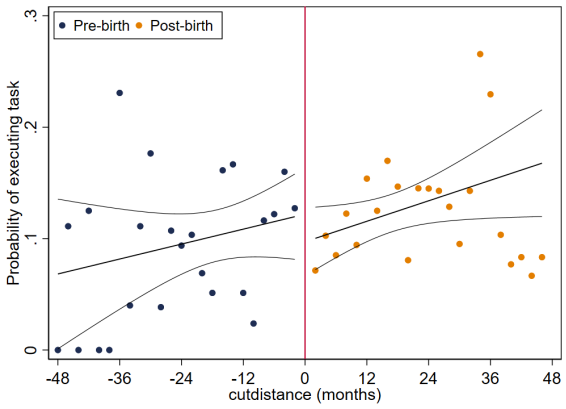


(d) Someone else always or usually

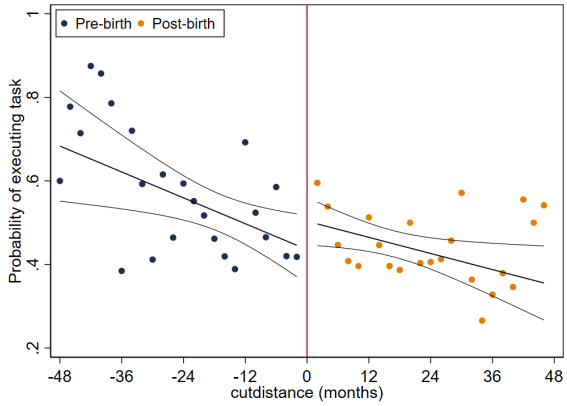
Figure 14: RD Estimates. Probability of doing task: Shopping for food



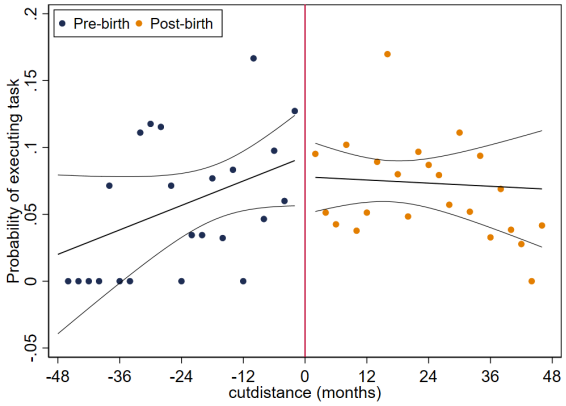
(a) Woman always or usually



(b) Man always or usually

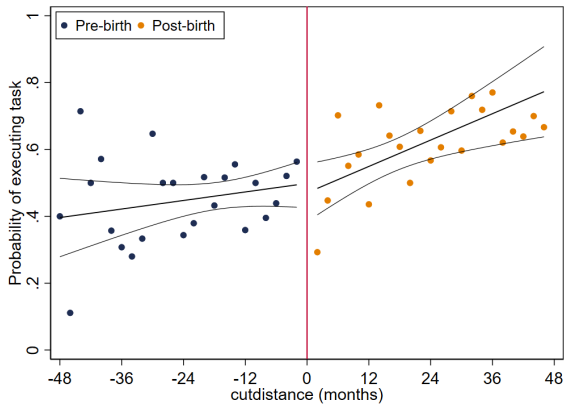


(c) Partners equally

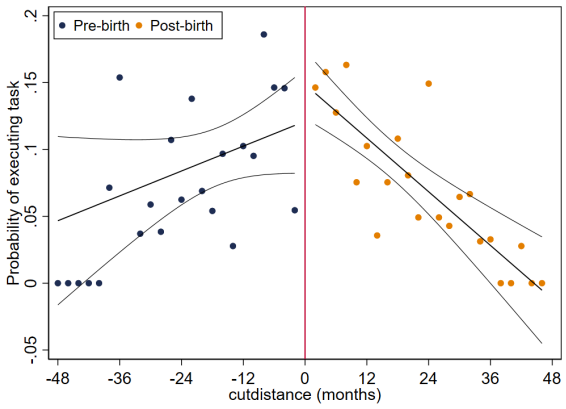


(d) Someone else always or usually

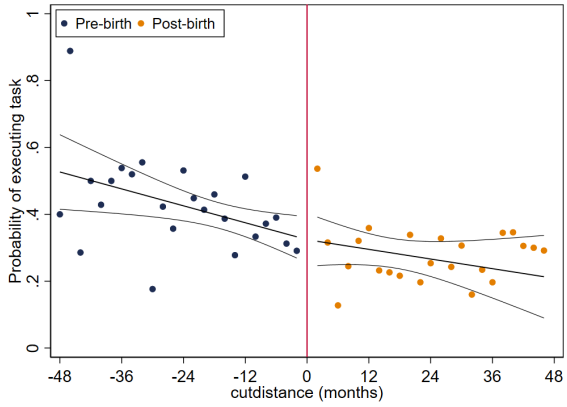
Figure 15: RD Estimates. Probability of doing task: Cleaning the house/vacuum cleaning



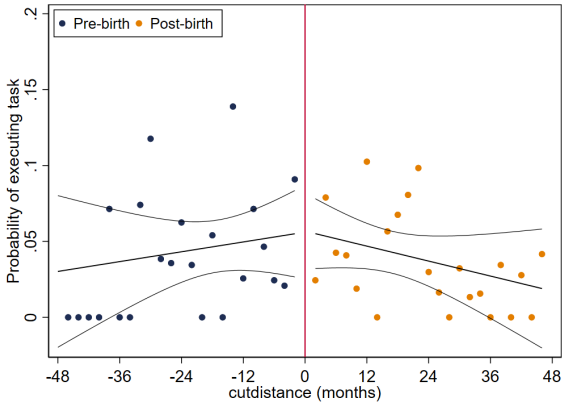
(a) Woman always or usually



(b) Man always or usually

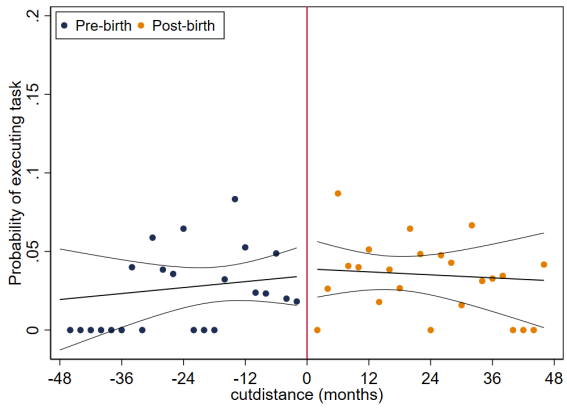


(c) Partners equally

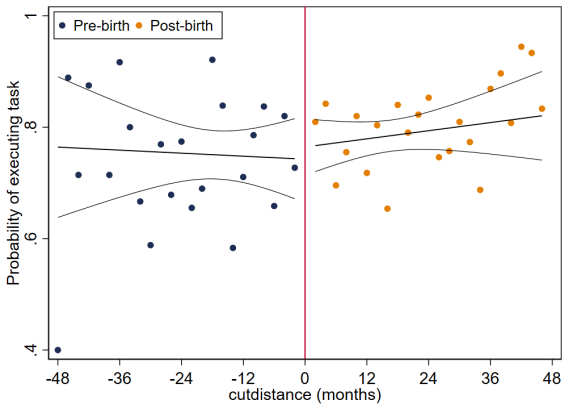


(d) Someone else always or usually

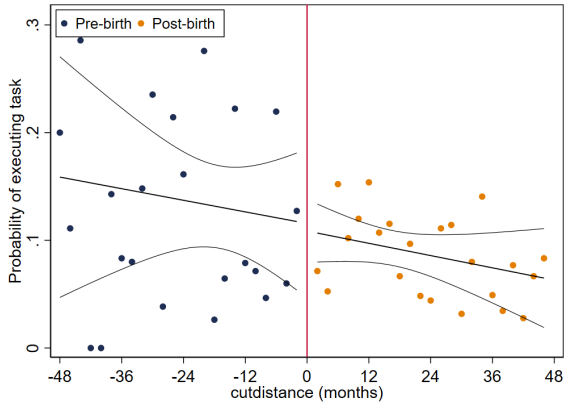
Figure 16: RD Estimates. Probability of doing task: Small repairs in, around the house



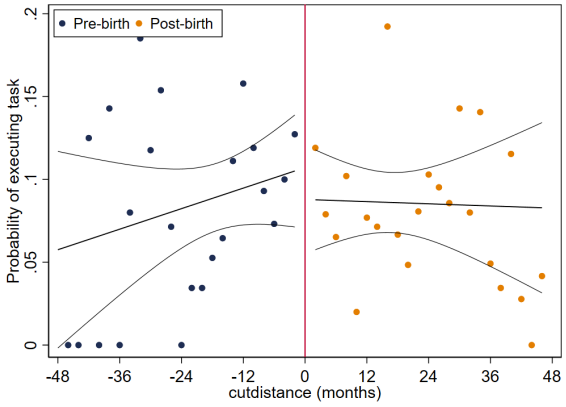
(a) Woman always or usually



(b) Man always or usually

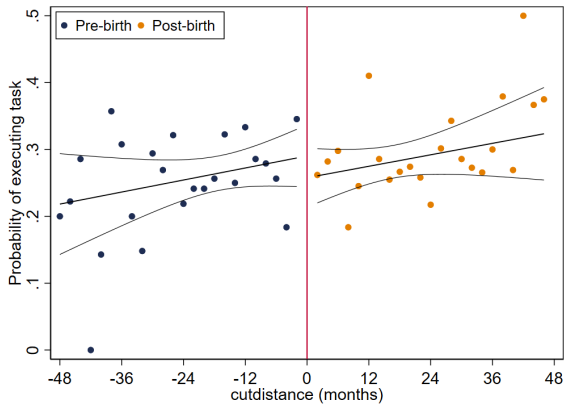


(c) Partners equally

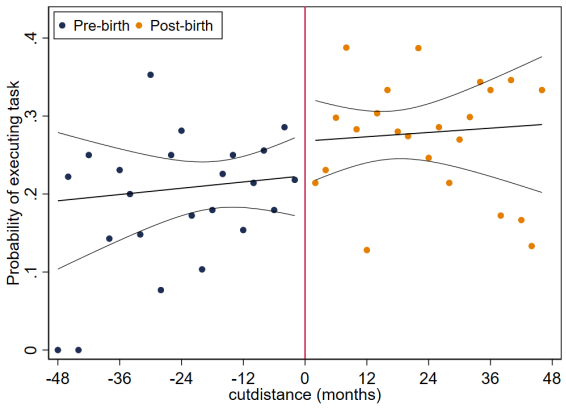


(d) Someone else always or usually

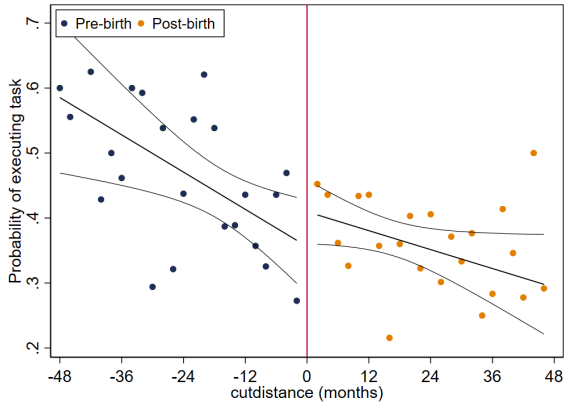
Figure 17: RD Estimates. Probability of doing task: Paying bills, financial records



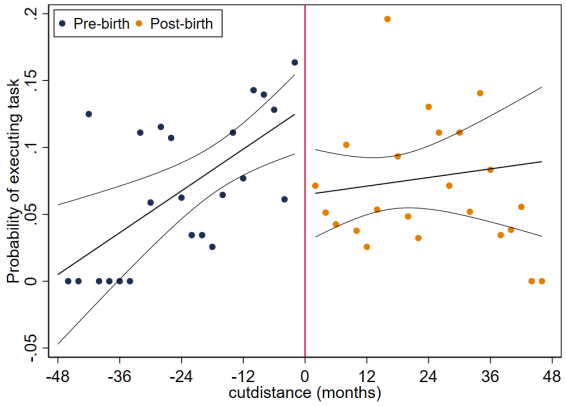
(a) Woman always or usually



(b) Man always or usually

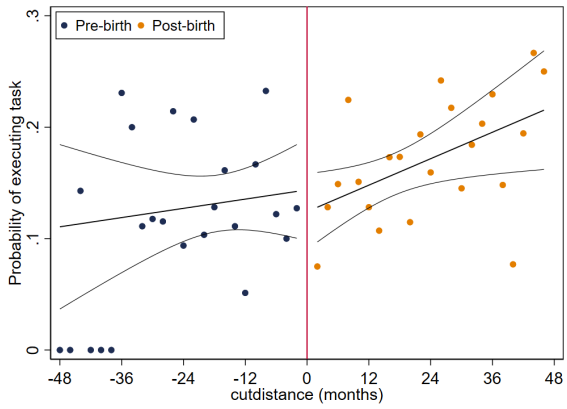


(c) Partners equally

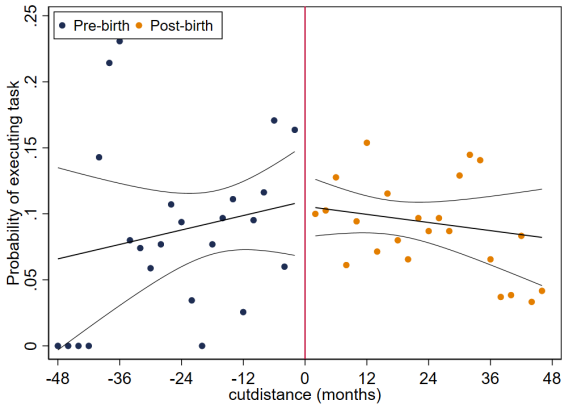


(d) Someone else always or usually

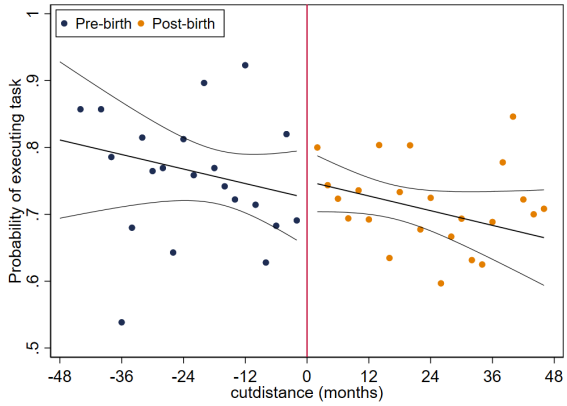
Figure 18: RD Estimates. Probability of doing task: Organizing joint social events



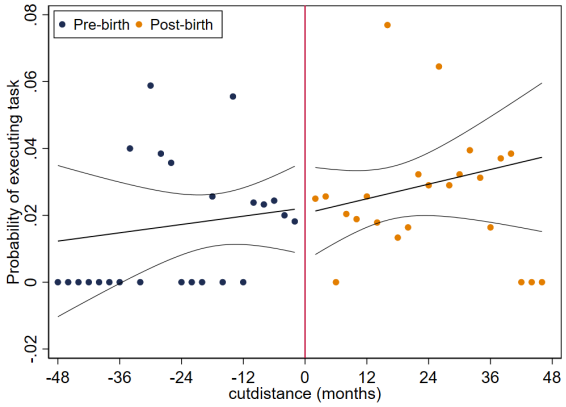
(a) Woman always or usually



(b) Man always or usually



(c) Partners equally



(d) Someone else always or usually

Table 12: Mean outcomes. Who does household tasks with children

	(1) Dressing	(2) Putting to bed	(3) Illness	(4) Leisure activities
Woman	0.76	0.66	0.84	0.32
	(0.03)	(0.03)	(0.02)	(0.03)
Man	0.01	0.04	0.03	0.02
	(0.01)	(0.01)	(0.01)	(0.01)
Equally	0.23	0.31	0.12	0.65
	(0.03)	(0.03)	(0.02)	(0.03)
Someone else	0.00	0.00	0.01	0.00
	(0.00)	(0.00)	(0.01)	(0.00)
Kids	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	282	283	272	281

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Who does household tasks with children, always or usually?

Table 13: Test for continuity in baseline covariates, by sex of respondent

	Traditional	Western	Age - Woman	Tertiary - Woman	Age - Man	Tertiary - Man
Panel A:						
Female Respondents	0.002	-0.038	-0.092	-0.076	0.668	0.314***
	(0.104)	(0.098)	(0.672)	(0.092)	(0.780)	(0.086)
Observations	316	316	316	316	316	316
Panel B:						
Male Respondents	-0.155	-0.018	1.579**	-0.065	0.092	0.130
	(0.124)	(0.117)	(0.714)	(0.112)	(0.886)	(0.104)
Observations	231	231	231	231	231	231

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table regress birth on a series of covariates. Estimates for the effect of birth on the age and education covariates are included for both men and women. "Traditional" refers to the dummy used as a proxy for sex role attitudes. It switches to 1 when the respondent agrees with a statement about working women. "Western" refers to the dummy describing whether the household resides in a western country. For both men and women the estimates for age and tertiary education are presented.

Table 14: RD Estimates. Probability of doing household task, always or usually

	(1) Preparing meals	(2) Doing dishes	(3) Food shopping	(4) Cleaning	(5) Repairs	(6) Record keeping	(7) Organizing events
Woman	0.026	0.014	-0.164**	-0.231***	0.003	-0.079	0.023
	(0.077)	(0.080)	(0.074)	(0.080)	(0.029)	(0.071)	(0.057)
Man	-0.049	-0.046	-0.070	0.061	0.007	-0.127*	-0.086*
	(0.031)	(0.041)	(0.051)	(0.053)	(0.069)	(0.068)	(0.051)
Both equally	0.035	0.059	0.226***	0.164**	-0.026	0.227***	0.033
	(0.061)	(0.070)	(0.081)	(0.077)	(0.050)	(0.078)	(0.073)
Someone else	-0.012	-0.028	0.008	0.006	0.017	-0.022	0.030
	(0.050)	(0.042)	(0.043)	(0.037)	(0.046)	(0.046)	(0.024)
Observations	570	563	571	567	565	568	569

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each column presents regression results decomposed into the different household tasks. Each row represent binary outcomes (=1) if the specified person is always/or usually doing the household task and zero otherwise. We include controls for the male and female household members age and whether they have obtained tertiary education or not and whether the respondent answering the survey was female as opposed to male. Controls for whether the respondent reside in a western European country, has traditional gender role views and a interaction term between the sex of the respondent is also added. All models are weighted by a triangular kernel. Each cell in the table present the coefficient of interest, which is Birth. This is the effect of giving birth on the likelihood of executing the specific household task always or usually for the potential task takers. The option equally refers to when the partners does it equally. Here the always/ usually description does not apply.

Table 15: Caption

	(1) Dressing	(2) Putting to bed	(3) Illness	(4) Leisure activities
Woman	0.70	0.52	0.78	0.19
	(0.05)	(0.06)	(0.05)	(0.04)
Man	0.01	0.03	0.05	0.01
	(0.01)	(0.02)	(0.03)	(0.01)
Equally	0.29	0.46	0.17	0.79
	(0.05)	(0.06)	(0.04)	(0.05)
Someone else	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Kids	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	79	79	76	78

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 16: Mean outcomes. Who does household tasks, always or usually?

	(1) Preparing meals	(2) Doing dishes	(3) Food shopping	(4) Cleaning	(5) Repairs	(6) Bills, financial records	(7) Organizing events
Panel A:							
Woman	0.57	0.47	0.30	0.45	0.03	0.26	0.13
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Man	0.05	0.10	0.10	0.09	0.75	0.21	0.09
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
Equally	0.27	0.35	0.53	0.41	0.13	0.45	0.76
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Someone else	0.11	0.08	0.07	0.05	0.09	0.08	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	651	648	652	648	648	649	652
Panel B:							
Woman	0.73	0.62	0.36	0.62	0.03	0.29	0.17
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Man	0.04	0.05	0.13	0.07	0.80	0.28	0.09
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Equally	0.14	0.27	0.43	0.27	0.09	0.36	0.71
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Someone else	0.09	0.06	0.07	0.04	0.08	0.08	0.03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Observations	1213	1204	1214	1203	1205	1211	1204

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 17: RD Estimates. Amount of tasks, by relative education levels

	(1) Woman	(2) Man	(3) Equally	(4) Someone else
Woman tertiary				
Birth	0.247	-0.008	-0.399	0.160
	(0.471)	(0.328)	(0.466)	(0.354)
Observations	157	157	157	157
Man tertiary				
Birth	0.318	0.119	-0.556	0.119
	(0.339)	(0.231)	(0.358)	(0.285)
Observations	308	308	308	308
Couple tertiary				
Birth	-0.148	-0.071	0.120	0.099
	(0.255)	(0.188)	(0.290)	(0.178)
Observations	435	435	435	435

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Heterogeneous effects across relative education levels. These outcome variables tell how many tasks a certain person within the household (man, woman, man and woman equally, someone else) always or usually takes on the responsibility for. The regression makes use of a triangular kernel weighting system. The table show coefficients for the variable of interest, "Birth", which should capture the effect of giving birth on the amount of tasks the specific household member is responsible for. We include controls for the male and female household members' age, as well as control for the sex of the respondent. Controls for whether the respondent reside in a western European country, has traditional gender role views, conditional on the sex of the respondent, are also added.

Table 18: RD Estimates. Probability of doing household task, always or usually

	(1) Preparing meals	(2) Doing dishes	(3) Food shopping	(4) Cleaning	(5) Repairs	(6) Bills, financial records	(7) Organizing events
Woman	0.015	0.066	-0.033	-0.014	0.005	-0.043	-0.003
	(0.039)	(0.041)	(0.039)	(0.041)	(0.016)	(0.038)	(0.031)
Man	0.001	-0.036*	-0.018	0.011	0.001	0.020	-0.020
	(0.018)	(0.021)	(0.027)	(0.024)	(0.036)	(0.036)	(0.025)
Both equally	-0.035	-0.034	0.046	-0.009	0.004	0.057	0.013
	(0.032)	(0.037)	(0.042)	(0.039)	(0.026)	(0.041)	(0.038)
Someone else	0.019	0.004	0.005	0.012	-0.010	-0.034	0.010
	(0.025)	(0.021)	(0.022)	(0.018)	(0.024)	(0.023)	(0.013)
Observations	1863	1851	1865	1850	1852	1859	1855

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each column presents regression results decomposed into the different household tasks. Each row represent binary outcomes (=1) if the specified person is always/or usually doing the household task and zero otherwise. We include controls for the male and female household members age and whether they have obtained tertiary education or not and whether the respondent answering the survey was female as opposed to male. Controls for whether the respondent reside in a western European country, has traditional gender role views and a interaction term between the sex of the respondent is also added. All models are weighted by a triangular kernel. Each cell in the table present the coefficient of interest, which is Birth. This is the effect of giving birth on the likelihood of executing the specific household task always or usually for the potential task takers. The option equally refers to when the partners does it equally. Here the always/ usually description does not apply.