Togetherness in the Household*

Sam Cosaert† Alexandros Theloudis‡ Bertrand Verheyden§

Revised version: October 11, 2021

Abstract

Spending time together with a spouse is a major gain from marriage. We extend the classical collective model of the household to allow for togetherness between spouses. Togetherness takes the form of joint leisure and joint care for children. Using revealed preferences conditions and Dutch data over years 2009-12, we find that households are willing to pay €1.2 per hour -10% of the average wage- to convert private leisure to joint, and €2.1 per hour to convert private childcare to joint. Our results suggest togetherness is an important component of household time use despite being overlooked in the economics literature.

Keywords: Joint leisure, joint childcare, private time use, collective model, irregular work, family labor supply, revealed preferences, gender gap, inequality, LISS.

JEL classification: D11, D12, D13, J22.

---

*We thank Olivier Bargain, Richard Blundell, Laurens Cherchye, Matthieu Delpierre, Thomas Demuynck, Olivier Donni, Mette Gørtz, Daniel Hamermesh, Jonas Jessen, Arthur Lewbel, Frederic Vermeulen and seminar and conference participants in various locations for many useful comments. We also thank the coeditor Leslie Marx and four anonymous referees for insightful comments that helped improve this paper. We are grateful to Società Italiana di Economia Pubblica for the 2019 prize in memory of Maria Concetta Chiuri. This research is part of the TIMING project of Cosaert and Theloudis, supported by the Luxembourg National Research Fund (C18/SC/12668748).

†Department of Economics, University of Antwerp; email: sam.cosaert@uantwerpen.be

‡Department of Econometrics & OR, Tilburg University; email: a.theloudis@gmail.com

§Luxembourg Institute for Socio-Economic Research; email: bertrand.verheyden@liser.lu

1 Introduction

This paper studies how couples with children allocate their time across paid work, leisure, and childcare. A distinguishing feature is our focus on togetherness: we divide the time each spouse spends on leisure and childcare to *private* (time spent alone) and *joint* (time spent together). In doing so, we admit that togetherness naturally requires spouses to synchronize their schedules so as to be physically together at the same time. We provide the first nonparametric characterization of togetherness which, among other things, allows us to quantify the value households assign to togetherness. Using data from Dutch households, we document large amounts of joint leisure, we recover bounds on joint childcare that is not directly observed in the data, and we find households value togetherness substantially. We conclude that togetherness is an important component of household time use despite it being relatively overlooked in the economics literature.

Togetherness. Spending time together with a partner is a major source of gain from marriage (e.g. Becker, 1973; Hamermesh, 2000). However, with the exception of Fong and Zhang (2001), Del Boca et al. (2019), and Browning et al. (2021), models of household time use typically abstract from togetherness; as a result, we know very little about how households value togetherness, what benefits and costs it accrues, or how it interacts with other time uses. Addressing these points is precisely our main goal.

We develop a theoretical model for household time use and consumption in which leisure and childcare comprise activities that spouses can carry out privately or jointly. Private activities are carried out by each spouse alone while joint activities are carried out simultaneously and together by both. For example, private leisure may include a lone stroll in the park while joint leisure may include a night out together at the movies. Private childcare may involve feeding a baby or tutoring a young schoolchild while joint childcare may involve both parents playing together with their child.[1]

Togetherness has benefits and costs. Joint leisure may be desirable by spouses on the grounds of companionship but may entail the loss of flexibility in the labor market. Consider a household in which the husband has a regular work schedule but the wife, perhaps due to the nature of her job, has an irregular or unpredictable work schedule (e.g. a doctor being on call). This household likely enjoys less togetherness than another one in which both spouses have regular work schedules. Togetherness requires synchronization of schedules which may be impossible if the timing of work is irregular, thus impossible without restricting one's flexibility at work and possibly reducing her earnings. An important implication therefore is that togetherness cannot be fully studied without also considering the joint *timing* of market work. Similarly, joint childcare may be beneficial

[1] Feeding a baby or tutoring a schoolchild may be carried out also jointly. These examples are offered to motivate ideas and their categorization as private or joint is not crucial for our analysis.

to children on child developmental grounds but may *also* entail the loss of benefits from specialization in the household. Children typically need attention for a given amount of time, say κ hours, which usually requires only one adult. Suppose that each parent has $\kappa/2$ hours available for childcare. If both supply it privately, they supply κ in total. If they supply it jointly, they only offer $\kappa/2$ while for the remaining time care must be provided by another, perhaps costly, caregiver.

Model. We start from a collective labor supply model in the spirit of Chiappori (1988). The collective model provides a flexible description of household behavior with the main assumption being that the decision-making process is Pareto efficient. Father, mother and children have different utility functions and the parents' bargaining powers and degree of caring for children may vary across households. Each parent values own private leisure and leisure jointly with the spouse. Each parent's private childcare and their joint childcare enter child utility. We focus on childcare as the main type of household work because parental involvement with children has important consequences for child development.[2] We do not restrict the way private and joint times enter utilities, thus we associate each such time use with, possibly, a different marginal effect. In this way joint times may yield marginal utility that differs from that of their private counterparts.

We capture the possibility that togetherness incurs a loss of flexibility at work by introducing different work schedules and giving spouses a choice over them. Specifically, we divide market work into two types: work that follows a *regular* schedule over time, and work that is *irregular*, i.e. that does not follow a regular schedule. We allow spouses to choose the amount of each type of work, regular or irregular, thereby also choosing their total market hours. Finally, we assume that the precise schedule of *irregular* work is unpredictable, as if workers are on-call or the schedule is determined by the employer at short notice. In line with Mas and Pallais (2017) workers have no control over the timing of *irregular* work, thus making synchronization of such work between spouses hard or impossible. Togetherness requires synchronization, therefore it limits workers' opportunity to do irregular work and decreases their flexibility in the labor market.

We capture the possibility that togetherness incurs the loss of benefits of specialization in the household by introducing mandatory childcare over a certain amount of time. This reflects the need of young children for continuous attention and generalizes easily to other non-market tasks. Childcare can be provided by parents or other caregivers. Togetherness means that the parents themselves care jointly for their child, which however limits their

[2] The impact of parental time on child development has been studied extensively, e.g. Cunha et al. (2010), Del Boca et al. (2014). These studies do not consider joint parental time. Del Boca et al. (2014) estimate, however, a complementary relationship between private parental childcare in child development. A special case of such complementarity is joint childcare in our paper. More recently, Del Boca et al. (2019) consider joint childcare as an input in the development process while Cano et al. (2019) and Le Forner (2021) show that joint childcare has substantial benefits for children's cognitive skills.

capacity to specialize in the provision of this service. Our model allows us to monetize both *forgone specialization* and *forgone flexibility* by means of forgone income, in a way that will be made clear subsequently.

Implementation based on revealed preference theory. We derive a nonparametric characterization of togetherness and establish testable necessary and sufficient conditions for data consistency with collective rationality. We apply our test to time use and consumption data from the Longitudinal Internet Studies for the Social sciences (LISS) in the Netherlands. We use cross-sectional variation in wages and allow for general heterogeneity across groups of households.

Our nonparametric approach fits in the literature on revealed preference characterizations of the collective model by Cherchye et al. (2007, 2009, 2011), offering three main advantages. First, it allows us to test consistency in a robust way, without resorting to parametric form assumptions. Any such assumption would inevitably restrict how togetherness interacts with other choices or how it accrues utility to household members. Given that togetherness is still largely understudied, we feel it is overly restrictive to discipline it through specific functional forms.[3] Second, it allows us to bound the fraction of time that spouses spend jointly on childcare as this is rarely observed in the data. Third, it yields global identification results (see Chiappori and Ekeland, 2009, for a discussion of identification in a collective environment).

Contribution & results. We offer four main contributions. First, we extend the classical collective model to allow for togetherness in the form of joint leisure and, separately, joint childcare. The closest paper to ours, Browning et al. (2021), estimates a collective model with joint leisure. Their main goal is to investigate the substitutability between private and joint leisure. Using Danish data on both types of leisure and a flexible parametric specification for utility, they find little substitution in practice. Unlike Browning et al. (2021), we consider various dimensions of togetherness: between two spouses (joint leisure), between parent and child (private childcare), and between both parents and child (joint childcare, as in Del Boca et al., 2019). Moreover, we admit that togetherness requires synchronization of time schedules and we explicitly model this.[4] Therefore, we let household members determine the amount of market work as well as, in part, its schedule. We formalize the gains and costs of joint leisure and joint childcare relative to their private counterparts and we allow the value of togetherness to differ from that of private time. This is because togetherness in our model is naturally associated with forgone specialization and forgone flexibility that Browning et al. (2021) abstract from.

[3]For example, Browning et al. (2021) assume separability between leisure time and household expenditure, often rejected in the data (e.g. Blundell et al., 2016). Our approach naturally relaxes this.

[4]There is empirical evidence that couples synchronize their work (e.g. Sullivan, 1996; Hallberg, 2003).

Second, we provide a nonparametric characterization of togetherness. To the best of our knowledge, this is the first paper to do so. This exercise is theoretically appealing for reasons explained above but also relevant in practice. We find that our model explains time use and consumption data substantially better than the classical collective model that does not distinguish between private and joint times (67% versus 6% pass rates).

Third, we show set identification of joint childcare when only total childcare is available per spouse. While Fong and Zhang (2001) provide original identification results for joint leisure (which we observe in our data), we extend identification to a second joint time use. In our most informative setting, parents spend between 3.8 and 11.9 weekly hours on joint childcare, representing up to 92% of total childcare of one parent. This result is useful when interest lies not only in the amount but also in the nature of childcare.

Fourth, we quantify the value of togetherness: we compute the household's willingness to pay to swap one unit of private time by each spouse with one unit of joint time. On average, households pay at least €1.22 per hour -10% of the average wage- to convert private leisure to joint, and at least €2.08 per hour -17% of the average wage- to convert private childcare to joint. Joint childcare is more expensive because it entails the loss of flexibility in the labor market *and* specialization in the household. So if parents freely choose joint childcare it must be because they value it relatively more.

Our results suggest that women are the likeliest to forgo work flexibility in order to increase togetherness. Work conditions (e.g. a historical wage gap) may have women work less. Our model describes how, for the sake of togetherness, this creates an incentive for women to restrict their work schedule, reduce irregular work, and forgo income associated with flexibility in the market. This reduces their earnings and reinforces the gender wage gap. This is in line with Goldin (2014) and Olivetti and Petrongolo (2016) who argue that the gender wage gap would be lower if women were as flexible as men in their work schedules. In spite of this, we argue that togetherness mitigates intra-household inequality in the sharing of resources, at least vis-à-vis a setting with private times only.

Related literature. Since Gary Becker's treatise on the family (Becker, 1973, 1981), household economics studies decisions made in the family.[5] Our model falls in the collective approach: formalized in Apps and Rees (1988) and Chiappori (1988, 1992), this approach respects methodological individualism but remains empirically tractable. While this literature has also studied home production (Chiappori, 1997) and public goods (Blundell et al., 2005), it is scarce in studying time use.[6] A distinct literature on the economics of time use, inaugurated by Becker (1965), is concerned with the allocation of time across work and non-work activities. Recent research defines and measures im-

[5] Extensive recent reviews of this literature are Donni and Chiappori (2011), Chiappori and Mazzocco (2017), and Greenwood et al. (2017).

[6] Browning and Gørtz (2012), Cherchye et al. (2012), Lise and Yamada (2019), and Browning et al. (2021) are notable exceptions. Only the fourth paper studies joint leisure.

portant concepts such as demand for time (Hamermesh and Biddle, 2018), the timing of work (Hamermesh, 1999), time strain (Hamermesh and Lee, 2007) and more. Our paper partly bridges the two strands of literature: along with modeling private and joint time in the household, we address Hamermesh (1998)'s critique that most theories and empirical analyses of time focus on the duration of a given activity rather than also its timing.

Our focus on togetherness should not be surprising. There has long been evidence that couples like spending time together (Sullivan, 1996; Hamermesh, 2000; Qi et al., 2017). Workers use regular time away from work in order to synchronize schedules within the family (Brown et al., 2011). However, synchronization is not always possible when the timing of work is unpredictable or determined at short notice. Cousins and Tang (2004) and Ruppanner and Maume (2016) show that employees in countries with shorter working hours report more work-family conflict. It turns out that the benefits of shorter workweeks quickly evaporate when associated with more variability/less predictability in hours worked (Voorpostel et al., 2010). Mas and Pallais (2017) found that employees are unwilling to pay for more control over the timing of work when the default option is to work regular hours but they will pay substantially to avoid working outside their regular hours. This is consistent with demand for togetherness which, as we argued above and develop subsequently, requires that spouses synchronize their work schedules.

Our interest in joint childcare is warranted by the implications it has for child development. Joint care implies a higher concentration of adults around the child (Folbre et al., 2005), which enhances communication (Bronfenbrenner, 1979) and the sense of closeness in the family (Crouter et al., 2004). Cano et al. (2019) and Le Forner (2021) study the impact of joint parental time on children's cognitive and non-cognitive skills; in both papers joint time is paramount for children's verbal skills and communication.[7] Del Boca et al. (2019) estimate a production function for child development and find that joint childcare is more productive than each parent's private childcare after age 5, and more productive than the *sum* of parents' private times after age 9. Joint care can be beneficial also to parents as it associates with higher parental well-being (Flood et al., 2020) and better performance in parental tasks (Stueve and Pleck, 2001). So the evidence points to joint childcare being an important actor in child development. However, the presence of children in practice often decreases parental time synchronization (van Klaveren and van den Brink, 2007; Barnet-Verzat et al., 2011; Bryan and Sevilla, 2017), perhaps due to the pressure childcare puts on parents' time (Presser, 1994). This is why it is important to study joint leisure alongside parents' work and childcare duties, as we do here.

The rest of the paper is organized as follows. Section 2 provides motivating empirical facts. Section 3 presents our model, section 4 the nonparametric characterization, section 5 the results, and section 6 offers several extensions. We conclude in section 7.

[7] In addition, Milkie et al. (2015) find that family time improves adolescents' self-concept and math performance and decreases drug use and behavioral problems. See also Offer (2013).

2 Basic empirical patterns

The main data we use come from the Longitudinal Internet Studies for the Social sciences (LISS) in the Netherlands (CentERdata, 2012). The LISS started in 2007 tracking a nationally representative sample of approximately 4,500 households with the aim to gather information on their members' life course and living conditions. The core study of the panel is repeated yearly collecting data on household composition, members' incomes, employment, housing etc., similar to other major surveys such as the PSID. In addition to the core study, researchers can propose additional questions to which participants respond up to a monthly basis. This feature makes the LISS rich on information spanning economic outcomes, time use, the timing of certain activities, and many more.

We draw data from three waves, covering calendar years 2009, 2010, and 2012. We need information on household consumption and members' time use, which are consistently available only in those years. Within each wave we observe certain variables with monthly frequency (income, employment) but most other variables only once (consumption, time use). Therefore we have complete information at most in three points in time. Data are collected by means of online survey questions requiring respondents to recall information from the past (e.g. a month or week). Cherchye et al. (2012) designed the first wave of the consumption and time use data and we refer to them for details.[8]

Our main sample includes married or permanently cohabiting couples with children up to 12 years old as, given our focus on childcare, it is unreasonable to admit families with older children. We further restrict this sample to households where both adult members participate in the labor market. While this introduces some selection (among couples with young children, 98% of men and 76% of women work), it is necessary to do so as our revealed preferences approach requires wages for both spouses. Our final sample includes 398 household×year observations. This may look small but it is comparable to other recent studies that require consumption and time use at the household level (e.g. Cherchye et al., 2012; Browning et al., 2021). We present summary statistics in section 5.1, details about the data and the sample selection in appendix A.1, and a comparison of our sample with one where we relax the market participation requirement in appendix A.2. Time allocations in our sample are remarkably similar to the larger sample.

A unique feature of our data is that we observe the time each spouse spends on leisure activities. Therefore, unlike most other data, we do not have to calculate leisure as the remainder time after accounting for work and other tasks. Moreover, we directly observe the fraction of leisure that spouses spend together with one another, which is what we call joint leisure.[9] Figure 1 sorts the households in our sample with respect to aggregate

[8]Further information about the LISS as well as access to the data is available at www.lissdata.nl. Households without Internet access were provided a connection and, if needed, a computer.

[9]The leisure questions in the data are *How much time did you spend in the last seven days on leisure activities?*, which we use for total leisure of each spouse, and *You have indicated that you spent [the*

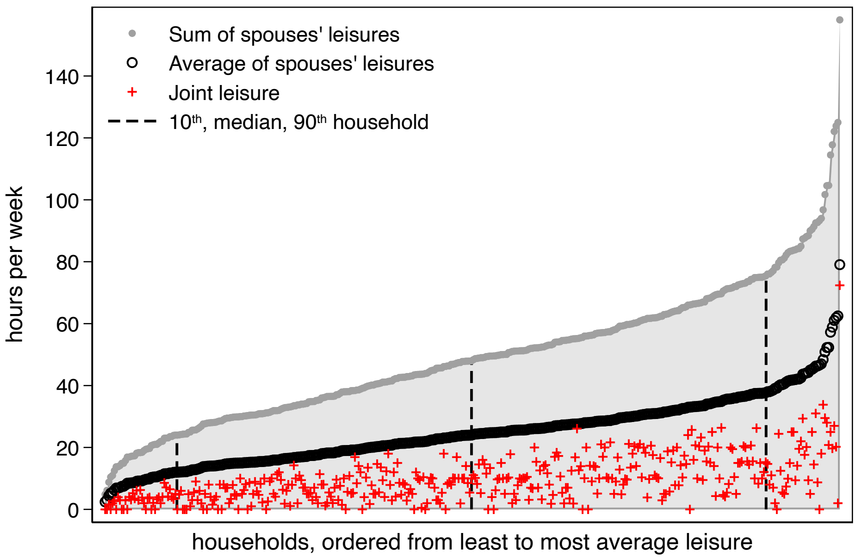


Figure 1 – Weekly leisure: aggregate, average, joint

leisure, namely the *sum* of spouses' total leisure times. Additionally it plots the *average* leisure between spouses, namely half the aggregate. Two points stand out. First, average leisure in the middle 80% of households is between 12 and 38 hours per week (24 for the median household). Second, most couples (91.7% of them) have some joint leisure. In the middle 80% of households this is between 1.5 and 20 hours per week, meaning that a large fraction of average leisure between spouses is actually joint leisure. As a result, adding up spouses' total leisures (top gray line in figure 1) overestimates how prevalent leisure is in a typical week as it double counts joint leisure.

Joint leisure in our data is retrospective so it may be subject to recall error. However, the LISS also has a small time diary in year 2013 where respondents use a mobile application to record their real-time time use. We do not use this information because of lack of accompanying consumption data and an extremely small sample size. However, we check the diary and confirm that joint leisure is prevalent and numerically close to our baseline figures above. We discuss the time diary further in appendix A.2.

Unlike joint leisure, we do not observe joint childcare. However, we do observe the total time each spouse spends on childcare (comprising private and joint childcare) and we plot it in figure 2. There are separate questions in the survey about time spent on chores and several other activities, so we are confident that the childcare data indeed reflect time devoted to children. Three points stand out. First, while mothers are the main carer of young children, fathers contribute to childcare substantially: the median mother supplies

answer to the first question] on leisure time activities, in the seven days preceding today. Please indicate how much of that time you spent together with your partner?, which we use for joint leisure. In case the spouses disagree on the amount of joint leisure, we use the minimum of their answers (see appendix A.1). Note that the questionnaires were administered in September/October and early December so we miss any potential seasonal patterns in leisure, for example due to summer holidays.

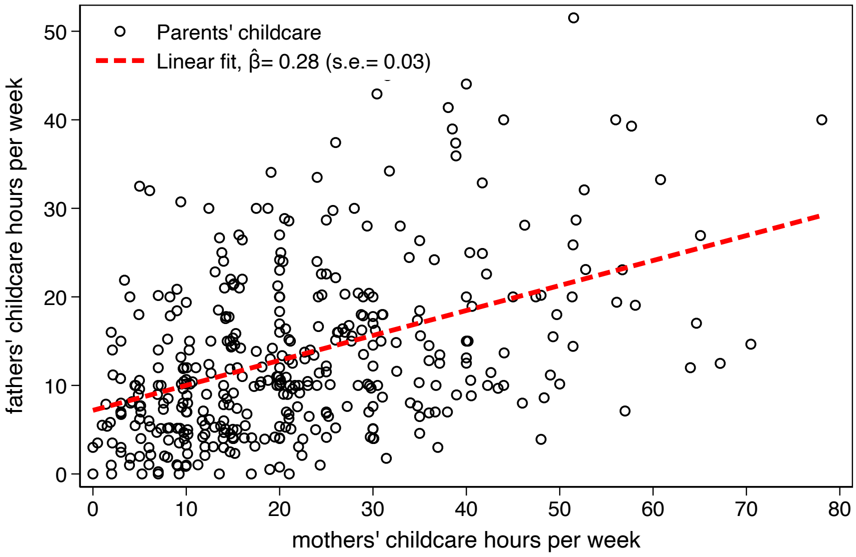


Figure 2 – Weekly childcare: mothers and fathers

20 weekly hours while the median father supplies 10.6. Second, only 0.5% of mothers and 2% of fathers do not do any childcare at all. In the majority of households *both* parents supply strictly positive hours, which is a necessary condition for a joint component in childcare. Third, mothers' and fathers' hours are positively correlated. While there are several underlying structures that may generate this correlation, one of them is when a large part of parents' childcare is joint. Our model will help assess this claim.

As we do not observe *joint* childcare in our data, we cannot say for sure what parents actually do during joint childcare. However, the survey asks parents about who engages in certain childcare activities, which we plot in figure 3. In the vast majority of households parents engage *equally* in going out with the children or talking with and advising them.

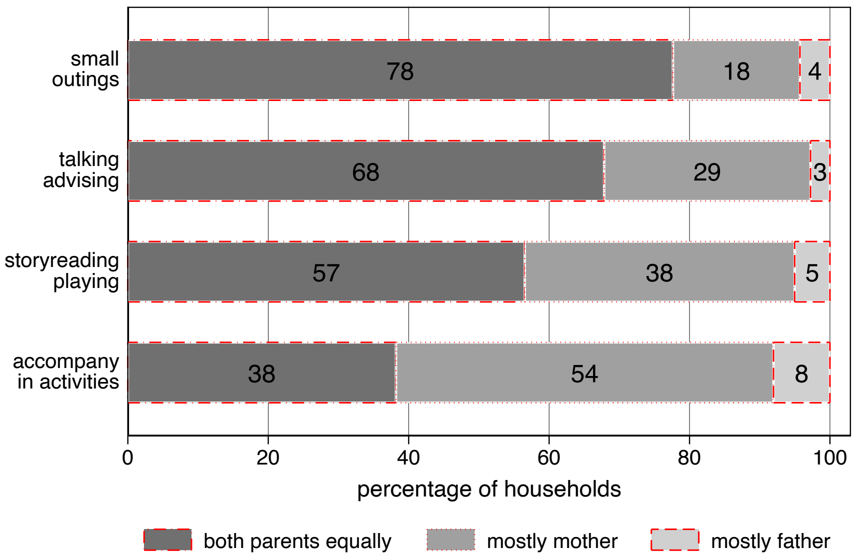


Figure 3 – Weekly childcare: engagement of parents in select childcare activities

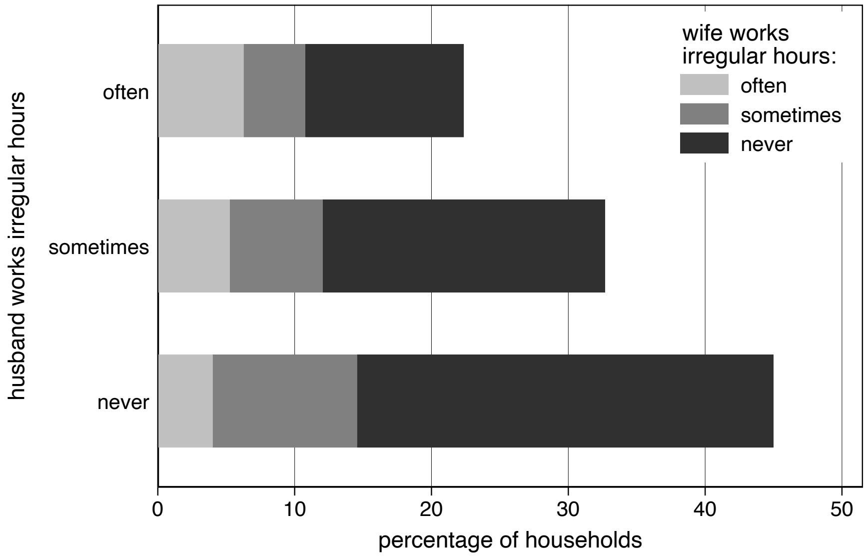


Figure 4 – Market work: incidence of irregular work

Of course this information is not conclusive about joint childcare (the parents could devote equal time to a given activity but carry it out privately) but it is *suggestive* that going out or talking/advising are candidate components of joint childcare in our data. This agrees with external information on joint childcare. Cano et al. (2019) use the Longitudinal Study of Australian Children and report that about a quarter of parents' joint childcare concerns educational activities (reading, playing educational games, lessons) while the rest concerns outings (walking, riding bicycle, visiting people and attending parties), eating together, and bathing.[10]

While the evidence so far points to some demand for togetherness, togetherness is not possible without the couple being physically together at the same time. The spouses may differ in the number of hours they devote to market work or the timing of such work. While we do not observe the precise timing of market work, we do observe the extent to which work follows a regular or irregular schedule. An irregular schedule likely limits the ability spouses have to synchronize their work and spend time together. Figure 4 plots the extent to which the spouses engage in irregular work.[11] Work for a large fraction of individuals, namely for about 55% of men and 37% of women, is *sometimes* or *often* irregular. Moreover, while the incidence of irregular work is prevalent in sectors such as retail food or health services (sectors known for irregular schedules and whose workers often report work-life conflict, e.g. Henly and Lambert, 2014; Golden, 2015), it is widespread across education groups and varies over time in a given household. Figure 5

[10] Cano et al. (2019) report that joint care takes up about 83% of total childcare of the father, which is very close to our estimate that joint care represents up to 92% of total childcare of one parent.

[11] The question is *Now follows a list of aspects that can characterize a job. Please indicate whether the following aspects are often, sometimes or never characteristic for your job. [...] Do you work irregular hours?* Other questions on the list include whether the job is tiring, mentally demanding etc.

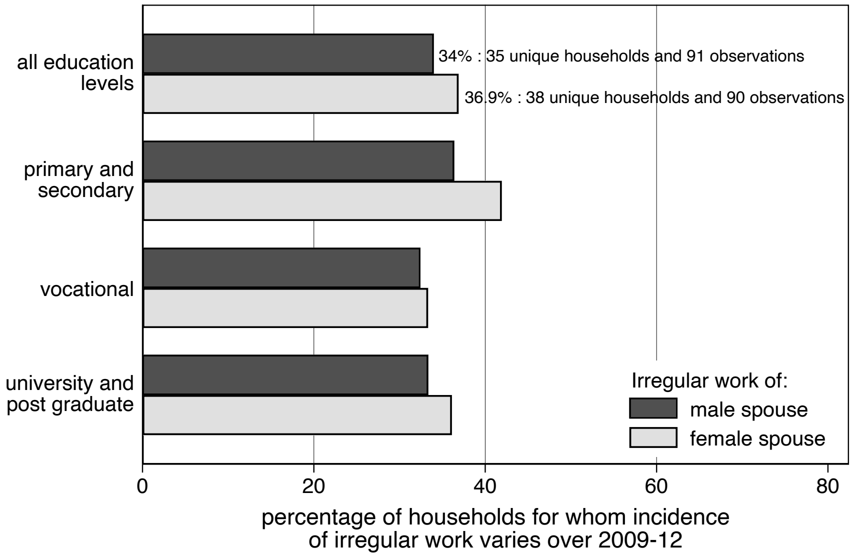


Figure 5 – Market work: within-subject variation in incidence of irregular work

Notes: The figure plots the percentage of households by education group for whom the incidence of irregular work varies over the length of our data. The incidence of irregular work takes categorical values for 'often', 'sometimes', 'never', as in figure 4. We measure this variation out of 103 unique households (contributing a total of 250 observations) that appear at least twice over 2009-2012.

illustrates that the incidence and frequency of irregular work in about 35% of households varies in the short term (over 2009-12). While there is some correlation with job switching, most of this variability occurs *without* people changing jobs or occupation.[12] This is important because it shows that one's incidence and frequency of irregular work is not a fixed attribute (on which spouses could perhaps match at marriage) but varies substantially with time. Our data are thus indicative of a large extent of possible misalignment between spouses' timing of market work. This misalignment matters for togetherness as it hampers the possibility the spouses spend time jointly. Our model allows for such misalignment enabling us to assess its implications for family time use.

Before introducing the model, we present descriptives for leisure and childcare. Table 1 presents the mean, median and bottom/top deciles of the distributions of leisure and childcare. It confirms that joint leisure is a substantial component of overall leisure while childcare is prevalent among both mothers and fathers. Appendix table A.1 presents correlations among these time uses. Childcare correlates positively between parents (echoing figure 2) but negatively with one's overall leisure reflecting a trade-off between childcare and leisure. By contrast, childcare correlates positively with *joint* leisure so families in which joint leisure is high, parental childcare is also high. If parental childcare increases with joint childcare, then togetherness is a possible explanation behind this pattern as it increases both joint leisure *and* parental childcare through its joint component.

Table 2 presents linear regressions of leisure and childcare on the gender wage gap,

[12] The probability of a job change among those for whom the incidence of irregular work is fixed over 2009-2012 is about 4.5% while it is 5.7% (men) and 7.9% (women) among those for whom it varies.

Table 1 – Descriptive statistics: weekly leisure and childcare

	mean	median	10^{th} pct.	90^{th} pct.
leisure male	25.9	25.0	9.0	43.0
leisure female	23.6	21.0	8.9	41.4
joint leisure	9.5	8.2	1.5	20.0
childcare male	13.3	10.6	3.1	26.9
childcare female	21.7	20.0	5.1	41.7

Notes: The table reports the average, median, 10^{th} and 90^{th} percentiles of leisure and childcare. All statistics are calculated over 398 household/year observations. See appendix A.1 for details on the sample and variable definitions.

the incidence of irregular work, and select demographics. A larger wage gap in favor of men associates with higher male and lower female leisure (perhaps due to improving the male's intra-household bargaining power), and with lower female childcare (perhaps due to higher wage women devoting more time to their children, e.g. Chiappori et al., 2017). The gender *age* gap does not really matter. Women's incidence of irregular work decreases joint leisure substantially, consistent with our argument that irregular work hampers the couple's possibilities for togetherness. Appendix table A.2 repeats these regressions introducing controls for education and occupation of each spouse. While post-

Table 2 – Linear regressions: weekly leisure and childcare

	(1) leisure male	(2) leisure female	(3) joint leisure	(4) childcare male	(5) childcare female
gender wage gap	0.012	-0.005	0.000	0.002	-0.014
	(0.005)	(0.006)	(0.006)	(0.008)	(0.008)
1[irregular work male]	0.059	-0.021	-0.062	0.101	0.032
	(0.072)	(0.066)	(0.086)	(0.093)	(0.072)
1[irregular work female]	-0.164	-0.116	-0.178	0.135	0.076
	(0.075)	(0.071)	(0.089)	(0.094)	(0.075)
gender age gap	-0.012	-0.004	-0.003	0.005	0.007
	(0.011)	(0.011)	(0.013)	(0.015)	(0.014)
1[child 4-6 yrs]	-0.049	0.099	-0.004	-0.204	-0.172
	(0.094)	(0.084)	(0.105)	(0.110)	(0.089)
1[child 7-12 yrs]	0.045	0.149	0.005	-0.592	-0.747
	(0.079)	(0.074)	(0.094)	(0.101)	(0.097)
constant	3.136	2.980	2.190	2.497	3.123
R^2	0.03	0.02	0.02	0.11	0.21
# of observations	394	395	365	390	396

Notes: The table reports coefficients and standard errors from linear regressions of log leisure and childcare on the gender wage gap (male−female wage), the gender age gap (male−female age), indicator variables for whether a spouse works irregular hours *at least sometimes* and for children's age (the excluded age category is children less than 4 years old), and a constant. Standard errors are clustered at the household level. The constant is always significant at the 1% level. The number of observations varies because not everyone has positive amounts of leisure and childcare (see e.g. figure 1 for joint leisure). Results in levels are qualitatively similar. See appendix A.1 for details on the sample and variable definitions.

graduate education seems to be associated with higher levels of joint leisure (reflecting perhaps better marital match among highly educated persons), education and occupation do not otherwise affect leisure or childcare. By contrast, the incidence of irregular work still affects leisure (and childcare, but less so) despite the large number of education and occupation controls. This is an important insight as it indicates that work schedules affect time use above and beyond one's education or occupation.

3 Collective model with private and joint time use

3.1 Model preliminaries

In the model a household consists of two adult members, subscripted by $m = \{1, 2\}$, and a number of young children. Each adult member, i.e. each spouse, is endowed with $\mathcal{T}_m$ units of time to allocate to leisure L_m, childcare T_m, and market work H_m. Therefore, each spouse has a time budget given by

$$L_m + T_m + H_m = \mathcal{T}_m. \tag{1}$$

The spouses may differ in time endowments $\mathcal{T}_m$ due to differences in the amount of sleep, personal care, or other uses of time that we abstract from such as household chores.[13]

Leisure and childcare comprise activities that the spouses can carry out *privately* or *jointly*. Let l_m denote spouse m's private leisure and l_J denote joint leisure. Similarly, let t_m denote m's private childcare and t_J joint childcare. In total, spouse m spends

$$\begin{aligned} l_m + l_J &= L_m \\ t_m + t_J &= T_m \end{aligned} \tag{2}$$

time on leisure and childcare respectively. This distinction between private and joint time use is in the epicentre of our paper. With the exception of time diaries and a limited number of surveys, however, private and joint times are typically not observed separately. In our data we observe joint leisure but we do not observe joint childcare. One of our contributions is to provide restrictions that separate private from joint childcare when only the *sum* is observed (Fong and Zhang, 2001, separate private from joint leisure).

Market work is strictly private. The spouses, however, have a choice over the amount of hours H_m as well as the timing or type of such work in a way that we define below.[14]

[13] This does not rule out that individuals also spend time on chores, at least so long as chores are weakly separable from other uses of time in the utilities we introduce subsequently.

[14] The cross-sectional distribution of market hours in the LISS is continuous between 25 and 45 weekly hours for men, and 10 and 40 hours for women. The mode is at 40 and 25 hours respectively. While this is comparable with the hours distribution in the US, as reported e.g. in the PSID, an advantage of our data is that it records the *weekly* hours of work, which people can arguably recall more easily. We treat the continuous distributions as evidence that people have choice over their market hours.

There are two reasons why we introduce (an element of) timing. First, the incidence of irregular or asynchronous work between spouses is empirically relevant (Hamermesh and Stancanelli, 2015, also figure 4) and has implications for family choices and welfare (Craig and Powell, 2012). Second, togetherness requires spouses to be physically together at the same time. If their work schedules are asynchronous, and this asynchronicity is driven by incentives or restrictions in the labor market, then their capacity to spend time jointly may be limited. Therefore the timing of market work is crucial for togetherness.

We operationalize timing as follows. We divide market work into two types: work that follows a *regular* schedule over time, and work that is *irregular*, i.e. that does not follow a regular schedule. Regular work is scheduled at specific hours and days and its precise timing is known or predictable. Irregular work is the opposite: its precise timing is not known in advance and may change from one week to the other. While we do not need to restrict regular or irregular work to specific times, the following examples may help fix ideas. Work is regular for a bank clerk who always only works 9am to 5pm on weekdays. Work is also regular for a part-time hotel concierge who works 11am to 10pm on Saturdays and Sundays. However, when the concierge is unexpectedly called in on a weekday (and perhaps receives a day off as a compensation), then this part of her work is irregular. Finally, a nurse who faces constantly changing shifts does irregular work. He is aware of his total hours of market work over a given period, but he does not know the exact timing of his work in advance, e.g. he cannot predict the exact timing of his shifts, say, four weeks ahead. Irregular work in our data is prevalent in (but not limited to) sectors such as retail food, culture, or health care, in which workers often report work-life conflict due to unpredictable schedules (e.g. Henly and Lambert, 2014; Golden, 2015).[15]

The spouses choose how many hours they supply to each type of work. We use h_m^R to denote m's hours of regular work and h_m^I to denote her hours of irregular work such that

$$h_m^R + h_m^I = H_m. \tag{3}$$

Regular work pays a wage w_m while irregular work pays a wage $w_m + p_m$, i.e. it pays a premium p_m. We think of this premium as weakly positive (i.e. $p_m \geq 0$) in order to reflect the compensation a worker may have to receive for the inconvenience (defined formally in section 3.3) of working outside her regular schedule. By allowing spouses a choice over regular and irregular work we will capture the main conceptual restriction relevant to togetherness, namely that time cannot be joint unless spouses synchronize to be physically together at the same time.

Each spouse has utility U_m over private and joint leisure, and parental consumption

[15] While the incidence of irregular work is more prevalent in certain jobs, appendix table A.2 establishes that irregular work affects leisure and childcare even after controlling for an array of occupations.

C_P, given by

$$U_m(l_m, l_J, C_P).^{16}$$

Children also have utility U_K which is a function of their parents' private childcare, joint childcare, and child consumption C_K, and is given by

$$U_K(t_1, t_2, t_J, C_K).^{17}$$

Private and joint times enter separately in U_m and U_K so as to associate such times with possibly different marginal utilities. If togetherness yields benefits to the household over and above private leisure or private childcare, then our formulation allows for this while still retaining traditional complementarities between different time uses. In other words, joint time does not necessarily have the same impact on utility as private time.

Finally, a standard budget constraint connects expenditure (parental consumption, child consumption, market childcare T_K) to non-labor income Y and earnings, given by

$$C_P + C_K + w_K T_K = Y + \sum_{m=1}^{2} w_m h_m^R + \sum_{m=1}^{2} (w_m + p_m) h_m^I. \tag{4}$$

As is typical, we treat the price of parental and child consumption as the numeraire. We price market childcare at w_K (for example, the wage of an external carer).[18]

3.2 Costs of togetherness

The potential benefits of togetherness stem from joint time having marginal utility that differs from that of private time, reflecting companionship and child developmental benefits respectively. Its costs pertain to *forgone flexibility* and *forgone specialization.*

Forgone flexibility. Forgone flexibility corresponds to the cost of togetherness overall, be it in the form of joint leisure or joint childcare. The idea in a nutshell is as follows: togetherness requires both spouses to be together at the same time. This is feasible if they synchronize their work schedules in such a way so as to be away from market work at the same time. In an environment where irregular work pays a premium but workers are unable to anticipate its precise timing or cannot plan this work in advance,

[16] Cosaert and Hennebel (2019) allow parental childcare to have an intrinsic leisure component but their unitary setting is not suitable to study togetherness in the household. Our choice to model consumption in the way we do is quite general as it allows C_P to be interpreted as purely public consumption but also, alternatively, as private consumption with positive externalities between spouses.

[17] U_K captures utility of all children in the household combined; see however section 6.2.

[18] We exclude market childcare T_K from children's utility U_K in order to simplify our nonparametric characterization in section 4. The exclusion effectively assumes that children's marginal utility of market childcare is zero, i.e. $\partial U_K / \partial T_K = 0$. Nevertheless, parents still purchase T_K because, as we show subsequently, children may require care at times when the parents cannot provide such care themselves.

synchronization may be impossible without restricting the flexibility of workers with respect to the type of work they do, thus without reducing their earnings.

More formally, togetherness requires to account for the *joint timing* of market work, something that the spouses' time budgets alone do not do. We make two conjectures:

- The spouses fully synchronize their regular hours h_m^R between them, so these hours always overlap. We think of the timing of regular work as stable or predictable over time, making it possible for spouses to synchronize their schedules between them. This is not saying that the spouses work the same amount of regular hours: one may work less and another more. But the person who works less is at work at the same time when her partner also is. This automatically implies that whoever is at work *for longer* limits the maximum amount of joint time available to the couple.

- The spouses cannot synchronize their irregular hours h_m^I between them or synchronize these hours with the other's regular work. We think of the timing of irregular work as unpredictable, determined at short notice by the employer or the circumstances (e.g. the concierge having to show up on weekdays to meet the occasional increased demand). Irregular work is associated with workers being on-call or generally unable to anticipate its precise schedule (e.g. Mas and Pallais, 2017). This makes it hard or impossible for spouses to synchronize their shifts between them, therefore irregular work by *either* spouse limits the maximum amount of time they can possibly spend jointly. We present in section 5.4 an alternative scenario in which individuals have some discretion to adjust their schedule during irregular hours and have it overlap partly with the partner's. We start, however, with the strong no-overlap assumption because it helps us crystallize our main ideas.

Putting these conjectures together, the maximum amount of time the spouses can possibly spend jointly (leisure and childcare) is given by

$$l_J + t_J \leq \mathcal{T}_m - \max\left\{h_1^R, h_2^R\right\} - h_1^I - h_2^I \tag{5}$$

and proven in appendix B.1. Maximum togetherness equals total time $\mathcal{T}_m$ (by construction the minimum $\mathcal{T}_m$ in the couple) but is reduced by the amount and type of spouses' market work. The more they work, the less time they can spend together. An increase in regular work by the person who works the most reduces maximum togetherness one-to-one; an increase by the spouse who works the least leaves it unchanged. By contrast, an increase in irregular work by either spouse reduces maximum togetherness one-to-one because irregular work is asynchronous.

Two remarks are in order. First, an increase in togetherness beyond the upper bound above requires at least one spouse to give up part of her market hours, thus forgo labor earnings. Although both types of work enter the boundary condition in (5), irregular work tightens togetherness *the most* because of our assumption that spouses cannot synchronize there. This facilitates the empirical implementation of our model because, as we show subsequently, the premium of irregular work conveniently disciplines the associated forgone flexibility. Second, our model allows that part of market work overlaps between spouses and another, empirically smaller, part does not. Thus, it offers a *generalization* to an environment where hours strictly only overlap or strictly do not. Without fundamentally changing our approach, we present in section 5.4 an extension in which individuals have some flexibility to adjust their schedule during irregular hours.

Forgone specialization. Forgone specialization is the cost of joint childcare. The idea in a nutshell is as follows: keeping everything else constant, the spouses must sacrifice one unit of private childcare t_1 *and* one unit of private childcare t_2 in order to obtain one more unit of joint childcare t_J. So the household must forgo two units of private childcare in total to gain only one unit of joint childcare.

More formally, specialization in our context is associated with the presence of young children in the household who require attention and care for $\mathcal{T}_K$ units of time. This gives rise to a childcare constraint (similar to a child time budget), given by

$$\sum_{m=1}^{2} t_m + t_J + T_K = \mathcal{T}_K. \tag{6}$$

The constraint involves all childcare choices that we explicitly model and reflects the loss of benefits of specialization when parents engage in joint childcare. Suppose each spouse has $\mathcal{T}_K/2$ hours available to look after children. If the spouses specialize and supply childcare privately, they supply $\mathcal{T}_K$ hours in total. But if they look after the children simultaneously and together, they supply $t_J = \mathcal{T}_K/2$ hours only. Even if such joint hours may be more beneficial to children and more pleasant to parents, the spouses need additional childcare to fill the gap of the remaining $\mathcal{T}_K/2$ hours. Having exhausted their own hypothetical hours, they must source such childcare elsewhere, possibly on the costly market (i.e. T_K). Forgone specialization thus incurs a cost of t_J hours that can be monetized in different ways, for example at the price of market childcare.

Childcare constraint (6) stems solely from admitting that children require care for $\mathcal{T}_K$ hours, as if it reflects a child's time budget. We allow $\mathcal{T}_K$ to be different from $\mathcal{T}_m$ because children do not typically require parental or market care during the entire day: grandparents may also contribute to childcare, older children spend time in school or with friends etc.[19] Similarly, the parents may be able to adapt $\mathcal{T}_K$ through care offered by other family members or through passive care by themselves while engaging in time

[19]Interestingly, writing down the child's time budget precisely allows us to identify joint and private childcare by means of a system of 3 equations (two parental and one child's time budgets) in 3 unknowns (t_J, t_1, t_2). Unfortunately this is not possible without precise information on $\mathcal{T}_m$ and $\mathcal{T}_K$.

uses that we abstract from (e.g. chores), etc. We thus consider $\mathcal{T}_K$ as the remaining time children require care for, net of all these other times.

Three remarks are in order here. First, the childcare constraint in an environment where togetherness is *not* explicitly modeled becomes $\sum_{m=1}^{2} T_m + T_K = \mathcal{T}_K$; this double-counts the time spent jointly and overestimates the amount of time for which children receive parental attention. Our constraint (6) is nested within this latter constraint only if $t_J = 0$; otherwise it improves over it by correctly counting private and joint childcare separately. Second, although we are slightly abusing the terminology, there are parallels between classical specialization (i.e. division of labor) and our notion of temporal specialization. The cost of togetherness that stems from the childcare constraint is mainly due to costly redundancies (loss of opportunities for division of childcare duties over the day/week) rather than loss of specialization in the pure sense. Finally, we have purposefully introduced the costs of togetherness prior to showing the household problem so as to highlight that such costs are independent of the decision process in the household.

3.3 Household problem and optimal time allocation

The household maximizes a weighted sum over the spouses' and child respective utilities in the traditional collective spirit of Chiappori (1988, 1992) and Dunbar et al. (2013). It chooses $\mathcal{C} = \{l_1,\ l_2,\ l_J,\ t_1,\ t_2,\ t_J,\ h_1^R,\ h_2^R,\ h_1^I,\ h_2^I,\ C_P,\ C_K,\ T_K\}$ in order to

$$\max_{\mathcal{C}}\ \mu_1 U_1(l_1, l_J, C_P) + \mu_2 U_2(l_2, l_J, C_P) + \mu_K U_K(t_1, t_2, t_J, C_K) \tag{P}$$

subject to

$$\begin{aligned}
& l_m + l_J + t_m + t_J + h_m^R + h_m^I = \mathcal{T}_m, \qquad m = \{1, 2\} \\
\lambda :\quad & C_P + C_K + w_K T_K = Y + \textstyle\sum_{m=1}^{2} w_m h_m^R + \sum_{m=1}^{2} (w_m + p_m) h_m^I \\
\tau_J :\quad & l_J + t_J \le \mathcal{T}_m - \max\left\{h_1^R, h_2^R\right\} - h_1^I - h_2^I \\
\tau_K :\quad & \textstyle\sum_{m=1}^{2} t_m + t_J + T_K = \mathcal{T}_K
\end{aligned}$$

and non-negativity constraints on consumption. The constraints, in the order they appear, are: the spouses' time budgets (1) after plugging in the leisure/childcare and market hours identities (2) and (3); the money budget (4); the upper bound on togetherness (5); and the childcare constraint (6).

Consistent with the collective approach to household behavior, each spouse is associated with a utility weight or intra-household bargaining power μ_m and children are associated with a utility weight μ_K. The latter reflects the degree that children's welfare matters for the household although it can be seen also as the weight parents put on a child development production function à la Del Boca et al. (2014). In fact the two formulations are equivalent, as also is one in which each parent enjoys utility from their children's utility through caring. We show this formally in appendix B.2.

Spousal bargaining powers and the weight on children's welfare may depend on wages, demographics, or distribution factors (variables that affect the utility weights but not preferences or the budget set; Bourguignon et al., 2009). Demographics will likely also affect the utility functions. Here we suppress this dependence in order to ease the notation but we address both issues in the empirical implementation of our model.

The parameters on the left of the constraints are their shadow prices. Technically, they reflect the marginal value of money or time accrued to (or expended by) the household while shifting a constraint at the margin. For example, τ_J is the marginal value of maximum synchronous time given the couple's work hours and schedules. To relax this constraint, a spouse must restrict her labor market flexibility (optimally give up irregular work as we explain below) so τ_J also reflects the cost of forgone flexibility. Similarly, τ_K is the marginal cost of childcare duties. To relax this constraint, the household can purchase market childcare so τ_K reflects the marginal value of market childcare; to tighten it, the household may increase joint childcare at the expense of private childcare so τ_K also reflects the cost of forgone specialization. The first order conditions help illustrate these equivalent interpretations. As usual, λ is the marginal value of total income.

Optimal time allocation. The household problem is convex so the first order conditions are necessary and sufficient. We present them explicitly below in order to characterize the benefits and costs of togetherness.

Let $U_{m,\mathfrak{c}}$ be the marginal utility of spouse m and $U_{K,\mathfrak{c}}$ the marginal utility of children, both with respect to variable $\mathfrak{c} \in \mathcal{C}$. Assuming $h_1^R > h_2^R$ without losing generality, and using the spousal time budgets in lieu of private leisure, the optimality conditions are

$$
\begin{aligned}
[l_J]: &\quad \mu_1 U_{1,l_J} + \mu_2 U_{2,l_J} = \mu_1 U_{1,l_1} + \mu_2 U_{2,l_2} + \tau_J \\
[t_m]: &\quad \mu_K U_{K,t_m} \leq \mu_m U_{m,l_m} - \tau_K \\
[t_J]: &\quad \mu_K U_{K,t_J} \leq \mu_1 U_{1,l_1} + \mu_2 U_{2,l_2} + \tau_J - \tau_K \\
[h_1^R]: &\quad \lambda w_1 = \mu_1 U_{1,l_1} + \tau_J \\
[h_2^R]: &\quad \lambda w_2 = \mu_2 U_{2,l_2} \\
[h_m^I]: &\quad \lambda(w_m + p_m) \leq \mu_m U_{m,l_m} + \tau_J \\
[C_P]: &\quad \mu_1 U_{1,C_P} + \mu_2 U_{2,C_P} = \lambda \\
[C_K]: &\quad \mu_K U_{K,C_K} = \lambda \\
[T_K]: &\quad \tau_K = \lambda w_K.
\end{aligned}
$$

The left hand side across all equations represents the marginal benefit of the respective variable while the right hand side its marginal cost. To obtain these, we assume interior solution for leisure (95% of households in the data have $l_m > 0$ and 91.7% have $l_J > 0$), regular hours (everyone in our sample works at least some regular hours), consumption (so the multipliers on the consumption non-negativity constraints become irrelevant) and market childcare (both parents are away at regular work simultaneously), but we allow for corners in the unobserved parental childcare variables as well as in irregular hours.

These equations describe the optimal family time allocation that satisfies the following features in equilibrium (all benefits and costs below are *marginal*):

- The benefit of joint leisure, captured by the weighted sum of the spouses' marginal utilities of l_J, equals the sum of two costs: the weighted sum of the spouses' marginal utilities of forgone private leisure and forgone flexibility when togetherness is at the upper bound of (5). Rearranging we obtain $\tau_J = \mu_1(U_{1,l_J} - U_{1,l_1}) + \mu_2(U_{2,l_J} - U_{2,l_2})$: in equilibrium, forgone flexibility τ_J equals the benefits of togetherness given by joint leisure's *additional* marginal utility over private leisure. We subsequently monetize forgone flexibility by means of forgone earnings.

- The benefit of private childcare is given by the child's marginal utility of t_m. Its cost equals the marginal utility of forgone private leisure, net of the savings τ_K the household achieves from sparing one unit of *market* childcare. The latter stems from the childcare constraint (6) and we monetize it subsequently.

- The benefit of joint childcare is given by the child's marginal utility of t_J. Its cost is the weighted sum of *both* spouses' marginal utilities of forgone private leisure, plus forgone flexibility τ_J as above, net of the savings τ_K in market childcare.

- The benefit of one regular hour of market work equals the wage rate w_m normalized by the price of consumption λ. Its cost is the marginal utility of forgone private leisure by whoever supplies such work and, in the case of spouse 1, the forgone benefits of togetherness given by τ_J. The latter arises because an additional hour of regular work by spouse 1 restricts togetherness by tightening constraint (5) (recall assumption $h_1^R > h_2^R$).

- The benefit of one irregular hour of market work equals the wage plus premium rate $w_m + p_m$ normalized by the price of consumption λ. Its cost is the marginal utility of forgone private leisure and the forgone benefits of togetherness given by τ_J. The latter is so because irregular work always restricts togetherness by tightening (5).

- The marginal utility of any type of consumption equals its price λ.

- Finally, the benefit τ_K of market childcare (which monetizes the savings the household achieves from providing childcare internally in the household) is equal to its cost given by the normalized unit price w_K.

There remains one last shadow price to characterize, namely τ_J or forgone flexibility when condition (5) binds. The simplest way to do this is when spouse 2 has some irregular hours. Then, a combination of the optimality conditions for h_2^R and h_2^I yields

$$\tau_J = \lambda p_2.^{20}$$

This indicates that joint leisure at the margin costs a (normalized) premium p_2. When maximum possible togetherness is tight, the household can enjoy more joint time if spouse 2 reduces her irregular work, that is if she gives up p_2. It is spouse 2 who gives up her premium because, unlike her husband, she does not have to give up the full $w_2 + p_2$ to increase togetherness; instead she can optimally replace irregular with regular work (thus still earn w_2 in lieu of $w_2 + p_2$) and still loosen (5). This is because we have assumed $h_1^R > h_2^R$ in this illustration, so an increase in h_1^R by spouse 1 will not loosen (5). The case of $h_1^R < h_2^R$ is analogous. This summarizes the core of our cost argument of togetherness when, due to irregular or unpredictable schedules, the spouses cannot fully synchronize their market work: at least one of them must forgo irregular work, thus forgo the associated earnings, if the couple wishes to increase togetherness.

Rearranging the first order conditions for joint leisure and all types of childcare yields

$$\mu_1(U_{1,l_J} - U_{1,l_1}) + \mu_2(U_{2,l_J} - U_{2,l_2}) = \tau_J \qquad = \lambda p_2 \tag{7}$$

$$\mu_K(U_{K,t_J} - U_{K,t_1} - U_{K,t_2}) = \tau_J + \tau_K = \lambda p_2 + \lambda w_K. \tag{8}$$

Both equations describe the relationship, in equilibrium, between the benefits and costs of joint time *over* their private counterparts. The left hand side of (7) reflects the utility benefits joint leisure renders over private leisure. These benefits come at the cost of forgone flexibility τ_J, on the right hand side, as the spouses reduce the extent of asynchronous work between them. This also provides a micro-foundation to the premium of irregular work: p_2 compensates for the inconvenience irregular work incurs by restricting togetherness in the household. The left hand side of (8) reflects the utility benefits to children when the parents engage in joint rather than private childcare. These benefits comes at the cost, on the right hand side, of forgone flexibility τ_J and forgone specialization τ_K, the latter monetized at the unit price of market childcare. Analogous statements describe the optimal family time allocation when $h_1^R < h_2^R$.

Our model implies an optimal time allocation that differs markedly from that of the standard collective model without joint times or irregular work. Given a set of prices, the optimal allocation in the standard model is determined by the marginal utility of (private) leisure and the marginal productivity of (private) childcare. In our model, the optimal allocation is also determined by the marginal utility of joint leisure, the marginal

[20] This becomes $\tau_J \geq \lambda p_2$ and serves as lower bound to τ_J, if spouse 2 does not do irregular work.

productivity of joint childcare, togetherness upper bound (5), and childcare constraint (6). The implications of our model for within-household inequality also differ from those of the classical model; we postpone this discussion to section 6.

4 Nonparametric characterization

We now formally introduce the notion of collective rationality with private and joint uses of time, and we provide a nonparametric characterization. This characterization allows us to test consistency of the model with the data without imposing specific functional forms on utility functions U_m and U_K. Moreover, it contains restrictions that enable us to separate private from joint childcare when only their sum is available per spouse. Our nonparametric conditions fit in the revealed preferences literature in line with Afriat (1967), Diewert (1973) and Varian (1982). Cherchye et al. (2007, 2009, 2011) developed the revealed preference characterization of the collective labor supply model.

We start in 4.1 by introducing the baseline characterization of collective rationality with private and joint uses of time. This characterization makes a formal distinction between private and joint times but, unlike the model of section 3, does not introduce the togetherness upper bound (5) (reflecting forgone flexibility) or childcare constraint (6) (reflecting forgone specialization). In this simpler model the gains from joint and private times are equal in equilibrium. Subsequently, we develop in 4.2 the nonparametric characterization of collective rationality that is fully consistent with our model of section 3. This is the characterization that allows us, among other things, to separate private from joint childcare. We assume throughout that utilities are well-behaved (i.e. continuous, concave, nonsatiated); no further restrictions are imposed on their functional form.

Setup. We build on the *Generalized Axiom of Revealed Preference* (GARP). The GARP constructs binary preference relationships between bundles of goods based on the affordability of these bundles. Consider hypothetical data $S = \{\mathbf{p}^{(v)};\ \mathbf{q}^{(v)}\}_{v \in V}$ with $\mathbf{p}^{(v)} \in \mathbb{R}^{|N|}_{++}$ the vector of strictly positive prices and $\mathbf{q}^{(v)} \in \mathbb{R}^{|N|}_{+}$ the bundle of consumption goods and time use in observation v. Let V be the set of observations. The following definition provides a formal statement of the revealed preference relations that apply to S.

Definition 1 Consider data $S = \{\mathbf{p}^{(v)};\ \mathbf{q}^{(v)}\}_{v \in V}$. The GARP asserts the existence of direct (R^0) and indirect preference (R) relations, defined at the *individual* utility level, such that for all $v, s \in V$:

1. if $\mathbf{p}^{(v)\prime}(\mathbf{q}^{(v)} - \mathbf{q}^{(s)}) \geq 0$ then $\mathbf{q}^{(v)}\ R^0\ \mathbf{q}^{(s)}$;
2. if $\mathbf{q}^{(v)}\ R^0\ \mathbf{q}^{(s1)},\ \mathbf{q}^{(s1)}\ R^0\ \mathbf{q}^{(s2)}, \ldots,\ \mathbf{q}^{(sk)}\ R^0\ \mathbf{q}^{(s)}$ then $\mathbf{q}^{(v)}\ R\ \mathbf{q}^{(s)}$;
3. if $\mathbf{q}^{(v)}\ R\ \mathbf{q}^{(s)}$ then not $\mathbf{p}^{(s)\prime}(\mathbf{q}^{(v)} - \mathbf{q}^{(s)}) < 0$.

In words, the first condition states that $\mathbf{q}^{(v)}$ is directly revealed preferred over $\mathbf{q}^{(s)}$ if $\mathbf{q}^{(s)}$ was in the budget set when $\mathbf{q}^{(v)}$ was chosen. The second condition imposes transitivity. The last condition closes the revealed preference argument stating that if $\mathbf{q}^{(v)}$ is revealed preferred over $\mathbf{q}^{(s)}$ (i.e. at least as good as $\mathbf{q}^{(s)}$), it must be the case that $\mathbf{q}^{(v)}$ was not strictly affordable when $\mathbf{q}^{(s)}$ was chosen. Otherwise there existed an alternative bundle $\tilde{\mathbf{q}} \neq \mathbf{q}^{(s)}$ in budget set s that would have strictly improved this consumer's wellbeing.

For the following sections, we make the notation compact by defining $\delta_{m,\mathfrak{c}} = \mu_m U_{m,\mathfrak{c}}/\lambda$ and $\delta_{K,\mathfrak{c}} = \mu_K U_{K,\mathfrak{c}}/\lambda$. Dividing by λ and multiplying by μ eliminates sensitivity of our variables to the cardinality of utility functions and utility weights. For example, δ_{m,l_m} is spouse m's marginal willingness to pay for her private leisure in the household. Finally, let us formally introduce the empirical content of our data, namely the variables that we observe in our dataset, given by

$$D_{\text{LISS}} = \left\{ w_m;\ l_m,\ l_J,\ T_m,\ h_m^R,\ h_m^I,\ C_P,\ C_K \right\}_{m\in\{1,2\}}.$$

This makes clear that we observe private leisure l_m separately from joint leisure l_J, but we only observe total childcare T_m per spouse. The main variables that we do not observe, other than the childcare components, are the work premium p_m and the price of market childcare w_K. We will explain subsequently how we get around this.

4.1 Collective rationality

We consider the simpler model that makes the explicit distinction between private and joint times but does not feature the togetherness upper bound (5) or childcare constraint (6). Definition 2 formally states our notion of collective rationality in this environment.

Definition 2 A household with observations V in D_{LISS} is *collectively rational* if and only if there exist weights $\mu_m^{(v)}$ and $\mu_K^{(v)}$, utility functions U_m and U_K, time variables $t_m^{(v)} + t_J^{(v)} = T_m^{(v)}$, for $m = \{1, 2\}$, such that the household solves problem (P) subject to time budget (1) and money budget (4).

In the absence of constraints (5) and (6) the optimality conditions associated with joint leisure and childcare become

$$\begin{aligned}
[l_J]: &\qquad \delta_{1,l_J} + \delta_{2,l_J} = w_1 + w_2 \\
[t_m]: &\qquad \delta_{K,t_m} = w_m \\
[t_J]: &\qquad \delta_{K,t_J} = w_1 + w_2
\end{aligned}$$

while the optimality condition associated with private leisure is $\delta_{m,l_m} = w_m$.[21] It follows

[21] From here onwards, the optimality conditions appear *as if* we assume an interior solution; in reality our revealed preferences approach can easily deal with corners.

that, in equilibrium, the *total* gains from joint time equal the *total* gains from (both spouses') private times. Nonetheless, the *individual* gain from joint leisure, captured by δ_{m,l_J}, may deviate from the *individual* gain from private leisure, captured by δ_{m,l_m}. After all, joint leisure takes the form of a public good within the household and the corresponding individual willingness to pay for it may vary.

Let us summarize what the goal in this characterization is, prior to showing our first proposition formally. To facilitate the exposition, let us define the variable subsets $S_m = \{\delta^{(v)}_{m,l_m}, \delta^{(v)}_{m,l_J}, \delta^{(v)}_{m,C_P};\ l^{(v)}_m, l^{(v)}_J, C^{(v)}_P\}_{v\in V}$ and $S_K = \{\delta^{(v)}_{K,t_1}, \delta^{(v)}_{K,t_2}, \delta^{(v)}_{K,C_K};\ \ T^{(v)}_1, T^{(v)}_2, C^{(v)}_K\}_{v\in V}$ for $m \in \{1,2\}$. We will test if S_m and S_K are consistent with GARP by checking if they satisfy the conditions of Definition 1. These conditions are defined at the individual utility level so S_1, S_2, and S_K serve the purpose of separating the variables that enter the male, female, and child utility respectively. Of course, some δ's in S_m are unobserved (the individual prices of joint leisure or public consumption are not part of D_{LISS}) but this is the purpose of the test anyway: if such prices exist and GARP is satisfied, then this tells us that the model fits the data and gives us information about, among other things, the value of togetherness in the household. This is the basis of testing and identification in the paper (and in similar revealed preferences settings), which will become clearer as we move along this section. We are now ready to formalize our test for collective rationality.

We obtain Proposition 1 by exploiting the problem's optimality conditions (statements *(a)*-*(d)*) and concavity of the underlying utility functions (statement *(e)*). This mimics the revealed preference conditions for collective rationality in Cherchye et al. (2007).

Proposition 1 Consider a household with observations V in D_{LISS}. The following statements are equivalent (proof in appendix C.1):

1. The household is *collectively rational.*

2. There exist shadow prices $\delta^{(v)}$ such that the following conditions hold for all $v \in V$:

   (a) $\delta^{(v)}_{m,l_m} = w^{(v)}_m$

   (b) $\delta^{(v)}_{1,l_J} + \delta^{(v)}_{2,l_J} = w^{(v)}_1 + w^{(v)}_2$

   (c) $\delta^{(v)}_{1,C_P} + \delta^{(v)}_{2,C_P} = 1$

   (d) $\delta^{(v)}_{K,t_m} = w^{(v)}_m$ and $\delta^{(v)}_{K,t_J} = w^{(v)}_1 + w^{(v)}_2$ and $\delta^{(v)}_{K,C_K} = 1$

   (e) S_1, S_2, S_K satisfy GARP as per Definition 1.

Statement *(a)* says that the individual willingness to pay for private leisure is equal to one's wage while that for joint leisure, as per *(b)*, sums up to $w^{(v)}_1 + w^{(v)}_2$. Statement *(c)* is the Lindahl-Bowen-Samuelson condition for the optimal provision of public goods, and stems from our assumption of Pareto efficiency: the individual willingness to pay for parental consumption sums up to the normalized consumption market price. Statement *(d)* says that the individual willingness to pay for private childcare is equal, as in the case of private leisure, to one's wage while that for joint childcare to the sum of wages. The willingness to pay for children's consumption equals the normalized consumption price. Finally, *(e)* states that the data S_1, S_2, S_K satisfy the requirements of the GARP. That S_K satisfies GARP is nontrivial; it follows from childcare identity (2) and the optimality conditions associated with childcare. See appendix C.1 for a detailed derivation.

We call Proposition 1 a *test* for collective rationality because it lays out necessary and sufficient conditions for the data D_{LISS} to reflect collectively rational behavior. The test, however, is written in terms of S_1, S_2, S_K and these involve also the unobserved δ's. If we find δ's for which the data in D_{LISS} satisfy the above conditions, then D_{LISS} reflects collectively rational behavior. These conditions, especially the inequalities of Definition 1, then offer us powerful restrictions that enable us to bound (i.e. set identify) certain δ's and other items as we illustrate subsequently.

One of our objectives has been to separate private from joint childcare which, unlike leisure, are not separately observed in our data. However, the characterization in Proposition 1 does not allow this as the conditions are independent of the composition of T_m. The underlying reason is that t_m and t_J are inputs into one utility function (the child's) and the gains from joint childcare in this simplified environment equal the sum of gains from private childcare. This is not the case in our full model which features the togetherness upper bound (5) and childcare constraint (6). The characterization in section 4.2 will allow us to separate private from joint childcare precisely because it exploits these additional constraints. Without them, the distinction between regular and irregular work is also meaningless and this is why Proposition 1 is independent of h_m^R and h_m^I.

Finally, the individual shadow prices of l_m are observed (the wages) but the *individual* shadow prices of l_J are not. So one may attempt to recover shadow prices δ_{m,l_J}, which are informative for the individual preferences for joint leisure. Understanding such preferences (and any gender differences in them) is important for understanding the consequences of intra-household wage variation for leisure inequality or, generally, for how leisure enters a generalized sharing rule as in Chiappori and Meghir (2014).[22]

4.2 Collective rationality with togetherness

We now introduce the nonparametric characterization of collective rationality considering all the constraints togetherness is associated with in our full model. Togetherness here may admit gains above and beyond the gains from spouses' private times. Household

[22] An alternative objective would be to separate private from joint *leisure* whenever total leisure is observed per spouse. This is in principle possible as shown by Fong and Zhang (2001) or Cherchye et al. (2013) in a revealed preferences context, even without considering the togetherness upper bound (5) or childcare constraint (6). However, treating both quantity l_J and prices δ_{1,l_J} and δ_{2,l_J} as unknown variables creates nonlinearities in the revealed preferences conditions: one can identify the price-commodity product but it is hard to separate the two. See also Cherchye et al. (2007).

members face a trade-off between these gains and the costs associated with the upper bound on joint time (5) and childcare constraint (6). As such, the term 'collective rationality with togetherness' refers hereafter to the environment with these specific constraints. Definition 3 formally states our notion of collective rationality with togetherness.

Definition 3 A household with observations V in D_{LISS} is *collectively rational with togetherness* if and only if there exist weights $\mu_m^{(v)}$ and $\mu_K^{(v)}$, utility functions U_m and U_K, time variables $t_m^{(v)} + t_J^{(v)} = T_m^{(v)}$, for $m = \{1, 2\}$, such that the household solves problem (P) as presented in section 3.3.

Let us summarize again what the goal in this characterization is. To facilitate the exposition, we again define variable subsets $S_m = \{\delta_{m,l_m}^{(v)}, \delta_{m,l_J}^{(v)}, \delta_{m,C_P}^{(v)};\ l_m^{(v)}, l_J^{(v)}, C_P^{(v)}\}_{v \in V}$ and $\widetilde{S}_K = \{\delta_{K,t_1}^{(v)}, \delta_{K,t_2}^{(v)}, \delta_{K,t_J}^{(v)}, \delta_{K,C_K}^{(v)};\ \ t_1^{(v)}, t_2^{(v)}, t_J^{(v)}, C_K^{(v)}\}_{v \in V}$ for $m \in \{1, 2\}$. We will test if S_m and $\widetilde{S}_K$ are consistent with GARP by checking if they satisfy the conditions of Definition 1. Recall that these conditions are defined at the individual utility level so S_1, S_2, and $\widetilde{S}_K$ serve the purpose of separating the variables that enter the male, female, and child utility respectively. $\widetilde{S}_K$ differs from S_K in the previous characterization because it involves the (unobserved) childcare components t and the (unobserved) childcare shadow prices $\delta_{K,t}$. GARP in Definition 1, together with the problem's first order conditions, offer restrictions that will help us discipline these items. So let us now formalize our main test for collective rationality which, like previously, we obtain combining the problem's optimality conditions with concavity of the underlying utility functions.

Proposition 2 Consider a household with observations V in D_{LISS}. The following statements are equivalent (proof in appendix C.2):

1. The household is *collectively rational with togetherness.*

2. There exist shadow prices $\delta^{(v)}$, wage premiums $p_m^{(v)}$, price of market childcare $w_K^{(v)}$, and time variables $t^{(v)}$ such that the following conditions hold for all $v \in V$:

   (a) $\begin{matrix} \delta_{1,l_1}^{(v)} = w_1^{(v)} - p_2^{(v)} \\ \delta_{2,l_2}^{(v)} = w_2^{(v)} \end{matrix}$ if $h_1^{R(v)} > h_2^{R(v)}$ and $\begin{matrix} \delta_{1,l_1}^{(v)} = w_1^{(v)} \\ \delta_{2,l_2}^{(v)} = w_2^{(v)} - p_1^{(v)} \end{matrix}$ if $h_1^{R(v)} < h_2^{R(v)}$

   (b) $\delta_{1,l_J}^{(v)} + \delta_{2,l_J}^{(v)} = w_1^{(v)} + w_2^{(v)}$

   (c) $\delta_{1,C_P}^{(v)} + \delta_{2,C_P}^{(v)} = 1$

   (d) $\delta_{K,t_m}^{(v)} = \delta_{m,l_m}^{(v)} - w_K^{(v)}$ and $\delta_{K,t_J}^{(v)} = w_1^{(v)} + w_2^{(v)} - w_K^{(v)}$ and $\delta_{K,C_K}^{(v)} = 1$

   (e) $t_m^{(v)} + t_J^{(v)} = T_m^{(v)}$

   (f) S_1, S_2, $\widetilde{S}_K$ satisfy GARP as per Definition 1.

Statements *(a)-(d)* above play a similar role to statements *(a)-(d)* in Proposition 1, while statement *(f)* is the reference to GARP like statement *(e)* in Proposition 1.

Proposition 2 has two distinctive features. First, the characterization involves not only wages w_m but also premium p_m. This is because our full model features the upper bound on joint time (5), and regular and irregular work affect this differently. We explained in section 3 that, when $h_1^R > h_2^R$, both regular and irregular work by spouse 1 put the same pressure on togetherness. By contrast, irregular work by spouse 2 puts *more* pressure on togetherness than regular work, at least for a given total number of hours, simply because spouse 2 can optimally increase togetherness by replacing one irregular with one regular hour at cost p_2. In this case the benefit of joint over spouses' private leisure is equal to the cost of forgone flexibility p_2 by spouse 2, in equilibrium. This is reflected in statements *(a)* & *(b)*. On the flip side, the agent who does irregular work even though she could optimally replace it with regular and improve togetherness, must be compensated (p_2) for the loss of togetherness due to her work. This explains why p_2 matters but p_1 drops out in the scenario where $h_1^R > h_2^R$. An analogous statement is true when $h_1^R < h_2^R$.[23]

The second distinctive feature is that the necessary and sufficient conditions are no longer independent of the components of T_m or market childcare price w_K. This is because our full model features the childcare constraint (6), t_m and t_J present trade-offs that matter for it, and market childcare enters the constraint. This is reflected in statement *(d)*, which defines the childcare shadow prices, and statement *(e)*, which requires the unobserved childcare components to add up to observed T_m.

We call Proposition 2 a *test* for collective rationality with togetherness because it lays out necessary and sufficient conditions for the data D_{LISS} to reflect collectively rational behavior. As in the earlier proposition, the test is written in terms of S_1, S_2, $\widetilde{S}_K$ that involve also unobserved δ's and unobserved childcare times t's. But the main idea behind testing is precisely this: if we find δ's and t's for which the data in D_{LISS} satisfy the above conditions, then D_{LISS} reflects collectively rational behavior. What is crucial for testing is the upper bound on joint time (5) and childcare constraint (6). The constraints drive a wedge between the benefits of private and joint times, evident from inspecting statements *(a)* & *(b)* (leisure) and *(d)* (childcare), enabling these time uses to have different implications for behavior. These type-of-time-unique shadow prices, together with the GARP conditions of Definition 1, offer us powerful restrictions to bound the δ's and the t's. This is the main idea behind set identification in our setting.[24]

[23]When $h_1^R > h_2^R$, our model cannot explain why spouse 1 does *regular* work if he can simply do *irregular* work and earn more. So we must have $p_1 = 0$ (so p_1 drops out of the first order conditions) or the household would be better off if spouse 1 worked irregular hours only. But everyone in our data does some regular work. To reproduce this while keeping $p_1 > 0$, our model would require an upper bound on irregular hours of spouse 1, perhaps stemming from labor market regulations or limits to the needs of employers. Technically, this constraint would introduce corners at both ends of spouse 1's choice set for hours; at the upper end $\lambda(w_1 + p_1) \geq \mu_1 U_{1,l_1} + \tau_J$, which clearly does not force p_1 to zero. As we do not want to complicate the model further, we simply assume $p_1 = 0$ when $h_1^R > h_2^R$.

[24]Proposition 2 disciplines the shadow price of joint childcare δ_{K,t_J}, which subsequently allows us to identify t_J (e.g. example 2 in section 4.3). By contrast, if l_J was unobserved, we would still have both quantity l_J and prices δ_{1,l_J} and δ_{2,l_J} as unknowns in the revealed preferences conditions because

Table 3 – Hypothetical data for example 1

	wage $w_1 = w_2$	private leisure $l_1 = l_2$	joint leisure l_J	parental consum. C_P	premium p_2
price regime A	30	30	40	75	15
price regime B	10	10	40	500	5

4.3 Examples

Before we move to our empirical exercise, we offer two examples that help understand the empirical content of Proposition 2 vis-à-vis Proposition 1. Both examples assume for simplicity that p_m and w_K are observed. This is not true in reality ($p_m, w_K \notin D_{\text{LISS}}$) and section 4.4 will explain how we implement the propositions in our *actual* data.

Example 1. This example illustrates how the upper bound on togetherness (5), which is part of the characterization in Proposition 2 but not Proposition 1, matters for the rationalization of leisure and consumption data. Consider the data in table 3 concerning a hypothetical household in price regimes A and B. Assume $h_1^R > h_2^R$ in both regimes. Spousal preferences U_m do not change across the regimes but bargaining power $\mu_m^{(v)}$ may vary between A and B. Wages are higher in price regime A; despite this, the spouses have more private leisure and less consumption in A compared to B.

We will check if the data are consistent with collective rationality in Proposition 1. To do so, we effectively check if the leisure and consumption bundles A and B satisfy the GARP. Definition 1 asserts that if bundle B was affordable at prices A, then bundle A (actually chosen at prices A) must be preferred over B. To obtain the budget set per spouse, we must multiply shadow prices and commodities and add them up across the commodities that enter each spouse's preferences. Then we observe that

$$\delta_{m,l_m}^{(A)} \times (l_m^{(A)} - l_m^{(B)}) + \delta_{m,l_J}^{(A)} \times (l_J^{(A)} - l_J^{(B)}) + \delta_{m,C_P}^{(A)} \times (C_P^{(A)} - C_P^{(B)}) > 0,$$

so, given GARP condition 1 in Definition 1, each spouse m 'revealed prefers' bundle A over B. Proposition 1 disciplines $\delta_{m,l_m} = w_m$ so the above expression is always positive because $\delta_{m,l_m}^{(A)} \times (l_m^{(A)} - l_m^{(B)}) = 30 \times 20 = 600$ and there is no price $\delta_{m,C_P}^{(A)} \leq 1$ for which $\delta_{m,C_P}^{(A)} \times (-425) \leq -600$. As bundle A is revealed preferred over B, then condition 3 in Definition 1 requires that

$$\delta_{m,l_m}^{(B)} \times (l_m^{(B)} - l_m^{(A)}) + \delta_{m,l_J}^{(B)} \times (l_J^{(B)} - l_J^{(A)}) + \delta_{m,C_P}^{(B)} \times (C_P^{(B)} - C_P^{(A)}) \leq 0.$$

the proposition only disciplines the *sum* of shadow prices of joint leisure. One could identify the price-commodity product of joint leisure but it would be hard to separate the two (see also footnote 22).

Table 4 – Hypothetical data for example 2

	wage $w_1 = w_2$	total childcare T_1	total childcare T_2	child consumption C_K	market price w_K	premium p_m
price regime A	20	20	20	75	3	0
price regime B	10	20	10	200	3	0

One can verify that it is impossible to meet the latter condition for both spouses m simultaneously, with $\delta_{1,C_P}^{(B)} + \delta_{2,C_P}^{(B)} = 1$. At least one spouse violates GARP and the data are therefore inconsistent with collective rationality *without* togetherness (Proposition 1). The main reason behind this is that the value of private leisure is too high at prices A ($\delta_{m,l_m}^{(A)} = 30$). Because of this, there are no feasible values for δ_{m,C_P} so that all shadow prices and data jointly satisfy Proposition 1.

The data are, however, consistent with collective rationality *with* togetherness. Statement *(a)* in Proposition 2 tells us that the shadow price δ_{m,l_m} of private leisure does not necessarily equal the hourly wage w_m. In the case when $h_1^R > h_2^R$, spouse 1's wage w_1 pays for two items: the value of own private leisure δ_{1,l_1} *and* the cost p_2 of tightening togetherness (5) as spouse 1 is the one whose work always reduces togetherness. Consequently, the true shadow price of 1's private leisure is lower than the wage (e.g. $\delta_{1,l_1}^{(A)} = w_1^{(A)} - p_2^{(A)} = 15$ at prices A). The interested reader can then verify that there exist consumption prices $\delta_{1,C_P}^{(A)} \geq 12/17$ and $\delta_{2,C_P}^{(B)} \leq 4/17$ for which the previous inequalities and, in fact, all conditions of Proposition 2 are satisfied.[25]

Example 2. This example illustrates how the childcare constraint (6), which is again part of the characterization in Proposition 2 but not Proposition 1, matters for the rationalization of childcare and child expenditure data. Consider again a household in two price regimes, A and B, with data given in table 4. Assume $h_1^R > h_2^R$ in both regimes. Utilities U_m and U_K do not change but utility weights $\mu_m^{(v)}$ and $\mu_K^{(v)}$ may vary between A and B. Wages are again higher at prices A; despite this, spouse 2 supplies more childcare and child consumption is lower in A compared to price regime B.

We will again check if the data are consistent with collective rationality in Proposition 1. To do so, we check whether bundles A and B satisfy the GARP. Invoking the opening condition in Definition 1, we obtain

$$\delta_{K,t_1}^{(A)} \times (T_1^{(A)} - T_1^{(B)}) + \delta_{K,t_2}^{(A)} \times (T_2^{(A)} - T_2^{(B)}) + \delta_{K,C_K}^{(A)} \times (C_K^{(A)} - C_K^{(B)}) > 0,$$

[25]Collective rationality with togetherness further implies $\delta_{1,l_1}^{(A)} = 15$, $\delta_{2,l_2}^{(A)} = 30$, $\delta_{1,l_1}^{(B)} = 5$, $\delta_{2,l_2}^{(B)} = 10$, $\delta_{1,l_J}^{(A)} + \delta_{2,l_J}^{(A)} = 60$, and $\delta_{1,l_J}^{(B)} + \delta_{2,l_J}^{(B)} = 20$. These shadow prices together show that spouse 1 has stronger preferences for consumption while spouse 2 has stronger preferences for leisure.

which is straightforward to check because $\delta_{K,t_m} = w_m$ without togetherness constraints and $\delta_{K,C_K} = 1$ as per Proposition 1. So the household 'revealed prefers' child bundle A over B. Then GARP condition 3 in Definition 1 requires that

$$\delta^{(B)}_{K,t_1} \times (T^{(B)}_1 - T^{(A)}_1) + \delta^{(B)}_{K,t_2} \times (T^{(B)}_2 - T^{(A)}_2) + \delta^{(B)}_{K,C_K} \times (C^{(B)}_K - C^{(A)}_K) \leq 0 \Rightarrow 25 \leq 0,$$

a violation. The reason behind this is the price of parental childcare that is too high in regime A ($w^{(A)}_m = 20$). We thus conclude that the data do not satisfy Proposition 1.

The data are, however, consistent with collective rationality *with* togetherness. Statement *(d)* in Proposition 2 tells us that the shadow price of parental childcare need not be equal to the hourly wage and that the underlying components of childcare may be priced differently. In the case when $h^R_1 > h^R_2$, spouse 2's wage w_2 pays for two items: the value of own private childcare δ_{K,t_2} *and* the cost w_K of tightening childcare constraint (6) as an hour away at work necessitates an hour of market childcare. Consequently, the true shadow price of 2's *private* childcare is lower than the wage (e.g. $\delta^{(A)}_{K,t_2} = w^{(A)}_2 - w^{(A)}_K = 17$ at prices A). Proposition 2 also disciplines the shadow price of *joint* childcare (e.g. $\delta^{(A)}_{K,t_J} = w^{(A)}_1 + w^{(A)}_2 - w^{(A)}_K = 37$). As all childcare shadow prices are known, the previous GARP inequalities only include one unknown: joint childcare t_J (private childcare is simply $T_m - t_J$). The interested reader can then verify that there are values for t_J for which both inequalities, and in fact the GARP, are satisfied (testing). Moreover, the inequalities bound the acceptable t_J and a feasible solution must have $t^{(A)}_J - t^{(B)}_J \geq 18.33$ (recovery or set-identification). The family has more joint childcare at prices A because it is relatively cheaper ($\delta^{(A)}_{K,t_J}/\delta^{(A)}_{K,t_m} = 37/17$) than at prices B ($\delta^{(B)}_{K,t_J}/\delta^{(B)}_{K,t_m} = 17/7$).[26]

The result $t^{(A)}_J - t^{(B)}_J \geq 18.33$ does not say much about the number of hours joint childcare really takes up in a typical week. Fortunately there is more information in our design that helps further discipline t_J. By construction, t_J is always between zero and the minimum individual total childcare, $t_J \in [0, \min\{T_1, T_2\}]$; this upper bound is because joint childcare cannot be larger than *either* parent's own childcare. This suggests that we should bring in these naive *theoretical* bounds into the problem, namely $0 \leq t^{(A)}_J \leq 20$ and $0 \leq t^{(B)}_J \leq 10$. Minimizing and then maximizing $t^{(v)}_J$ subject to the naive bounds and the condition $t^{(A)}_J - t^{(B)}_J \geq 18.33$ (in other words, minimizing and maximizing $t^{(v)}_J$ subject to the naive bounds and the conditions of Proposition 2) yields $18.33 \leq t^{(A)}_J \leq 20$ and $0 \leq t^{(B)}_J \leq 1.67$. These are the applicable bounds for joint childcare in the hypothetical data of table 4. We show all steps in detail in appendix C.3.

[26]Note that our characterization does not use information on the quantity of market childcare T_K because of the exclusion of T_K from children's utility U_K.

4.4 Empirical implementation

While the previous examples illustrate the empirical content of our propositions, testing and identification in our *actual* data requires that we address three final practical issues. First, our data are a short unbalanced panel so we must switch from a setting in which households are observed at multiple prices (how we presented the examples) to a setting in which households may be observed only once. Second, the wage premium p_m and the childcare price w_K are unobserved. Third, the examples were made easy by the fact that the hypothetical data only included two observations. Checking for consistency with multiple observations can quickly become tedious so we must reformulate our testable conditions into a well-defined program. We detail all three issues below.

Cross-sectional setting and heterogeneity. There are two main ways to implement our testable conditions. One is to apply them to each household separately, thus exploit the fact that, given the longitudinal data, we observe each household multiple times. This is the approach we took in the earlier examples and one taken by several other revealed preferences studies (e.g. Cosaert, 2018; Blow et al., 2020). Its main advantage is that it allows for time-invariant heterogeneity across households, equivalent to household-specific preferences. Our data, however, are a short unbalanced panel and many households are only observed once. So the longitudinal approach would not work here.

Another way is to group together households with *similar* demographics and apply our testable conditions to each group separately. We assume effectively that households within a group share the same underlying utility functions U_m and U_K but otherwise have different wages as wages vary in the cross-section.[27] This is then equivalent to observing the same household (multiple households in reality, but they all share the same preferences) at multiple prices (wages). This is the approach taken in Famulari (1995), Cherchye and Vermeulen (2008), and other revealed preferences studies. Its main advantage is that it allows for testing even in settings when households may be observed only once. It also allows for general forms of heterogeneity *across* groups as well as heterogeneity in utility weights μ_m and μ_K *within* groups. This is more general than it first seems: even if utility functions are constant within a group, the shadow prices $\delta_{m,\mathfrak{c}} = \mu_m U_{m,\mathfrak{c}}/\lambda$ and $\delta_{K,\mathfrak{c}} = \mu_K U_{K,\mathfrak{c}}/\lambda$ associated with time inputs of each person (husband, wife, child) can still vary across households in a group because we allow for general inter-household heterogeneity in the utility weights μ_m and μ_K.

We follow the latter cross-sectional approach given the specificities of our data. Based on underlying demographics, we split the 398 households in our sample into 36 groups; we provide the empirical details in the next section. Each group has about 11 observations.

[27] There is no set of universal rules for grouping but some guiding principles can be formulated: (1) the tests must have sufficient power, so prices (wages) must vary within a group, and (2) the determinants of the grouping should reflect important sources of heterogeneity in preferences for time use.

We carry out testing and identification as illustrated in the examples, albeit treating the observations within a group as if they reflect the same household at multiple prices.

Informed grid search for $\mathbf{p_m}$, $\mathbf{w_K}$. Prices p_m and w_K enter the leisure and childcare shadow prices (the δ's) in Proposition 2. These prices are not available in our data, which renders all leisure and childcare shadow prices ultimately unobserved. This presents a major complication because with both prices (the δ's) and quantities (t_J) unobserved, our propositions are overly permissive. Consider example 1 in the previous section. With δ_{1,l_1} unobserved, nearly any consumption price $\delta_{m,C_P} \leq 1$ can rationalize the data and Proposition 2 has no real empirical bite. Similarly, with δ_{K,t_m}, δ_{K,t_J}, and t_J unobserved in example 2, data D_{LISS} will likely always appear consistent with Proposition 2 and any observation would fit its conditions for some childcare shadow prices and quantities. To make progress, we cannot keep p_m and w_K unobserved.

The simplest way to fix this is to use external estimates for p_m and w_K. However, forcing homogeneous prices across households may bias our test to any direction. So we take a slightly different approach and ask: *Are there reasonable prices p_m and w_K for which the data satisfy the conditions of Proposition 2?* This implies a search over 'reasonable' prices p_m and w_K. If the search is successful, that is, if there exist reasonable prices p_m and w_K for which a given group of households satisfies the conditions of Proposition 2, then the data are consistent with collective rationality with togetherness (testing). Based on the admissible prices, we then set identify t_J as illustrated in example 2.

Our search for prices p_m and w_K is *informed*. Pay for work at night, early morning, or late evening is regulated in the Netherlands by collective bargaining agreements. A study of Dutch collective bargaining agreements by Kuiper et al. (2014) reports that for the period of our study the average pay for work at night is 47% on top of a worker's regular wage, while that for work in early morning or late evening is 25% and 42% respectively. While our concept of irregular work does not necessarily reflect nights, early mornings, or late evenings (weekday mornings may be 'irregular' for the concierge), practically a large part of irregular work will likely take place at those times. So we search for a reasonable p_m (specifically, for a markup on one's regular wage) over a grid in the range $[0; 0.5]$, i.e. $0 \leq p_m \leq 0.5 \times m$'s regular wage, which clearly includes Kuiper et al. (2014)'s markups.[28]

We follow a similar approach for the unit price of market childcare. While w_K is about €6/hour in our time frame (Dutch Coalition for Community Schools, 2013), the Dutch government subsidizes childcare with poorer households receiving higher subsidies. Households themselves finance between 4% and 66% of the cost, with an average of about 25% (Kok et al., 2011). So poorer households use childcare effectively for free, while richer households pay a substantial part of its cost. We operationalize these facts by searching

[28]Regular wages are not observed either but they (w_m^R) can be recovered from the condition $h_m^R w_m^R + h_m^I (w_m^R + p_m) = w_m H_m$, with the right-hand side observed in the data.

for w_K (specifically, for a markup on the household's lowest wage) over a grid in $[0; 0.33]$, i.e. $0 \leq w_K \leq 0.33 \times$ household's lowest wage. The lowest point on the grid reflects that childcare is free for some households while the highest point is about 66% of the public unit price (i.e. €4 on average). We make this depend on household wages to reflect that better-paid parents typically opt for higher quality (thus pricier) childcare.

If there exist p_m and w_K in their respective grids that make a group of households fit the conditions of Proposition 2, then this group is consistent with collective rationality with togetherness. Importantly, we search for markups on a group-by-group basis, therefore allowing for markup heterogeneity across groups.[29] Of course there may be multiple markups that rationalize the data. p_m and w_K enter the shadow prices of leisure and childcare (the δ's), so they matter -among other things- for the value of joint over private time use. So for the purpose of recovering the δ's we select the *lowest* markups, thus assigning togetherness the lowest possible value that is consistent with the data.

Mixed integer linear programming. Multiple observations per household group and the inherent logical nature of our problem (Definition 1 includes `if/then` statements, e.g. in the opening and concluding conditions respectively) requires that we recast the problem in terms of a well-defined mathematical program. Our choice is a mixed integer linear programming problem, which is standard in the revealed preferences literature. The mixed integer program involves binary variables that capture the revealed preference relations of the GARP. Given values for p_m and w_K on the grid, the program searches for shadow prices δ and, where applicable, childcare variables t so that all conditions of Proposition 1 or Proposition 2 hold simultaneously. The program retains all the features of testing and set identification precisely as we illustrated in the earlier examples. The program details are inevitably technical, so we relegate them to appendix C.4.

5 Application

5.1 Data

We use the results established in the previous section to test data consistency with our model, quantify the value households assign to togetherness, and obtain bounds on unobserved joint childcare. As we said in section 2, we draw data from the 2009, 2010, and 2012 waves of the LISS (CentERdata, 2012). The survey is administered via multiple online questionnaires so we merge information from various modules to assemble a sample that includes all the variables in D_{LISS}. We apply two main margins of selection in

[29] Allowing markups to vary freely across households makes our testable conditions overly permissive. Fixed markups within groups facilitates their interpretation as policy parameters, although p_m and w_K still vary across households due to variation in underlying wages. The revealed preference characterization with fixed markups is sufficient but not necessary for collective rationality with togetherness.

Table 5 – Sample summary statistics, weekly amounts

	male spouse		female spouse	
	mean	st.d.	mean	st.d.
age	41.08	5.61	38.66	5.41
1[university education or similar]	0.36	0.48	0.38	0.49
private leisure	16.49	12.32	14.16	11.44
childcare	13.30	9.82	21.67	14.36
regular work hours	32.86	9.57	20.47	9.24
1[irregular work hours > 0]	0.55	0.50	0.37	0.48
irregular work hours	8.26	9.13	3.13	4.87
observed hourly wage (€)	12.37	5.02	11.64	4.40
	household			
	mean		st.d.	
age youngest child	5.77		3.59	
number of children	2.16		0.83	
1[married]	0.83		0.37	
joint leisure	9.45		7.56	
minimum of parents' childcare	11.44		8.51	
parental consumption (€)	592.82		223.12	
child consumption (€)	75.88		52.75	
	sample			
# of household groups	36			
# of households in a group	avg. = 11.86 (st.d. = 3.07)			
household/year observations	398			

Notes: The table reports summary statistics (mean and standard deviation) in our estimation sample. All time and monetary amounts are weekly. Monetary amounts are expressed in 2012 prices; the monetary series are deflated using the Consumer Price Index available by Statistics Netherlands.

the baseline: both spouses must participate in the labor market, and the couple must have at least one child up to 12 years old. Together with other standard data requirements (e.g. no missing demographics), these two criteria result in a baseline sample of 398 household×year observations. We observe 148 households only once, 59 households twice, and 44 household thrice (i.e. a total of 251 unique households). Appendix A.1 details the construction of our sample and defines what each variable precisely measures.

Table 5 presents summary statistics. On average the male spouse is 41 years old and has a 36% probability of being university educated. The female spouse is almost 39 years old and is 38% likely to be university educated. He has about 16 hours of private leisure per week (she has 2 hours less) and spends about 13 hours on childcare (she spends almost 22); these numbers echo our introductory table 1. We calculate the minimum of parents' reported childcare, i.e. $\min\{T_1, T_2\}$, and we find it equal to about 11 hours per week.

The male's hourly wage for regular and irregular work combined after taxes are paid is about €12.4 while the female's is €11.6. While certain income components in the

Netherlands are subject to joint taxation between spouses, labor income is taxed separately (Dutch Belastingdienst, 2020). So while we do not explicitly model labor income taxation, our characterization is still valid with respect to *after-tax* wages. A household has on average 2.16 children (19% of households have 1 child, 54% have two children) and spends €593 on parental consumption and €76 on child consumption per week.

We do not separately observe the hours of regular and irregular work but we do observe total hours of work per spouse and frequency indicators for the incidence of irregular work (appendix A.1). These indicators capture whether work is *often*, *sometimes*, or *never* irregular and they form the basis for our motivating figure 4. We translate this information by assigning to irregular work 50% of a spouse's total market hours H_m in case irregular work occurs *often*, and 25% in case it occurs *sometimes*. All remaining market hours are regular. Based on this translation, the male spouse records an average of 33 hours of regular work while the female spouse records approximately 20. In terms of irregular work, he works on average for about 8 hours while she works for a bit more than 3. Irregular work is prevalent; consistent with figure 4, 55% of men and 37% of women in our sample report at least some irregular work. We think of the numbers above as a natural benchmark for the qualitative frequency indicators. We check, however, the robustness of our results with respect to a different quantification, namely 66% (*often*) - 33% (*sometimes*). We present these results in section 5.4. None of our baseline results change meaningfully with the quantification of the qualitative frequency indicators.

Given the short unbalanced panel, we must implement our testable conditions cross-sectionally. Based on observed demographics, we pool similar households together and form 36 household groups, each one comprising about 11 observations on average. We assign observations to groups given values for: calendar year (for 2009, 2010, 2012), age of youngest child (for 0-5, 6-12), average age of parents (for 25-34, 35-44, 45-60), and education of parents (2 values; cutoff at the median of the product of the parents' schooling years within a calendar-time, age-of-child, and age-of-parents cell; this corresponds to intermediate vocational education, similar to junior college in the US). We then carry out testing and identification exactly as illustrated in the examples of section 4.3, treating the observations within a group as if they reflect the same household at multiple prices.

5.2 Testing and recovery

Testing. Our empirical analysis starts with a formal comparison of the different collective models with private and joint time use. The first model in the comparison reflects our notion of collective rationality in Definition 2 (CR in short). This model makes a formal distinction between private and joint times, but the gains from joint time equal the sum of gains from private times in equilibrium. There is no forgone flexibility or forgone specialization. The most general model in this comparison reflects our notion of collec-

Table 6 – Pass rates of collective rationality in various models

model	pass rate (max. 1.0)	number of groups (max. 36)	number of observations (max. 398)
CR	0.25	9	87
T–CR	0.67	24	250
A–CR	0.06	2	19

Notes: The table reports the pass rates of collective rationality (CR), collective rationality with togetherness (T–CR), and collective rationality with aggregate times only (A–CR), as well as the associated number of household groups and household observations that are rational in each case.

tive rationality with togetherness in Definition 3 (T–CR in short). This model makes an explicit distinction between private and joint times, and allows that the marginal gains from joint time exceed the sum of gains from the spouses' private times in equilibrium. According to this model, household members weigh these gains against the costs of forgone flexibility and specialization. The test of T–CR is the only one that carries out a grid search over p_m and w_K, in search for a wage premium and a price of market childcare that allow us to rationalize the underlying group of households. Finally, the most restrictive specification that we consider is a collective model that makes no distinction between private and joint time use. Welfare thus depends on individual total times only. We refer to this model as collective rationality with aggregate times, A–CR.

Table 6 reports the pass rates of these three specifications. The pass rate captures the proportion of groups whose behavior is fully consistent with the conditions under consideration. We apply this 'sharp' test to each group separately, thereby accounting for general forms of preference heterogeneity between groups. A–CR describes the behavior of only 6% of groups, or 19 households in total. CR describes the behavior of 25% of groups, or 87 households in total. Finally, T–CR is consistent with 67% of groups, that is 24 groups or 250 households in total. The large difference in favor of T–CR already indicates that joint time indeed generates gains over and above the sum of private times as household members incur the costs of forgone flexibility and forgone specialization.[30]

Pass rates provide empirical support for our novel notion of T–CR. Next, we compute the discriminatory power and predictive success of our models, focusing on CR and T–CR. *Power* is one minus the pass rate of random data. It therefore addresses the question whether the models are sufficiently strong to reject consistency for random, simulated observations. The stronger the discriminatory power of a model, the more it should reject consistency of these random data. *Predictive success*, introduced by Selten (1991) and axiomatized by Beatty and Crawford (2011), summarizes pass rates and power. It

[30]In all cases we implement a scenario in which either C_P is shared equally or parents have the same individual valuations for it. Without this restriction, C_P is a public good with unobserved Lindahl prices. For that case, we find pass rates of 31%, 31%, and 75% for A–CR, CR, and T–CR respectively.

Table 7 – Power and predictive success of collective rationality in various models

model	discriminatory power	predictive success
CR	0.82 (0.20)	0.07 (0.44)
T–CR	0.62 (0.24)	0.29 (0.38)

Notes: The table reports the average and standard deviation (in parentheses) of power and predictive success of collective rationality (CR) and collective rationality with togetherness (T–CR) across groups. See text for definitions.

is the sum of pass rates and power minus one. Predictive success therefore captures the degree to which a model describes the observed data better than the random data. It is expressed on a continuum from -1 to 1. A predictive success score of 1 indicates that 100% of the actual data pass the model while 100% of the random data violate it. This is the best possible scenario. Vice versa, a predictive success score of -1 indicates that 100% of the actual data violate the model while 100% of the random data pass it. Positive predictive success scores are desirable.

To compute power, we follow a procedure outlined by Bronars (1987), which is based on Becker (1962)'s notion of irrational behavior. We construct artificial data by drawing random time and consumption bundles given a household's income and wages. We test consistency of our groups with CR and T–CR, while replacing the groups' actual data with the newly generated random data. We repeat this procedure 100 times per group which gives us 3,600 separate tests of rationality; all details of testing in the random data are exactly as testing in the actual data. The results are in table 7. CR rejects consistency, on average, for 82% of groups. Although this may seem high, note that CR describes the actual data only 7% better than the random data. We then compute discriminatory power in T–CR. Our model rejects consistency for 62% of groups. Although this index is only slightly better than half, T–CR describes the actual data 29% better than the random data. This provides empirical support for our model.[31]

While table 7 presents *mean* power and predictive success among our 36 groups, we can also shed light on the underlying *distribution* of power and predictive success as we compute both measures for each group separately. The results in figure 6 confirm empirical support for T–CR. The distribution of predictive success without togetherness (dashed line) clearly has much more mass between -1 and 0, and T–CR generally dominates CR.

Recovery of joint childcare. We found that T–CR has better goodness-of-fit without

[31]There is no straightforward statistical framework to compare average predictive success scores because the result for each group depends on a complex interaction of multiple inequalities, corresponding to different households. Yet, a predictive success gap of 0.22 is large in the literature, see e.g. Cosaert (2018), Demuynck and Seel (2018), and Blow et al. (2020). Demuynck and Seel (2018) find that consumption models with endogenous consideration sets outperform the neo-classical model of utility maximization; the corresponding predictive success gap in their paper is about 0.06.

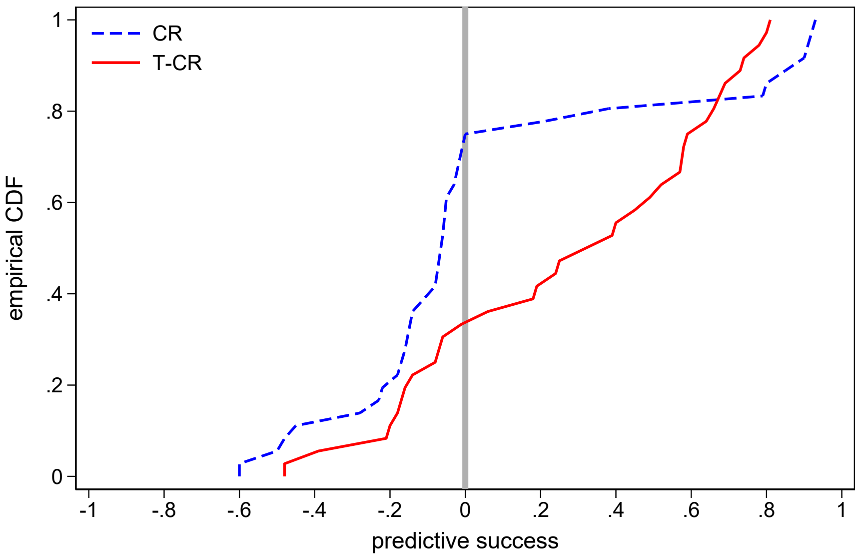


Figure 6 – Distribution of predictive success in various models

Notes: The figure plots the distribution of predictive success for collective rationality (CR) and collective rationality with togetherness (T–CR) among 36 groups of households. See main text for definitions.

being overly permissive. Next, we use T–CR to separate private from joint childcare when only individual total childcare is available. To facilitate the exposition of the results, our object of interest is the proportion $t_J/\min\{T_1, T_2\}$ that reflects how much of the naive maximum amount of joint childcare is supplied jointly. We apply the procedure described in section 4.3 to recover bounds on the average $t_J/\min\{T_1, T_2\}$ within each group that is consistent with collective rationality with togetherness.[32]

Table 8 presents the distribution of bounds for all 24 groups consistent with T–CR. On average, joint childcare amounts to at least 2% and at most 99% of its naive maximum. As expected, there is some heterogeneity across groups. The maximum lower bound at 10% reflects there is a group in which households supply at least 10% of childcare jointly; the minimum upper bound at 90% suggests that households in another group supply at most 90% of childcare jointly. As our groups differ with respect to demographics (age of children, age and education of parents) we can in principle investigate whether our bounds relate systematically to them. This is, however, hard to do in practice as, given we identify one set of bounds for each group of households, we only have 24 observations to assess a statistical hypothesis. Moreover, our bounds are admittedly not sharp.

Our bounds are not sharp because we have kept our framework general, refraining from imposing additional restrictions. However, we explore one set of restrictions in appendix C.3.2 and we see that the bounds can be sharpened substantially. Specifically, we find that parents spend on average between 3.8 and 11.9 hours per week on joint childcare, representing up to 92% of total childcare of one parent. We report the sharper

[32]Specifically, we minimize/maximize $\alpha = (1/V)\sum_v t_J^{(v)}/\min\{T_1^{(v)}, T_2^{(v)}\}$ across households v in a group, subject to the constraints in Proposition 2. See appendix C.3 for details.

Table 8 – Distribution of bounds on mean proportion of joint childcare

mean $(t_J/\min\{T_1,T_2\})$ within group	estimated lower bound	estimated upper bound
Distribution across groups:		
min	0	0.90
mean	0.02	0.99
max	0.10	1

Notes: The table reports the distribution *across groups* of the bounds on the mean *within-group* proportion of joint childcare, i.e. mean $(t_J/\min\{T_1,T_2\})$ within group. The results are over 24 groups of households consistent with collective rationality with togetherness (T–CR).

bounds in table C.1 and visualize them in figure C.1. Our sharper bounds are consistent with the information from the LISS time diary described in appendix A.2.

5.3 The value of togetherness

We have the necessary ingredients to identify the marginal value of joint leisure relative to private leisure $(\delta^{(v)}_{1,l_J}+\delta^{(v)}_{2,l_J}-\delta^{(v)}_{1,l_1}-\delta^{(v)}_{2,l_2})$ and the marginal value of joint childcare relative to private childcare $(\delta^{(v)}_{K,t_J}-\delta^{(v)}_{K,t_1}-\delta^{(v)}_{K,t_2})$. These values indicate how much the household is willing to pay for an hour of joint time *relative* to the price it pays for an hour of private time by each spouse. The T–CR model and its goodness-of-fit allow us to pin down lower bounds on both values. First, the model tells us that in equilibrium the value of joint leisure equals the premium for irregular work p_m (reflecting forgone flexibility), while the value of joint childcare equals the premium for irregular work p_m and the price of market childcare w_K (reflecting forgone flexibility and forgone specialization).[33,34] Second, the goodness-of-fit in table 6 indicates which households behave consistently with T–CR and which not. We found that 250 households are rationalizable with togetherness.

Figure 7 presents the results. The dotted red line presents the distribution of the underlying value of joint leisure across households. The mean value is €1.22, indicating that on average households are willing to pay at least €1.22 for an hour of joint leisure *over* the cost of an hour of private leisure by *each* spouse. This corresponds to 10% of the average wage in our sample. In addition, 5% of households are willing to pay at least €4/hour to replace an hour of private leisure by each spouse with an hour of joint leisure.

The solid blue line represents the distribution of the value of joint childcare across households. Values lie between €0 and €10 per hour of joint time. The mean value is €2.08, indicating that on average households are willing to pay at least €2.08 for an hour of joint childcare *over* the cost of an hour of private childcare by *each* spouse. This

[33] Recall that the marginal value of joint over private childcare may reflect stronger preferences for this time use or a higher productivity of joint time; the two interpretations are indistinguishable.

[34] We obtain lower bounds because we select from our original grid search the combination that yields the *smallest* value for togetherness $\sum p_m + w_K$ within a given group.

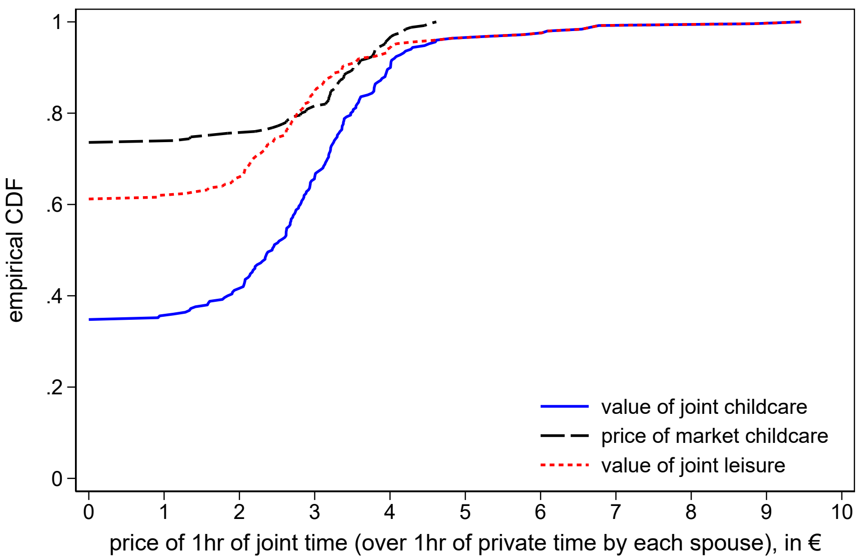


Figure 7 – Distribution of value of togetherness

Notes: The figure plots the distribution of the value of togetherness based on forgone specialization and forgone flexibility among 250 households consistent with collective rationality with togetherness (T–CR).

corresponds to almost 17% of the average wage in our sample. Digging a bit deeper into the distribution, we see that 5% of households are willing to pay at least €4.5/hour. By contrast, for about 33% of households the willingness to pay for joint childcare is 0. This corresponds to households that are also consistent with CR.

These results provide us with a decomposition of the value of joint childcare in terms of the price of irregular work (forgone flexibility) and the price of market childcare (forgone specialization). The former also corresponds to the marginal value of joint leisure. Values for t_J equal to the price of market childcare indicate forgone specialization but no forgone flexibility. Values for t_J equal to the value of joint leisure alone indicate forgone flexibility but no forgone specialization. The dashed black, resp. dotted red, lines in figure 7 plot the distribution of these components. The value of joint childcare frequently comprises both items and we cannot exclude either as a possible cost of togetherness.

5.4 Robustness of empirical results

We repeat our analysis in sections 5.2 and 5.3 under six independent scenarios in order to investigate the robustness of our findings. These scenarios take into account commuting, the possibility that spouses have the discretion to determine the timing of part of irregular work, an alternative quantification of the frequency indicators for irregular work, measurement error in childcare, and the overall quality of our time use data.

The first scenario augments market work with commuting. Similar to pure market work, commuting implies that individuals spend time away from their spouse. We thus adapt regular and irregular hours, wages, and the upper bound on joint time accordingly.

The second and third robustness checks relax the assumption that individuals have no control over their irregular work schedules and that irregular work never overlaps. This exercise introduces the idea that individuals may have limited control over the timing of their irregular work and, as such, spouses may synchronize part of the irregular hours between them or synchronize them with their partner's regular work. Suppose that a fraction γ of irregular hours are flexible. Then condition (5) becomes $l_J + t_J \leq \mathcal{T}_m - \max\left\{h_1^R + \gamma h_1^I, h_2^R + \gamma h_2^I\right\} - (1-\gamma)h_1^I - (1-\gamma)h_2^I$; $\gamma = 0$ is our original no overlap set-up whereas $\gamma = 1$ assumes perfect overlap between irregular hours. The updated cost of forgone flexibility is $\tau_J = \lambda p_2/(1-\gamma)$. We explore several values for γ but here we only report results for $\gamma \leq 0.1$ and $\gamma \leq 0.2$.[35]

The fourth extension revisits the frequency indicators for the incidence of irregular work. Subjects in the LISS work irregular hours *often*, *sometimes*, or *never*. We translated this information by assigning to irregular hours 50% (resp. 25%) of a spouse's total market hours in case she reported irregular work to occur *often* (resp. *sometimes*). We test the sensitivity of this translation by considering 66% for *often* and 33% for *sometimes*.

Our fifth check addresses measurement error in individual total childcare T_m. We treat *true* $\widetilde{T}_m$ as unobserved and we let $\widetilde{T}_m \in [0.99 \times T_m; 1.01 \times T_m]$. This admits a 1% deviation from T_m in both directions. The conditions of Proposition 2 are linear in the new variable $\widetilde{T}_m$ (appendix C.4), which makes the procedure easy to implement.

Our final check assesses the robustness of our results with respect to the quality of our time use data. The LISS includes an exhaustive list of questions aimed to elicit the respondents' time use over all 168 hours in a week. So the extent to which the weekly time budgets add up to 168 hours is indicative of the quality of the data or the prevalence of errors. About 62% of households have their time budgets add up to about 168 hours (we allow for rounding errors of ± 4 hours). While in our baseline estimation we normalized time use in the remaining households to add up to 168, this robustness check uses only the 62% of households whose time budgets add up to 168 hours by default.

Table 9 reproduces testing (tables 6 and 7) and estimation of the value of togetherness (figure 7) for all scenarios. Four remarks are in order. First, the pass rate, power, and predictive success of collective rationality with togetherness (pred. succ. T–CR) are similar across our baseline and all robustness checks. Second, predictive success of collective rationality with togetherness consistently exceeds predictive success of collective rationality without togetherness (pred. succ. CR). Third, the mean value of one unit of joint leisure (over and above one unit of private leisure per spouse) lies between €1/hour and €1.6/hour, with our baseline at €1.22/hour in the middle of that range. Fourth, the value of joint childcare lies between about €1.7/hour and €2.3/hour, with our baseline again in the middle of that range. Overall, all results are similar to our baseline.

[35] We try values for γ up to 0.5, and we do this per household group, such that the resulting pass rate is as high as possible.

Table 9 – Robustness: fit and value of togetherness

		Robustness scenarios					
		(1)	(2)	(3)	(4)	(5)	(6)
	Base-line	Com-muting	10% control	20% control	66–33–0	Meas. error	Time budget
Testing (pass rates, power, predictive success)							
pass CR	0.25	0.25	0.25	0.25	0.25	0.36	0.25
pass T–CR	0.67	0.64	0.69	0.69	0.67	0.67	0.67
pass # groups	24	23	25	25	24	24	8
pass # HHs	250	242	263	263	250	250	103
power CR	0.82	0.83	0.82	0.82	0.82	0.78	0.89
power T–CR	0.62	0.62	0.61	0.61	0.62	0.59	0.68
pred. succ. CR	0.07	0.08	0.07	0.07	0.07	0.14	0.14
pred. succ. T–CR	0.29	0.26	0.31	0.30	0.29	0.26	0.35
Value of joint leisure $\sum_m \delta_{m,l_J} - \sum_m \delta_{m,l_m}$							
first quartile	0	0	0	0	0	0	0
median	0	0	0	0	0	0	1.81
mean	1.22	0.98	1.51	1.39	1.21	1.19	1.59
third quartile	2.60	2.00	2.73	2.68	2.53	2.44	2.93
max	9.46	13.73	21.59	21.59	9.46	10.75	5.90
Value of joint childcare $\delta_{K,t_J} - \sum_m \delta_{K,t_m}$							
first quartile	0	0	0	0	0	0	0
median	2.44	2.39	2.57	2.51	2.44	1.31	2.54
mean	2.08	1.99	2.33	2.21	2.07	1.67	2.08
third quartile	3.30	3.00	3.39	3.38	3.30	3.04	3.20
max	9.46	13.73	21.59	21.59	9.46	10.75	5.90

Notes: The table reports pass rates, power, and predictive success of collective rationality (CR) and collective rationality with togetherness (T–CR), and the distribution of the value of joint leisure and joint childcare among households consistent with T–CR. We introduce the following scenarios: (1) market work augmented with commuting time, (2/3) individuals control up to 10% (resp. 20%) of their irregular work, (4) *often* and *sometimes* assign 66% and 33% of market hours to irregular work, (5) reported total childcare may deviate from real total childcare by 1% (measurement error), (6) subset of households for whom the weekly time budgets add up to 168±4 hours. This subsample has fewer observations than our baseline so we introduce a smaller number of household groups (12) with assignment done on the basis of calendar year (for 2009, 2010, 2012), child age (for 0-5, 6-12), and education (2 values, cutoff at intermediate vocational education).

6 Discussion and extensions

We now highlight the implications of our model for the gender wage gap and for inequality. We also discuss how different fertility circumstances, assortative matching at marriage, and general intertemporal considerations may be approached in our setting.

6.1 Policy topics

Work schedules and gender wage gap. Togetherness requires that the spouses synchronize their work so as to be physically together at the same time. So it is interesting to study the interaction between the value of togetherness and work schedules. We shut down forgone specialization (as if households have free access to external childcare; this

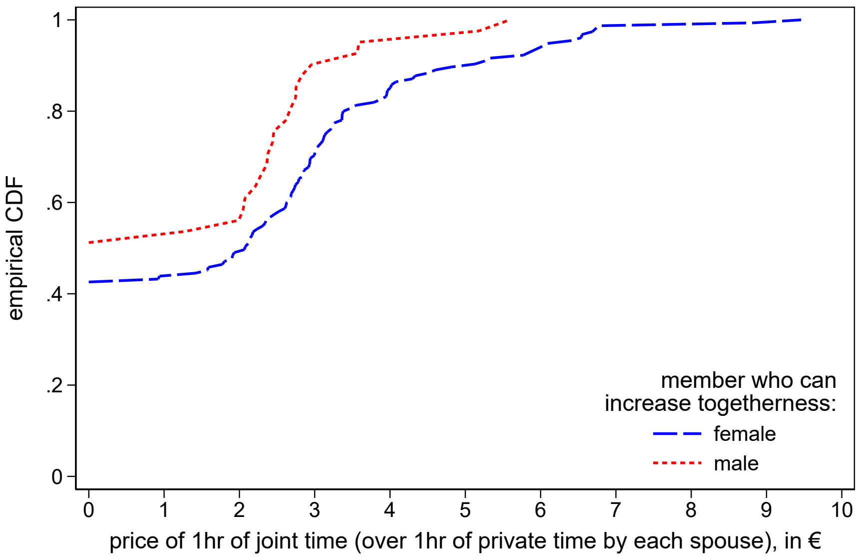


Figure 8 – Distribution of value of togetherness (forgone flexibility only)

Notes: The figure plots the distribution of the value of togetherness based on forgone flexibility among 196 households consistent with collective rationality with togetherness (T–CR) and $w_K = 0$.

does not qualitatively affect our subsequent results) and restrict attention to forgone flexibility (5).[36] Here we put forward the idea that one household member may forgo a wage premium in order to increase the amount of joint time in the household.

The individual who supplies less labor during regular hours has the opportunity to increase togetherness at cost p_m instead of $w_m + p_m$. The reason is simple: she can increase togetherness without working fewer hours in total; she can simply replace one hour of irregular with one hour of regular work. The latter does not put pressure on togetherness given the overlap in partners' regular hours. Her husband cannot do this as any increase in his regular hours immediately reduces maximum potential togetherness. About 80% of rational households in our sample have $h_1^R > h_2^R$, assigning member 1 to men and member 2 to women. So women are more frequently in the position where it is relatively cheaper to increase togetherness at cost p_m only. More women than men then have an incentive to drop irregular work, thus possibly compress their hourly earnings.

In addition, a positive correlation between the value of togetherness and the likelihood of $h_1^R > h_2^R$ may compress female irregular hours even further. Figure 8 plots the value of joint time for households in which women may forgo earnings ($h_1^R > h_2^R$, dashed blue line), in this case p_2, and for households in which men may forgo earnings ($h_1^R < h_2^R$, dotted red line), in this case p_1. The value of togetherness is higher in households where the female is the member with the less costly opportunity to increase joint time. In particular, the added value of togetherness is €0/hour for the median household with $h_1^R < h_2^R$ whereas it is slightly more than €2/hour for the median household with $h_1^R > h_2^R$.[37]

[36]Note that 53% of groups in our sample (196 households) are consistent with T–CR and $w_K = 0$.
[37]This is so even if the value of togetherness (monetized by the value of forgone flexibility here) should

In summary, our model implies that whoever supplies the least regular hours is the person who more easily also forgoes labor market flexibility. It is simply less costly for them to replace an hour of irregular with an hour of regular work, thus forgo premium p. A historical gender wage gap may have women as the secondary earner, thus working fewer regular hours. Our model then says that in such cases women will likely forgo the irregular wage premium, hence forgo earnings, thereby *reinforcing* the gender wage gap further. Moreover, in those cases (80%) the value of togetherness is higher than in the opposite cases, and this is unrelated to any pre-existing wage gap. Therefore, there is an even stronger incentive for alignment of work schedules, making women the ultimate person who sacrifices labor market flexibility and earnings. This is consistent with Goldin (2014), Olivetti and Petrongolo (2016) and Cubas et al. (2019), who relate the timing of female work to the gender wage gap. Complementary to them, we argue that the timing of women's work is partly driven by a demand for togetherness in the household.

This simple insight has important implications for gender gap policy and intra-household inequality. Suppose one wants to eliminate the gender wage gap without taking the joint timing of market work into consideration. Policy would mandate, as is often done, equality in the wages offered to men and women, even if wages would typically still differ across education groups, occupations, or job features. Time flexibility on the job is one such feature. So even if policy instructs that $w_1 = w_2 = w$ and $p_1 = p_2 = p$, this will not put an end to the gender gap if women still forgo the premium, e.g. if women do not earn a bonus because they are not as flexible in the timing of work. So while equalizing hourly rates between genders goes a long way towards pay parity, it does not eliminate the pay gap unless both genders have similar schedules at work. The core of our policy argument holds even if irregular work doesn't pay a premium as in our model. This argument is: with preferences for togetherness, one household member may have to choose working times in such a way such that those times are synchronized with his/her partner's. Togetherness requires synchronization. Currently, it is women who work fewer overall hours so it is women who can restrict the timing of their market work without decreasing their total hours. If there is a benefit to working unrestricted times (e.g. prospect of future promotion), then women may lose this benefit.[38]

in principle be lower when women work fewer regular hours, as p_2 (the value of togetherness in this case) is generally lower than p_1 (the value of togetherness in the opposite case). p_m is a markup over w_m and women's wages are typically lower than men's. In the data, however, households with $h_1^R > h_2^R$ are more likely to face a *binding* upper bound (5), thus requiring large values for p_2 to rationalize the data.

[38]Given a preexisting gap in regular hours, gender differences in work schedules arise in our model endogenously, i.e. as a choice, due to demand for togetherness. One may argue that policy should do nothing to eliminate a timing-of-work related pay gap that is the outcome of people's choices. This line of argument is, however, incomplete. Togetherness improves marital stability and child development with welfare-improving consequences for individuals and the economy. If women disproportionately pay the cost of this welfare, then there is scope for policy to compensate them for this. Inflexibility at work may also limit the careers women pursue or the degrees they study for (e.g. Bronson, 2014; Goldin, 2020). While we will not delve more into this discussion here, the core of our argument is that gender pay disparities cannot be remedied without a more holistic approach that also includes the timing of work.

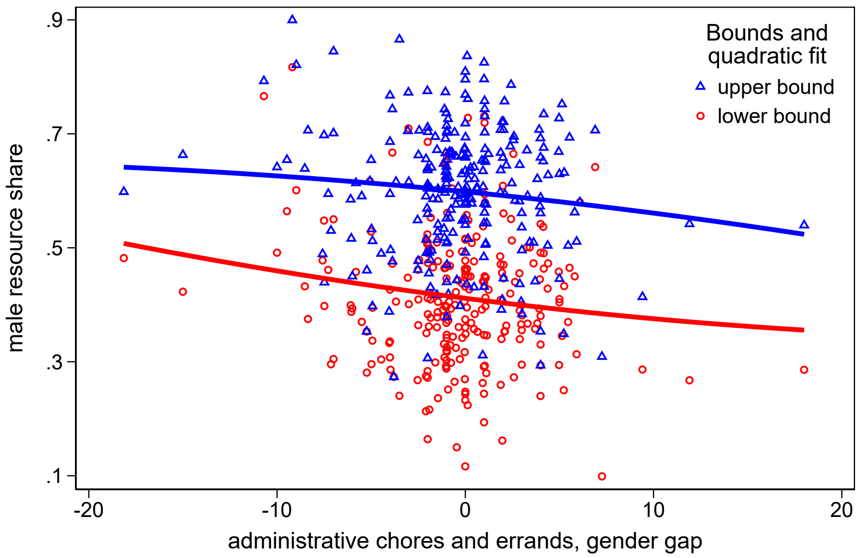


Figure 9 – Male resource shares against gender gap in administrative chores

Notes: The figure plots upper and lower bounds on $\eta_1 / \sum_m \eta_m$ among 250 households consistent with T–CR.

That women forgo flexibility and possibly sacrifice the premium of irregular work or other benefits has additional adverse consequences. A larger wage gap depresses women's human capital, reinforcing the wage disparities further. It also contributes to power disparities in the family, opening up a vicious circle of inequality in which the wage gap not only feeds into lower future female wages but also into a less equal distribution of goods in the family. Fortunately there is a limit to this, warranted by the plethora of public goods and togetherness itself.

Inequality and resource shares. The implications of our model for intra-household inequality differ markedly, at least as far as leisure is concerned, from those of the classical collective model. The standard model suggests that an increase in one spouse's wage may improve own bargaining power, thus increase labor supply and decrease leisure *of the partner*. In our model, however, joint leisure will increase with own bargaining power and partially offset the decrease in the partner's total leisure. This implies that resource shares in our setting should generally be less unequal between spouses than in a setting in which goods are private and togetherness does not exist. In other words, even if the gender wage gap opens up a vicious circle of inequality, inequality in the household should generally remain low because of togetherness and other public goods.

To shed light on intra-household inequality in our data, we estimate bounds on each spouse's resource share in the household, also known in the literature as *sharing rule*. The sharing rule is a sufficient statistic for the decentralized representation of the household, namely one in which each spouse acts rationally given their individual share of household resources. Following Chiappori and Meghir (2014) who also work with public goods, we define spouse m's resource share as $\eta_m = (\delta_{m,l_m} \times l_m) + (\delta_{m,l_J} \times l_J) + (\delta_{m,C_P} \times C_P)$. This 'personalized' budget consists of spouse m's valuation of private leisure multiplied by the amount of private leisure, her willingness to pay for joint leisure multiplied by the amount of joint leisure, and her share of household consumption. We estimate lower and upper bounds on the resource shares, specifically on $\eta_m/(\eta_1 + \eta_2)$, following Cherchye et al. (2011) and a sequence of papers thereafter; see also appendix C.4.

We estimate the average male resource share (half the sum of lower and upper bounds) at about 0.5, i.e. an almost equal sharing between spouses. This is consistent with our intuition that togetherness and public goods (C_P is a large part of resources) offset inequalities in the household. We do not find many significant correlations between the bounds and potential distribution factors, but we do find that higher *female* education reduces men's share, as also does a larger *male* share of administrative chores (e.g. errands). Figure 9 plots the bounds against the male share of administrative chores, along with the best quadratic fit.

6.2 Discussion

Children. We have presented our model *as if* households are homogeneous with respect to the number of children. In reality, heterogeneity in this number may matter for parental time use. For example, the very concept of togetherness likely differs between a 1-child family and a family of 4 children (the *age* of the youngest child may also matter; we account for this empirically in the way we group households).

Our model can be easily adapted to different children numbers. We can introduce utilities U_{K_1}, U_{K_2}, ... U_{K_N} and time budgets $\mathcal{T}_{K_1}$, $\mathcal{T}_{K_2}$, ... $\mathcal{T}_{K_N}$ for the 1$^{\text{st}}$, 2$^{\text{nd}}$, ... N^{th} child respectively and derive testable conditions for each child separately. The implementation of those conditions, however, requires information on childcare devoted to each child and on how this overlaps across children. The latter is unlikely to be available even in time diaries so our baseline formulation serves as a tractable approximation to this more detailed structure. Our formulation assumes that U_K and $\mathcal{T}_K$ capture utility and time of all children combined, as if parents 'bundle' the children together (the typical approach in applied work, e.g. Cherchye et al., 2012). An alternative but somewhat related interpretation is that U_K and $\mathcal{T}_K$ capture utility and time of the *youngest* child because the parents' time use is likely mostly responsive to the needs of that child.

Our model also abstracts from a fertility *choice* on the intensive (number of children, spacing) or extensive margin (to have children or not). One may argue that the pregnancy decision is more central to togetherness than how much one works in the market or what type of work they do. At least the extensive margin is clearly important. Without children, one element of togetherness (joint childcare) is no longer available; it is not clear, however, how this may affect joint leisure. If childlessness reflects a low togetherness 'type', then l_J should be low and the trade-off between regular and irregular work irrelevant. But if it reflects that a couple is so engaged in work that it lacks togetherness, e.g. work constraint (5) binds at extremely low levels for l_J, then this may have prevented the couple from having children in the first place. This suggests that time use should be studied alongside fertility, an interesting avenue of future research especially given the scarcity of collective models of fertility (a notable exception is Eckstein et al., 2019).

For now we take an empirical approach to these issues. We repeat our application twice: first, over sample families with 1-2 children and, separately, with 3+ children; second, over a sample of childless couples that fulfill all other selection criteria (see appendix A.2). Neither addresses the joint choice of fertility and time use but they shed light on how different fertilities affect the rationalization of the data or the value of togetherness. Comparing pass rates in families with 1-2 and 3+ children enables us to check if the model unrealistically 'bundles' several children together. The application to childless couples is informative about the relevance of togetherness in the absence of children.

Appendix table D.1 shows the results (columns 1-3). In both 1-2 and 3+ children samples, predictive success is larger in T–CR than in CR and the difference is more pronounced in the first sample (success rate 0.37 in T–CR vs. 0.02 in CR). This suggests that 'bundling' multiple children together may be overly simplistic and the model offers a better representation of families without too many children. The value of l_J is similar in the two samples but the value of t_J is substantially lower in the latter, indicating that private and joint childcares may not be too different with 3+ children. Without children, predictive success in T–CR and CR is similar (about 0.3) but much larger than in the model without joint leisure (-0.04 in A–CR; not shown in the table). This suggests that joint leisure remains an important component of household time use even though the only relevant togetherness constraint in childless couples, work constraint (5), seems to not be binding (if it were, the difference between T–CR and CR would be pronounced).[39]

Marriage. Assortative matching, or the choice whom to marry, is another relevant margin from which the model abstracts. One may argue that persons with strong preferences for togetherness may match positively on regular occupations because this permits large amounts of togetherness. Ignoring this primary margin may bias the true value of jointness. For this couple $p_m = 0$ (no irregular work) and we would likely underestimate the value of togetherness because we use the premium to pin down this value. By contrast, people who match on irregular occupations may in reality be low togetherness 'types'. We will mistakenly think that the second couple has little togetherness because of its irregular schedule, i.e. work constraint (5) binds at low levels of $l_J + t_J$, and place too much emphasis on this, thus overstating the true value of jointness.

[39] A note of caution is due here. Pass rates and predictive success are good measures for comparing different models on the same data. In this and the following exercise, we compare the same model in different datasets. Different datasets have different amounts of wage variation, different group sizes, etc. Especially with small sample sizes, this can make comparison of the performance measures difficult.

We take again an empirical approach and repeat our application among couples positively matched. This does not address the joint choice of marriage and time use but it sheds light on how matching may affect the value of togetherness. We first explore matching on occupational sector. Interestingly, there is little positive matching there: appendix figure D.1 shows that only 13% of couples share the same occupation. This small number of positively matched does not allow us to repeat our exercise on those matched on occupation (an obvious caveat here is that we cannot measure *negative* matching on occupation), so we turn to education where a clearer pattern emerges. The partners have the same level of schooling in about 42% of couples (figure D.1); this almost equal split of the sample allows us to repeat our exercise on those positively versus non-positively matched on education. Table D.1 (columns 4-5) shows that the results do not differ much between the two samples or from the baseline. The value of l_J is somewhat lower in the positively matched (which is counterfactual and likely reflects our understatement of togetherness there). However, the difference in predictive success between T–CR and CR is substantially more pronounced among those with similar levels of schooling, suggesting that togetherness is more relevant among the *positively* matched.

Lifecycle. Our model is static and, as such, it ignores the dynamics of household decision making or how choices made today affect future outcomes. The household problem in this paper should be seen as the static component in a broader intertemporal problem that includes education/occupation choices, marriage and pregnancy, wealth accumulation, or intertemporal substitution in leisure. Optimality in the intertemporal sense requires optimality in the static problem, *given* education, occupation, partner, fertility, wealth, available time etc. Moreover, we do not strictly need to group households in the static problem with respect to wealth or available time. Non-labor income Y or time endowments $\mathcal{T}_m$ and $\mathcal{T}_K$ do not enter the testable conditions in Proposition 2; Y and $\mathcal{T}$ affect the *level* of utility, which is irrelevant for what we do. Of course, while the stock of wealth or time may not matter, a *choice* over savings or available leisure next period (for example, through choosing how much annual leave to redeem today vs. retain for tomorrow) would likely bias our results insofar as couples make such choices in reality but we neglect them in the model. These are all potentially important issues which we leave for future research.

7 Conclusion

Spending time together with a partner is a major source of gain from marriage. However, models of household time use typically treat time as a commodity consumed by household members privately rather than also jointly. As a result, we know very little about how households value togetherness, what benefits and costs it accrues, or how it interacts with other time uses. Addressing these points has been this paper's main goal.

We extend the classical collective model to include togetherness in the form of joint leisure and joint childcare. Togetherness naturally requires that spouses synchronize their work schedules so as to be physically together at the same time. We have kept our analysis nonparametric providing the first nonparametric characterization and empirical implementation of togetherness based on revealed preferences.

Our model fits the data substantially better than the classical model. We estimate the value of togetherness and we find it substantial: households pay on average 17% of the hourly wage to convert an hour of childcare from private to joint. For leisure this stands at 10% of the wage. We estimate that joint childcare takes up almost 92% of overall childcare of one parent. The results suggest that togetherness is an important component of household time use despite it being relatively overlooked in the literature.

Our model offers a few simple policy instruments, such as the wage premium for certain types of work or the cost of market childcare. The core of our policy argument, however, is that togetherness comes along with certain costs which likely accrue differently to men and women in the household. In our model, these costs are forgone flexibility in the labor market and forgone specialization at home. Policies interested in togetherness should therefore closely look into these costs.

Appendices

A Data appendix

A.1 Construction of baseline sample and variables

The main data we use come from the Longitudinal Internet Studies for the Social sciences in the Netherlands (CentERdata, 2012). The LISS consists of different questionnaires, each with its own data covering a given topic. We combine data from the (1.) *Household Box* for baseline demographics and income, (2.) *Core Study 5 Family and Household* for household composition, the presence and age of children etc., (3.) *Core Study 6 Work and Schooling* for working hours and timing of work, and (4.) *Assembled Study 34 Time Use and Consumption* for consumption and time use. The latter module, which provides the bulk of our data, is described in detail in Cherchye et al. (2012). All modules above are administered to the same subjects.

We draw data from three waves, covering calendar years 2009, 2010, and 2012. Although the LISS is running continuously since late 2007, it is only in those three years that we can obtain consistent consumption and time use data. Our sample consists of 398 household×year observations on households with children up to 12 years old and in which both spouses participate in the labor market. We construct this sample as follows:

1. Within our time frame, we select households where two spouses are present.

2. We drop households who experience marital status changes within a given year as their behavior may be constrained by such changes. We also drop those with inconsistent information on gender and year of birth, homosexual couples, as well as couples where both spouses declare themselves as household heads (this makes hard to assign information subsequently).

3. We require that a household has completed the *Household Box*, *Core Study 5*, *Core Study 6*, and *Assembled Study 34*; then we merge data per household across studies. Up to this point, our simple selection criteria result in approximately 3,130 household×year observations, already a fraction of the approximately 10,200 household×year observations present prior to any selection (the last number counts unique households in *Core Study 5* in years 2009, 2010, and 2012).

4. Both spouses must be between 25-60 years old (we do not want time use to be restricted by statutory retirement or schooling) and participate in the labor market. The latter restricts our sample to 1,410 household×year observations.

5. Households must have at least one child up to 12 years old. This leaves us with 477 observations representing another large, but inevitable given our focus, sample cut.

6. Finally, we require that the age and gender of children is reported consistently, that parents have non-missing childcare, and that consumption is not missing, unusable, or zero. Our final sample has 398 household×year observations.

Further we define the variables we use in the paper. The respondents submit answers via online questionnaires that refer to a typical week or a typical month in the past.

Joint leisure is the leisure time that spouses spend together with one another. Both spouses report this (variables `bf09a064-65` in wave 2009). Differences in spouses' response are in most cases negligible; there are some households, however, where the difference is large but seemingly unrelated to household characteristics or other time use variables. We opted for using the minimum of the spouses' respective responses as this is the only choice that guarantees non-negative private leisure. *Private leisure* is the difference between individual total leisure and joint leisure. *Individual total leisure* is the time that each spouse spends on leisure activities (variables `bf09a021-22` in 2009).

Individual total childcare is the time each spouse spends on activities with children (variables `bf09a013-14` in wave 2009). The LISS contains a large array of separate questions on various time uses (e.g. chores, personal care, commuting, and more), so we are confident that the childcare measure indeed reflects time devoted to children.

Market work is the time each spouse spends working for pay on the market, including time on a second job (if any). These are variables `cw09b127` for the main job and `cw09b144` for a second job from Core Study 6. For commuting (robustness check 1 in section 5.4) we use `bf09a007/bf09a008`. We impose a theoretical maximum of 84 weekly market hours (12 hours per day $\times$ 7 days per week). We split market work into regular and irregular hours according to the rule described in section 5.1; to implement this, we use indicator variables for how often one works irregular hours (variable `cw09b425` in wave 2009). To check the sectors where irregular hours are concentrated (sections 2 and 3.1), we use information on the employment sector of individuals (variable `cw09b402`).

We require that the weekly time budget of each spouse adds up to 168 hours. To confirm this we use information on several time uses recorded in Assembled Study 34. Interestingly, the vast majority of households have their time uses add up to exactly (or very close to) 168 hours; for the rest we make a normalization ourselves.

Hourly wages are calculated as monthly earnings over monthly hours of work. *Monthly earnings* is the personal monthly income net of taxes (variable `nettoink_f`), which, unless directly reported by the individual, is imputed by the LISS based on gross income.

Parental consumption is the raw sum of expenditure on various goods, net of goods for children. These include food (at home or away), excluding food of children, tobacco products, clothing, personal care products and services, medical costs, leisure activities costs, costs of further schooling, donations and gifts, rent, household utilities, transport costs, insurance costs, alimony and financial support for children not living at home, costs of debts and loans, and home maintenance costs (variables `bf09a066-70`, `bf09a072-77`,

Table A.1 – Linear regressions: joint variation between leisure and childcare

	(1) leisure male	(2) leisure female	(3) joint leisure	(4) childcare male	(5) childcare female
leisure male		-0.103 (0.056)	0.467 (0.051)	-0.161 (0.083)	-0.070 (0.066)
leisure female	-0.135 (0.075)		0.609 (0.054)	-0.037 (0.094)	-0.427 (0.110)
joint leisure	0.455 (0.072)	0.451 (0.049)		0.023 (0.077)	0.167 (0.080)
childcare male	-0.092 (0.046)	-0.016 (0.042)	0.013 (0.046)		0.352 (0.045)
childcare female	-0.044 (0.042)	-0.202 (0.050)	0.107 (0.050)	0.382 (0.058)	

Notes: The table reports coefficients and st.errors from linear regressions of log leisure and childcare on themselves and a constant. St.errors are clustered at the household level. In column 4 we report a coefficient on female childcare at 0.382. This differs from $\widehat{\beta} = 0.28$ in fig. 2 because the time use variables here are in log. Results in levels are qualitatively similar.

`bf09a095-103` in wave 2009). Several of those items (those considered public expenditure) are reported by both spouses. The responses match surprisingly well. We opt for using the female responses as they include slightly fewer zeros. We have dropped one household where husband-wife responses differ by orders of magnitude.

Child consumption is the raw sum of expenditure on goods for children. These include food (at home or away), clothing, personal care products and services, medical costs, leisure activities costs, costs of schooling, and gifts and presents (variables `bf09a093`, `bf09a105-113` in wave 2009). We impute child consumption for four households for which it is missing. The imputation is statistical: we regress C_K (non-missing in the majority of households) on a large array of demographics, work hours, and parents' consumption, and we predict it for the households for which it is missing.

Table A.1 presents the joint variation between uses of time (each time use is regressed on all other uses and a constant). The linear regression coefficients are similar to correlations because the various time uses have similar cross-sectional variances. Table A.2 presents linear regressions of leisure and childcare on the gender wage gap, the incidence of irregular work, select demographics, and education and occupation dummies.

A.2 Non-participation, time diary, and childless couples

Our estimation method requires wages for both spouses, thus we select households in which both spouses participate in the labor market. Here we first explore how leisure and childcare in our sample compare to leisure and childcare among couples in which spouses do not necessarily participate in the market, but for whom all other selection criteria (e.g. age, presence of young children) still apply. This larger sample has 530 household/year observations. Among them, 98% of men and 76% of women work.

Table A.2 – Linear regressions: leisure and childcare with controls for education and occupation

	(1) leisure male	(2) leisure female	(3) joint leisure	(4) childcare male	(5) childcare female
gender wage gap	0.012 (0.005)	-0.004 (0.006)	-0.001 (0.006)	0.007 (0.008)	-0.013 (0.008)
1[irregular work male]	0.042 (0.076)	-0.018 (0.065)	-0.059 (0.087)	0.111 (0.103)	0.021 (0.073)
1[irregular work female]	-0.146 (0.078)	-0.075 (0.075)	-0.126 (0.092)	0.156 (0.096)	0.075 (0.079)
gender age gap	-0.009 (0.011)	-0.004 (0.011)	-0.005 (0.013)	0.005 (0.014)	0.008 (0.014)
1[child 4-6 yrs]	-0.045 (0.096)	0.067 (0.086)	-0.006 (0.104)	-0.229 (0.112)	-0.156 (0.094)
1[child 7-12 yrs]	0.061 (0.083)	0.112 (0.081)	-0.031 (0.096)	-0.592 (0.102)	-0.734 (0.101)
Male education:					
secondary	-0.036 (0.195)		0.583 (0.265)	-0.093 (0.277)	
vocational	0.020 (0.190)		0.495 (0.268)	0.083 (0.268)	
university	-0.004 (0.203)		0.331 (0.270)	-0.067 (0.285)	
post graduate	0.107 (0.219)		0.821 (0.296)	0.291 (0.309)	
Male occupation:					
higher professional	0.002 (0.154)		0.048 (0.182)	-0.354 (0.191)	
middle professional	0.045 (0.166)		-0.062 (0.183)	-0.282 (0.169)	
commercial	-0.041 (0.162)		0.015 (0.172)	-0.297 (0.164)	
other office work	-0.084 (0.174)		0.003 (0.179)	-0.382 (0.172)	
skilled work	0.077 (0.183)		0.064 (0.212)	-0.158 (0.199)	
semi-skilled	0.156 (0.185)		-0.038 (0.207)	-0.100 (0.222)	
manual work	-0.269 (0.279)		-0.470 (0.353)	-0.269 (0.407)	
Female education:					
secondary		0.182 (0.176)	0.236 (0.394)		-0.081 (0.278)
vocational		0.106 (0.173)	0.261 (0.385)		-0.105 (0.279)
university		0.099 (0.187)	0.110 (0.389)		-0.045 (0.282)
post graduate		0.289 (0.208)	0.355 (0.412)		-0.112 (0.335)
Female occupation:					
higher professional		-0.269 (0.187)	0.082 (0.235)		0.214 (0.253)
middle professional		-0.129 (0.165)	0.174 (0.226)		-0.078 (0.238)
commercial		-0.119 (0.212)	0.034 (0.245)		0.083 (0.246)
other office work		0.051 (0.176)	0.313 (0.228)		-0.054 (0.246)
skilled work		0.120 (0.180)	0.157 (0.469)		-0.094 (0.269)
semi-skilled		0.283 (0.213)	0.619 (0.271)		-0.158 (0.306)
manual work		0.099 (0.231)	0.235 (0.322)		0.040 (0.294)
constant	3.118 (0.253)	2.876 (0.242)	1.275 (0.506)	2.734 (0.334)	3.227 (0.357)
R^2	0.05	0.06	0.12	0.14	0.22
# of observations	391	394	361	387	395

Notes: The table reports coefficients and standard errors from linear regressions of log leisure and childcare on the gender wage gap (male−female wage), the gender age gap (male−female age), indicator variables for whether a spouse works irregular hours *at least sometimes*, for children's age (excluding children less than 4 years old), for education (excluding primary education), for profession, and a constant. Profession takes discrete values for academic or independent occupation (architect, physician, scholar, engineer; excluded category), higher professional (manager, director, supervisory civil servant), middle professional (teacher, artist, nurse, social worker), commercial (department manager, shopkeeper), other office work (administrative assistant, accountant, sales assistant), skilled work (car mechanic, foreman, electrician), semi-skilled (driver, factory worker), and manual work (cleaner, packer). Standard errors are clustered at the household level. The number of observations varies because not everyone has positive amounts of leisure and childcare (see e.g. figure 1 for joint leisure) or because occupation is missing. Results in levels are qualitatively similar.

Table A.3 – Descriptive statistics, sample unconditional on market participation: weekly leisure and childcare

	mean	median	10^{th} pct.	90^{th} pct.
leisure male	26.4	25.0	10.0	44.5
leisure female	23.9	21.0	8.0	41.0
joint leisure	9.5	8.0	1.0	20.0
childcare male	12.9	10.0	3.0	24.8
childcare female	23.4	20.0	5.7	45.0

Notes: The table reports the average, median, 10^{th} and 90^{th} percentiles of leisure and childcare. All statistics are calculated over 530 household/year observations that satisfy all our baseline sample criteria except labor market participation.

Table A.4 – Linear regressions, sample unconditional on market participation: joint variation between leisure and childcare

	(1) leisure male	(2) leisure female	(3) joint leisure	(4) childcare male	(5) childcare female
leisure male		-0.116	0.499	-0.189	-0.055
		(0.050)	(0.042)	(0.068)	(0.058)
leisure female	-0.133		0.585	-0.027	-0.349
	(0.059)		(0.044)	(0.074)	(0.092)
joint leisure	0.449	0.457		0.091	0.105
	(0.058)	(0.043)		(0.064)	(0.069)
childcare male	-0.106	-0.013	0.057		0.316
	(0.037)	(0.036)	(0.041)		(0.041)
childcare female	-0.030	-0.167	0.064	0.308	
	(0.033)	(0.042)	(0.041)	(0.046)	

Notes: The table reports coefficients and standard errors from linear regressions of log leisure and childcare on themselves and a constant. Standard errors are clustered at the household level. The regressions run over the larger sample that satisfies all our baseline sample criteria except labor market participation. Results in levels are qualitatively similar.

Table A.3 mimics table 1 in the main text, and table A.4 mimics table A.1. The distributions of leisure and childcare in the larger sample look remarkably similar to our baseline sample. The only noticeable difference is that women spend on average a bit more time on childcare than before (23.4 vs. 21.7 weekly hours). The correlations between the time use variables look similar to the baseline. Again, childcare correlates negatively to overall leisure but positively to joint leisure.

The LISS also includes a small time diary conducted in year 2013. The diary, officially *Assembled Study 122 Time Use*, contains detailed time use for 224 individuals. Using a mobile application, eligible subjects record their activities in 10-minute intervals in a typical weekday and weekend. They also record with whom, if any, they carry out a given activity, which is what allows us to categorize time as private or joint.

There are three reasons why we cannot use the time diary in our empirical exercise. First, consumption and time use data are available in years 2009, 2010, 2012, but *not* in 2013. Second, our model is for married couples with young children. Once we apply this selection in the time diary and remove non-respondents, we are left with only 12 observations without even conditioning on labor market participation. Third, no more than one person per household participates in the diary so one cannot use it to check what *both* spouses do at a time. Nevertheless, we observe among the 12 individuals that: 1.) weekly overall and joint leisure is effectively similar to the first three lines of tables 1 and A.3; 2.) joint childcare exists, it is prevalent, and it accounts on average for about 40% of a parent's overall childcare, consistent with our results in appendix table C.1.

In section 6.2 we repeat our empirical exercise over a sample of childless couples. This includes couples who do not (yet) have children but for whom all other selection criteria (e.g. age, participation) still apply. This sample has 346 household/year observations.

B Model appendix

B.1 Upper limit on joint time

Here we show how we obtain condition (5) in the main text. Let $\mathcal{T}_m$ be the total amount of time after sleep, personal care, chores etc. If this is different between spouses, the applicable $\mathcal{T}_m$ is the lowest between the two. Togetherness cannot be larger than this, therefore $l_J + t_J \leq \mathcal{T}_m$. The spouses, however, work privately on the market, so their available right hand side time is actually even lower. So we need to subtract the amount of time they are at work. If there is no irregular work, we subtract the *maximum* of the spouse's regular hours as we have assumed these hours always overlap. The spouse who works for longer limits maximum togetherness, therefore $l_J + t_J \leq \mathcal{T}_m - \max\left\{h_1^R, h_2^R\right\}$. If there is irregular work, we further subtract the total amount of such work by *either* spouse as we have assumed irregular hours never overlap. This gives us

$$l_J + t_J \leq \mathcal{T}_m - \max\left\{h_1^R, h_2^R\right\} - h_1^I - h_2^I$$

which is condition (5) in the main text. Note that we could, in principle, tighten this further by also introducing private leisure and private childcare. This would require assumptions on the joint timing of these activities, i.e. which of these private activities may overlap between spouses and to which extent. We refrain from making such assumptions as, unlike market work, there is little empirical guidance on this issue. Finally, this condition can be easily adapted if workers have some control over the schedule of irregular work (i.e. our robustness checks 2 and 3 in section 5.4).

B.2 Equivalent formulations of household problem

In this appendix we reformulate our baseline problem (P) in three alternative ways. We show that a formulation in which the couple cares about child utility insofar as it does not drop below a lower bound, one in which the parents obtain utility from the utility of their children through caring, and one in which child utility is in fact child quality produced by parental inputs are all *equivalent*. To simplify the illustration, we rewrite private leisure as $l_m = \mathcal{T}_m - l_J - t_m - t_J - h^R_m - h^I_m$. We plug this expression whenever l_m appears below, instead of treating the time budgets as separate constraints.

Let the objective function in our baseline problem (P) be $\mathcal{U}^0(\mathcal{C}) = \mu_1 U_1(l_1, l_J, C_P) + \mu_2 U_2(l_2, l_J, C_P) + \mu_K U_K(t_1, t_2, t_J, C_K)$. The Lagrangian function is then given by

$$\begin{aligned}\mathcal{L}^0(\mathcal{C}) = &\ \mathcal{U}^0(\mathcal{C}) + \lambda\Big(Y + \textstyle\sum_{m=1}^2 w_m h^R_m + \sum_{m=1}^2 (w_m + p_m) h^I_m - C_P - C_K - w_K T_K\Big) \\ &+ \tau_J\Big(\mathcal{T}_m - \max\left\{h^R_1, h^R_2\right\} - h^I_1 - h^I_2 - l_J - t_J\Big) \\ &+ \tau_K\Big(\textstyle\sum_{m=1}^2 t_m + t_J + T_K - \mathcal{T}_K\Big).\end{aligned}$$

We use $\mathcal{L}^0(\mathcal{C})$ to obtain the problem's optimality conditions in section 3.3.

Reformulation 1: lower bound to child utility. Suppose that instead of $\mathcal{U}^0(\mathcal{C})$, the household objective function is $\mathcal{U}^1(\mathcal{C}) = \mu_1 U_1(l_1, l_J, C_P) + \mu_2 U_2(l_2, l_J, C_P)$ as in the public goods collective model of Blundell et al. (2005). The problem is subject to the same constraints as in the baseline, and an additional constraint

$$\mu_K : \qquad U_K(t_1, t_2, t_J, C_K) \geq u_K$$

which ensures that parents care for children's welfare insofar as it does not drop below an exogenous threshold u_K. The Lagrangian function is then given by

$$\begin{aligned}\mathcal{L}^1(\mathcal{C}) = &\ \mathcal{U}^1(\mathcal{C}) + \mu_K\Big(U_K(t_1, t_2, t_J, C_K) - u_K\Big) \\ &+ \lambda\Big(Y + \textstyle\sum_{m=1}^2 w_m h^R_m + \sum_{m=1}^2 (w_m + p_m) h^I_m - C_P - C_K - w_K T_K\Big) \\ &+ \tau_J\Big(\mathcal{T}_m - \max\left\{h^R_1, h^R_2\right\} - h^I_1 - h^I_2 - l_J - t_J\Big) \\ &+ \tau_K\Big(\textstyle\sum_{m=1}^2 t_m + t_J + T_K - \mathcal{T}_K\Big).\end{aligned}$$

The threshold u_K can be dropped from the first line because it does not depend on choice variables $\mathcal{C}$. Then $\mathcal{L}^1(\mathcal{C})$ becomes exactly the same as $\mathcal{L}^0(\mathcal{C})$.

Reformulation 2: caring. Suppose that the parents are not "egoistic" but they care for their children's welfare à la Becker. Let each spouse's preferences be given by an altruistic index $W_m = U_m(l_m, l_J, C_P) + \theta_m U_K(t_1, t_2, t_J, C_K)$. Here θ_m captures the degree to which parent m cares for his/her child's welfare. Suppose the household only maximizes the weighted sum of the parents' indexes. The household objective function can then be written as:

$$\begin{aligned}\mathcal{U}^2(\mathcal{C}) = \mu_1 W_1 + \mu_2 W_2 &= \mu_1\big(U_1 + \theta_1 U_K\big) + \mu_2\big(U_2 + \theta_2 U_K\big)\\ &= \mu_1 U_1 + \mu_2 U_2 + (\mu_1\theta_1 + \mu_2\theta_2)U_K\\ &= \mu_1 U_1 + \mu_2 U_2 + \widetilde{\mu}_K U_K.\end{aligned}$$

The last equality uses a simple redefinition of the utility weight, i.e. $\widetilde{\mu}_K = \mu_1\theta_1 + \mu_2\theta_2$. This is possible as we do not impose restrictions on μ_1, μ_2, $\widetilde{\mu}_K$. Replacing $\mathcal{U}^0(\mathcal{C})$ in the baseline with $\mathcal{U}^2(\mathcal{C})$, we obtain Lagrangian $\mathcal{L}^2(\mathcal{C})$ that is exactly the same as baseline $\mathcal{L}^0(\mathcal{C})$. This is in fact a well-known result (e.g. Chiappori, 1992). The crucial assumption is the separability between U_m and U_K in the caring functions. Chiappori (1992) shows extensions to weaker forms of separability than the strong version that we used above.

Reformulation 3: child quality. Let child quality Q_K be produced by a weakly increasing function of parental time inputs and child expenditure, $Q_K = F(t_1, t_2, t_J, C_K)$. F is a child development production function à la Del Boca et al. (2014). The parents enjoy utility from child quality and each spouse's preferences are an index $W_m = U_m(l_m, l_J, C_P) + \alpha_m U_K(Q_K)$. Here $\alpha_m > 0$ is the individual weight on child quality while U_K is the weakly increasing utility over it. Given the production function underlying Q_K, we rewrite the child quality utility term as $\widetilde{U}_K(t_1, t_2, t_J, C_K) = U_K F(t_1, t_2, t_J, C_K)$ where $\widetilde{U}_K$ is simply a composite function. The rest of the discussion is then similar to the previous formulation. The household objective function is:

$$\begin{aligned}\mathcal{U}^3(\mathcal{C}) = \mu_1 W_1 + \mu_2 W_2 &= \mu_1\big(U_1 + \alpha_1 \widetilde{U}_K\big) + \mu_2\big(U_2 + \alpha_2 \widetilde{U}_K\big)\\ &= \mu_1 U_1 + \mu_2 U_2 + (\mu_1\alpha_1\widetilde{U}_K + \mu_2\alpha_2\widetilde{U}_K)\\ &= \mu_1 U_1 + \mu_2 U_2 + \widetilde{\mu}_K \widetilde{U}_K.\end{aligned}$$

The last equality simply redefines the child quality term as $\widetilde{\mu}_K\widetilde{U}_K = \mu_1\alpha_1\widetilde{U}_K + \mu_2\alpha_2\widetilde{U}_K$. This is possible for the same reasons as in the case of caring. Replacing $\mathcal{U}^0(\mathcal{C})$ in the baseline with $\mathcal{U}^3(\mathcal{C})$, we get Lagrangian $\mathcal{L}^3(\mathcal{C})$ that is exactly the same as baseline $\mathcal{L}^0(\mathcal{C})$.

Our baseline formulation and the three alternative ones are all equivalent to each other. It is impossible to separate them (at least not without additional information) and the interpretation one chooses to put forward (e.g. caring, child development etc.) is down to taste. While we choose a baseline formulation in which childcare superficially appears to affect only children's but not parents' utility, alternative formulations 2 & 3 highlight that childcare may also enter parents' utility through caring or child quality. So, regardless of the specific interpretation one puts forward, our formulation is consistent with parents enjoying time with their children because such time improves children's welfare.

C Revealed preferences appendix

C.1 Proof of Proposition 1

Since Proposition 1 is a special case of Proposition 2, we do not repeat the full proof here but we refer the reader to appendix C.2 for the full proof. Here we prove that if $\tilde{S}_K$ satisfies GARP, then S_K also satisfies GARP.

Let $\widetilde{S}_K = \{\delta^{(v)}_{K,t_1}, \delta^{(v)}_{K,t_2}, \delta^{(v)}_{K,t_J}, \delta^{(v)}_{K,C_K};\ t^{(v)}_1, t^{(v)}_2, t^{(v)}_J, C^{(v)}_K\}_{v\in V}$ and $S_K = \{\delta^{(v)}_{K,t_1}, \delta^{(v)}_{K,t_2}, \delta^{(v)}_{K,C_K};$ $T^{(v)}_1, T^{(v)}_2, C^{(v)}_K\}_{v\in V}$. Note that $\delta^{(v)}_{K,C_K} = 1$ from the problem's first order conditions. It follows that

$$\begin{aligned}
&\delta^{(v)}_{K,t_1}(t^{(s)}_1 - t^{(v)}_1) + \delta^{(v)}_{K,t_2}(t^{(s)}_2 - t^{(v)}_2) + \delta^{(v)}_{K,t_J}(t^{(s)}_J - t^{(v)}_J) + (C^{(s)}_K - C^{(v)}_K) \le 0\\
\Leftrightarrow\ & w^{(v)}_1(t^{(s)}_1 - t^{(v)}_1) + w^{(v)}_2(t^{(s)}_2 - t^{(v)}_2) + (w^{(v)}_1 + w^{(v)}_2)(t^{(s)}_J - t^{(v)}_J) + (C^{(s)}_K - C^{(v)}_K) \le 0\\
\Leftrightarrow\ & w^{(v)}_1(T^{(s)}_1 - T^{(v)}_1) + w^{(v)}_2(T^{(s)}_2 - T^{(v)}_2) + (C^{(s)}_K - C^{(v)}_K) \le 0.
\end{aligned}$$

To obtain the second line, we use that $\delta^{(v)}_{K,t_m} = w^{(v)}_m$ and $\delta^{(v)}_{K,t_J} = w^{(v)}_1 + w^{(v)}_2$ consistent with the first order conditions associated with Proposition 1. To obtain the third line, we use that $t^{(v)}_1 + t^{(v)}_J = T^{(v)}_1$ and $t^{(v)}_2 + t^{(v)}_J = T^{(v)}_2$. Summarizing, any revealed preference relation that exists for $\tilde{S}_K$ exists also for S_K, and vice versa. We may therefore conclude that S_K satisfies GARP if and only if $\tilde{S}_K$ satisfies GARP. The equivalence holds also with corners $t^{(v)}_J = 0$. Rationality implies $\delta^{(v)}_{K,t_J} \le w^{(v)}_1 + w^{(v)}_2$ because it also implies $\delta^{(v)}_{K,t_J} = w^{(v)}_1 + w^{(v)}_2$. Next, $\delta^{(v)}_{K,t_J} \le w^{(v)}_1 + w^{(v)}_2$ implies rationality because $\delta^{(v)}_{K,t_J}(t^{(s)}_J - 0) \le (w^{(v)}_1 + w^{(v)}_2)(t^{(s)}_J - 0)$, thus constructing the same preference relations as before.

C.2 Proof of Proposition 2

In the first part of the proof ((1) $\Rightarrow$ (2)), we use that a necessary condition for consistency with a convex problem is that the Karush-Kuhn-Tucker conditions hold. Assuming continuity and concavity of U_m and U_K we replace the partial derivatives of the utility functions with suitable super-gradients to obtain

$$\begin{aligned}
&u^{(s)}_m - u^{(v)}_m \le \eta^{(v)}_m \left[\delta^{(v)}_{m,l_m}(l^{(s)}_m - l^{(v)}_m) + \delta^{(v)}_{m,l_J}(l^{(s)}_J - l^{(v)}_J) + \delta^{(v)}_{m,C_P}(C^{(s)}_P - C^{(v)}_P)\right]\\
&u^{(s)}_K - u^{(v)}_K \le \eta^{(v)}_K \left[\delta^{(v)}_{K,t_1}(t^{(s)}_1 - t^{(v)}_1) + \delta^{(v)}_{K,t_2}(t^{(s)}_2 - t^{(v)}_2) + \delta^{(v)}_{K,t_J}(t^{(s)}_J - t^{(v)}_J) + (C^{(s)}_K - C^{(v)}_K)\right],
\end{aligned}$$

with $\eta^{(v)}_m = \lambda^{(v)}/\mu^{(v)}_m$ and $\eta^{(v)}_K = \lambda^{(v)}/\mu^{(v)}_K$. Finally, given that $\{\delta^{(v)}_{m,l_m}, \delta^{(v)}_{m,l_J}, \delta^{(v)}_{m,C_P}\}$ and $\{\delta^{(v)}_{K,t_1}, \delta^{(v)}_{K,t_2}, \delta^{(v)}_{K,t_J}, 1\}$ act as prices in these Afriat-like inequalities (in the latter set, $\delta^{(v)}_{K,C_K} = 1$ is the price of child expenditure), and $\{l^{(v)}_m, l^{(v)}_J, C^{(v)}_P\}$ and $\{t^{(v)}_1, t^{(v)}_2, t^{(v)}_J, C^{(v)}_K\}$ as quantities, we can reformulate these conditions in terms of consistency of $\{\delta^{(v)}_{m,l_m}, \delta^{(v)}_{m,l_J}, \delta^{(v)}_{m,C_P};$ $l^{(v)}_m, l^{(v)}_J, C^{(v)}_P\}_{v\in V}$ and $\{\delta^{(v)}_{K,t_1}, \delta^{(v)}_{K,t_2}, \delta^{(v)}_{K,t_J}, 1; t^{(v)}_1, t^{(v)}_2, t^{(v)}_J, C^{(v)}_K\}_{v\in V}$ with GARP (Varian, 1982).

In the second part of the proof ((2) $\Rightarrow$ (1)), we have to show that there exist weights

$\mu_m^{(v)}$ and $\mu_K^{(v)}$, and utility functions U_m and U_K that rationalize the data with togetherness, provided that the conditions in Proposition 2 hold. To this end, we first construct the utility functions. Let

$$U_m(l_m, l_J, C_P) = \min_{(s)} u_m^{(s)} + \eta_m^{(s)} \left(\delta_{m,l_m}^{(s)}(l_m - l_m^{(s)}) + \delta_{m,l_J}^{(s)}(l_J - l_J^{(s)}) + \delta_{m,C_P}^{(s)}(C_P - C_P^{(s)})\right)$$
$$U_K(t_1, t_2, t_J, C_K) = \min_{(s)} u_K^{(s)} + \eta_K^{(s)} \left(\delta_{K,t_1}^{(s)}(t_1 - t_1^{(s)}) + \delta_{K,t_2}^{(s)}(t_2 - t_2^{(s)}) + \delta_{K,t_J}^{(s)}(t_J - t_J^{(s)}) + (C_K - C_K^{(s)})\right).$$

One can verify that these piecewise linear functions are nonsatiated, continuous and concave. After all, the minimum of concave functions is concave. One can also show that $U_1(l_1^{(v)}, l_J^{(v)}, C_P^{(v)}) = u_1^{(v)}$. By definition of U_1:

$$U_1(l_1, l_J, C_P) = \min_{(s)} u_1^{(s)} + \eta_1^{(s)} \left(\delta_{1,l_1}^{(s)}(l_1 - l_1^{(s)}) + \delta_{1,l_J}^{(s)}(l_J - l_J^{(s)}) + \delta_{1,C_P}^{(s)}(C_P - C_P^{(s)})\right).$$

This implies $U_1(l_1^{(v)}, l_J^{(v)}, C_P^{(v)}) \leq u_1^{(v)}$. We cannot have that $U_1(l_1^{(v)}, l_J^{(v)}, C_P^{(v)}) < u_1^{(v)}$ because this violates the Afriat inequalities. Suppose that $U_1(l_1^{(v)}, l_J^{(v)}, C_P^{(v)})$ is at its minimum in situation r, i.e.

$$U_1(l_1^{(v)}, l_J^{(v)}, C_P^{(v)}) = u_1^{(r)} + \eta_1^{(r)} \left(\delta_{1,l_1}^{(r)}(l_1^{(v)} - l_1^{(r)}) + \delta_{1,l_J}^{(r)}(l_J^{(v)} - l_J^{(r)}) + \delta_{1,C_P}^{(r)}(C_P^{(v)} - C_P^{(r)})\right).$$

Then $u_1^{(v)} > U_1(l_1^{(v)}, l_J^{(v)}, C_P^{(v)})$ implies

$$u_1^{(v)} > u_1^{(r)} + \eta_1^{(r)} \left(\delta_{1,l_1}^{(r)}(l_1^{(v)} - l_1^{(r)}) + \delta_{1,l_J}^{(r)}(l_J^{(v)} - l_J^{(r)}) + \delta_{1,C_P}^{(r)}(C_P^{(v)} - C_P^{(r)})\right),$$

a contradiction of the Afriat inequalities implied by Proposition 2. This shows that $U_1(l_1^{(v)}, l_J^{(v)}, C_P^{(v)}) = u_1^{(v)}$. A similar reasoning gives that $U_2(l_2^{(v)}, l_J^{(v)}, C_P^{(v)}) = u_2^{(v)}$ and $U_K(t_1^{(v)}, t_2^{(v)}, t_J^{(v)}, C_K^{(v)}) = u_K^{(v)}$.

Second, we need to show that the data under consideration maximize the constructed utility functions. In other words, for any l_m, l_J, h_m^R, h_m^I, t_m, t_J, T_K, C_P, C_K that satisfy

- $C_P + C_K + w_K^{(v)}T_K - \sum_{m=1}^{2}(w_m^{(v)}h_m^R + (w_m^{(v)} + p_m^{(v)})h_m^I)$
  $\leq C_P^{(v)} + C_K^{(v)} + w_K^{(v)}T_K^{(v)} - \sum_{m=1}^{2}(w_m^{(v)}h_m^{R\,(v)} + (w_m^{(v)} + p_m^{(v)})h_m^{I\,(v)})$;
- $l_m + l_J + t_m + t_J + h_m^R + h_m^I = l_m^{(v)} + l_J^{(v)} + t_m^{(v)} + t_J^{(v)} + h_m^{R\,(v)} + h_m^{I\,(v)}$;
- $\sum_{m=1}^{2} t_m + t_J + T_K \geq \sum_{m=1}^{2} t_m^{(v)} + t_J^{(v)} + T_K^{(v)}$; and
- $l_J + t_J + \max\{h_1^R, h_2^R\} + h_1^I + h_2^I \leq l_J^{(v)} + t_J^{(v)} + \max\{h_1^{R(v)}, h_2^{R(v)}\} + h_1^{I(v)} + h_2^{I(v)}$,

it must be that $\sum_m \mu_m^{(v)} U_m(l_m, l_J, C_P) + \mu_K^{(v)} U_K(t_1, t_2, t_J, C_K) \leq \sum_m \mu_m^{(v)} U_m(l_m^{(v)}, l_J^{(v)}, C_P^{(v)}) + \mu_K^{(v)} U_K(t_1^{(v)}, t_2^{(v)}, t_J^{(v)}, C_K^{(v)})$. Supplementary to our construction of utility functions, we also choose $\mu_m^{(v)} = 1/\eta_m^{(v)}$ and $\mu_K^{(v)} = 1/\eta_K^{(v)}$. Thus

$$\mu_1^{(v)} U_1(l_1, l_J, C_P) + \mu_2^{(v)} U_2(l_2, l_J, C_P) + \mu_K^{(v)} U_K(t_1, t_2, t_J, C_K)$$

$$
\begin{aligned}
\leq\ & \textstyle\sum_m \mu_m^{(v)} u_m^{(v)} + \mu_K^{(v)} u_K^{(v)} \\
& + \textstyle\sum_m \left(\delta_{m,l_m}^{(v)}(l_m - l_m^{(v)}) + \delta_{m,l_J}^{(v)}(l_J - l_J^{(v)}) + \delta_{m,C_P}^{(v)}(C_P - C_P^{(v)}) \right) \\
& + \delta_{K,t_1}^{(v)}(t_1 - t_1^{(v)}) + \delta_{K,t_2}^{(v)}(t_2 - t_2^{(v)}) + \delta_{K,t_J}^{(v)}(t_J - t_J^{(v)}) + (C_K - C_K^{(v)}) \\
=\ & \textstyle\sum_m \mu_m^{(v)} u_m^{(v)} + \mu_K^{(v)} u_K^{(v)} \\
& + \textstyle\sum_m (w_m^{(v)} - p_{-m}^{(v)})(l_m - l_m^{(v)}) + (w_1^{(v)} + w_2^{(v)})(l_J - l_J^{(v)}) + (C_P - C_P^{(v)}) \\
& + \textstyle\sum_m (w_m^{(v)} - p_{-m}^{(v)} - w_K^{(v)})(t_m - t_m^{(v)}) + (w_1^{(v)} + w_2^{(v)} - w_K^{(v)})(t_J - t_J^{(v)}) + (C_K - C_K^{(v)}) \\
=\ & \textstyle\sum_m \mu_m^{(v)} u_m^{(v)} + \mu_K^{(v)} u_K^{(v)} \\
& + \textstyle\sum_m (w_m^{(v)} - p_{-m}^{(v)})(l_m - l_m^{(v)}) + (w_1^{(v)} + w_2^{(v)})(l_J - l_J^{(v)}) + (C_P - C_P^{(v)}) \\
& + \textstyle\sum_m (w_m^{(v)} - p_{-m}^{(v)})(t_m - t_m^{(v)}) + (w_1^{(v)} + w_2^{(v)})(t_J - t_J^{(v)}) \\
& - w_K^{(v)}(t_1 + t_2 + t_J - t_1^{(v)} - t_2^{(v)} - t_J^{(v)}) + (C_K - C_K^{(v)}) \\
\leq\ & \textstyle\sum_m \mu_m^{(v)} u_m^{(v)} + \mu_K^{(v)} u_K^{(v)} \\
& + \textstyle\sum_m (w_m^{(v)} - p_{-m}^{(v)})(l_m - l_m^{(v)}) + (w_1^{(v)} + w_2^{(v)})(l_J - l_J^{(v)}) + (C_P - C_P^{(v)}) \\
& + \textstyle\sum_m (w_m^{(v)} - p_{-m}^{(v)})(t_m - t_m^{(v)}) + (w_1^{(v)} + w_2^{(v)})(t_J - t_J^{(v)}) + w_K^{(v)}(T_K - T_K^{(v)}) + (C_K - C_K^{(v)}) \\
=\ & \textstyle\sum_m \mu_m^{(v)} u_m^{(v)} + \mu_K^{(v)} u_K^{(v)} \\
& + \textstyle\sum_m (w_m^{(v)} - p_{-m}^{(v)})(L_m + T_m - L_m^{(v)} - T_m^{(v)}) + \textstyle\sum_m p_{-m}^{(v)}(l_J + t_J - l_J^{(v)} - t_J^{(v)}) \\
& + w_K^{(v)}(T_K - T_K^{(v)}) + (C_P - C_P^{(v)}) + (C_K - C_K^{(v)}) \\
\leq\ & \textstyle\sum_m \mu_m^{(v)} u_m^{(v)} + \mu_K^{(v)} u_K^{(v)} \\
& - \textstyle\sum_m (w_m^{(v)} - p_{-m}^{(v)})(h_m^R + h_m^I - h_m^{R\,(v)} - h_m^{I\,(v)}) \\
& - \textstyle\sum_m p_{-m}^{(v)} \left(\max\{h_1^R, h_2^R\} + h_1^I + h_2^I - \max\{h_1^{R(v)}, h_2^{R(v)}\} - h_1^{I(v)} - h_2^{I(v)} \right) \\
& + w_K^{(v)}(T_K - T_K^{(v)}) + (C_P - C_P^{(v)}) + (C_K - C_K^{(v)}) \\
\leq\ & \textstyle\sum_m \mu_m^{(v)} u_m^{(v)} + \mu_K^{(v)} u_K^{(v)} \\
& - \textstyle\sum_m \left(w_m^{(v)}(h_m^R - h_m^{R\,(v)}) + w_m^{(v)}(h_m^I - h_m^{I\,(v)}) \right) + p_o^{(v)}((h_n^R - h_n^{R(v)}) + (h_n^I - h_n^{I\,(v)})) \\
& - p_o^{(v)} \left((h_n^R - h_n^{R(v)}) + (h_1^I - h_1^{I(v)}) + (h_2^I - h_2^{I(v)}) \right) \\
& + w_K^{(v)}(T_K - T_K^{(v)}) + (C_P - C_P^{(v)}) + (C_K - C_K^{(v)}) \\
=\ & \textstyle\sum_m \mu_m^{(v)} u_m^{(v)} + \mu_K^{(v)} u_K^{(v)} \\
& - w_n^{(v)} H_n - w_o^{(v)} h_o^R - (w_o^{(v)} + p_o^{(v)}) h_o^I + w_n^{(v)} H_n{}^{(v)} + w_o^{(v)} h_o^{R(v)} + (w_o^{(v)} + p_o^{(v)}) h_o^{I(v)} \\
& + C_P + C_K + w_K^{(v)} T_K - C_P^{(v)} - C_K^{(v)} - w_K^{(v)} T_K^{(v)} \\
\leq\ & \textstyle\sum_m \mu_m^{(v)} u_m^{(v)} + \mu_K^{(v)} u_K^{(v)} \\
=\ & \mu_1^{(v)} U_1(l_1^{(v)}, l_J^{(v)}, C_P^{(v)}) + \mu_2^{(v)} U_2(l_2^{(v)}, l_J^{(v)}, C_P^{(v)}) + \mu_K^{(v)} U_K(t_1^{(v)}, t_2^{(v)}, t_J^{(v)}, C_K^{(v)}).
\end{aligned}
$$

The first relation in the sequence (inequality) follows by construction of U_m and U_K. The second (equality) uses the conditions of Proposition 2 to replace shadow prices with wages and premiums. The third (equality) simply rearranges the terms. The fourth relation (inequality) follows from the childcare constraint. The fifth (equality) rearranges terms and uses the leisure and childcare identities (2). The sixth (inequality) stems from the time budgets and the upper bound on togetherness. The seventh relation (inequality) is due to concavity of function $-\max\{h_1^R, h_2^R\}$. Let index $o \in \{1,2\}$ denote the spouse who works the least regular hours and let $n \in \{1,2\}$ denote the spouse who works the most: $h_n^{R(v)} \geq h_o^{R(v)}$. Then $\max\{h_1^R, h_2^R\} - \max\{h_1^{R(v)}, h_2^{R(v)}\} \geq h_n^R - h_n^{R(v)}$. Notice that $p_n^{(v)} = 0$ follows from Proposition 2. The eighth relation (equality) rearranges terms and uses the hours identity (3). The ninth relation (inequality) is due to the budget constraint.

For each bundle l_m, l_J, h_m^R, h_m^I, t_m, t_J, T_K, C_P, C_K that respects the budget constraint, the parental time constraints, the childcare constraint and the upper bound on joint time use, we have thus shown that

$$\begin{aligned} & \mu_1^{(v)} U_1(l_1, l_J, C_P) + \mu_2^{(v)} U_2(l_2, l_J, C_P) + \mu_K^{(v)} U_K(t_1, t_2, t_J, C_K) \\ \leq \ & \mu_1^{(v)} U_1(l_1^{(v)}, l_J^{(v)}, C_P^{(v)}) + \mu_2^{(v)} U_2(l_2^{(v)}, l_J^{(v)}, C_P^{(v)}) + \mu_K^{(v)} U_K(t_1^{(v)}, t_2^{(v)}, t_J^{(v)}, C_K^{(v)}). \end{aligned}$$

We conclude that the constructed weights $\mu_m^{(v)}$, $\mu_K^{(v)}$ and utility functions U_m, U_K effectively provide a collective rationalization with togetherness.

C.3 Set identification of joint childcare

C.3.1 Illustration

The second example of section 4.3 shows that the hypothetical data of table 4 are not consistent with *collective rationality*, as described in Proposition 1. Let us now show that the data *are* consistent with *collective rationality with togetherness*, as per Proposition 2. Invoking the opening condition in Definition 1, bundle A is revealed preferred over bundle B if

$$\delta_{K,t_1}^{(A)} \times (t_1^{(A)} - t_1^{(B)}) + \delta_{K,t_2}^{(A)} \times (t_2^{(A)} - t_2^{(B)}) + \delta_{K,t_J}^{(A)} \times (t_J^{(A)} - t_J^{(B)}) + (C_K^{(A)} - C_K^{(B)}) > 0.$$

This should then imply (invoking the closing condition of Definition 1) that

$$\delta_{K,t_1}^{(B)} \times (t_1^{(B)} - t_1^{(A)}) + \delta_{K,t_2}^{(B)} \times (t_2^{(B)} - t_2^{(A)}) + \delta_{K,t_J}^{(B)} \times (t_J^{(B)} - t_J^{(A)}) + (C_K^{(B)} - C_K^{(A)}) \leq 0.$$

One can use the childcare identity (2) to express private childcare as a function of total childcare and joint childcare. Given that the shadow prices are all disciplined by Proposition 2, it is easy to show that the inequalities are both satisfied if $t_J^{(A)} - t_J^{(B)} \geq \frac{55}{3}$. This offers us some information about unobserved joint childcare in our data: joint childcare is higher at prices A than prices B by at least $\frac{55}{3}$ because joint childcare at A is relatively cheaper ($\delta_{K,t_J}^{(A)}/\delta_{K,t_m}^{(A)} = 37/17$) than joint childcare at prices B ($\delta_{K,t_J}^{(B)}/\delta_{K,t_m}^{(B)} = 17/7$).

While the above application of GARP (Definition 1) illustrates the workings of set identification of childcare in our setting, the result $t_J^{(A)} - t_J^{(B)} \geq \frac{55}{3}$ does not say much about the number of hours joint childcare really takes up in a typical week. Perhaps joint childcare at prices A takes up 100 hours and joint childcare at B takes up zero hours, or perhaps joint childcare at A is about 20 hours while joint childcare at B only one hour. Fortunately there is more information in our design that helps further discipline joint childcare. By construction, joint childcare is always between 0 and the minimum individual total childcare, $t_J \in [0, \min\{T_1, T_2\}]$. This upper bound is because joint childcare cannot be larger than *either* parent's own childcare, so a value of 100 is not a possibility here. This suggests that we should bring in these naive *theoretical* bounds into the problem, namely $0 \leq t_J^{(A)} \leq 20$ and $0 \leq t_J^{(B)} \leq 10$.

Suppose $t_J^{(B)}$ is at its lowest theoretical value (0); then $t_J^{(A)} - t_J^{(B)} \geq \frac{55}{3}$ implies that $t_J^{(A)} \geq \frac{55}{3}$. Suppose, alternatively, that $t_J^{(B)}$ is at its maximum theoretical value (10); $t_J^{(A)} - t_J^{(B)} \geq \frac{55}{3}$ implies that $t_J^{(A)} \geq \frac{85}{3}$ which is higher than the maximum theoretical value for $t_J^{(A)}$ (20). So $t_J^{(B)}$ cannot be at its maximum theoretical value and the sole new information out of this line of argument is that $t_J^{(A)} \geq \frac{55}{3}$. Similarly, suppose $t_J^{(A)}$ is at its updated lowest value ($\frac{55}{3}$); then $t_J^{(A)} - t_J^{(B)} \geq \frac{55}{3}$ implies that $t_J^{(B)} = 0$. Suppose, alternatively, that $t_J^{(A)}$ is at its maximum theoretical value (20); $t_J^{(A)} - t_J^{(B)} \geq \frac{55}{3}$ implies that $t_J^{(B)} \leq \frac{5}{3}$. Combining all this, we conclude that $\frac{55}{3} \leq t_J^{(A)} \leq 20$ and $0 \leq t_J^{(B)} \leq \frac{5}{3}$ are the applicable bounds of joint childcare in the hypothetical data of table 4.

When there are more than two price regimes, as there usually are in practice, the calculations above may become tedious. A somewhat related but slightly more structured way to do the same thing is to work with $\alpha^{(v)} = t_J^{(v)} / \min\{T_1^{(v)}, T_2^{(v)}\}$, namely the fraction that joint childcare at prices v is of its naive theoretical maximum. We can then choose $t_J^{(v)}$'s in order to minimize and maximize $\alpha = (\sum_v 1)^{-1} \sum_v \alpha^{(v)}$ subject to $t_J^{(A)} - t_J^{(B)} \geq \frac{55}{3}$, thus obtain the lowest and highest fractions of joint childcare (out of its naive theoretical maximum) that rationalize the data. This is precisely what we do with our actual data in section 5.2. In our hypothetical data of table 4, the minimization yields $t_J^{(A)} = 18.33$ and $t_J^{(B)} = 0$, corresponding to a lower bound on α of 0.458, while the maximization yields $t_J^{(A)} = 20$ and $t_J^{(B)} = 1.67$, corresponding to an upper bound on α of 0.583.

C.3.2 Optional additional restrictions for empirical implementation

Following the above illustration, table 8 reports bounds on the average $t_J / \min\{T_1, T_2\}$ within household groups consistent with collective rationality with togetherness (Proposition 2). These bounds can be tightened further by imposing additional structure on the problem. One appealing restriction is $\alpha - \Delta \leq t_J^{(v)} / \min\{T_1^{(v)}, T_2^{(v)}\} \leq \alpha + \Delta$ for a choice of Δ. This guarantees that the proportion of joint childcare for a given household v within a group does not deviate too much from the mean proportion in the group. For example, $\Delta = 0.1$ implies that the proportion of joint childcare of any given household does not deviate by more than 10 percentage points from the group average.

Table C.1 – Distribution of bounds on weekly t_J, after additional restrictions

t_J	estimated lower bound	estimated upper bound	naive upper bound
Distribution across households:			
min	0	0	0
first quartile	0	5.12	5.37
median	0	10.00	10.53
mean	3.75	11.92	12.90
third quartile	5.54	16.65	17.78
max	34.75	51.49	51.49

Notes: The table reports the distribution *across households* of the bounds on joint childcare t_J (weekly hours). The results are over 250 households consistent with collective rationality with togetherness (T–CR).

The choice of Δ will matter for the results and an unrealistic value may mean households violate the conditions of Proposition 2. So an interesting choice for Δ is to fix it at the lowest possible level per group that still allows to collectively rationalize the households with togetherness. This effectively minimizes heterogeneity in $t_J^{(v)} / \min\{T_1^{(v)}, T_2^{(v)}\}$ across households in the same group, though it may well increase heterogeneity across different groups. We apply this restriction and we observe that our bounds are sharpened substantially. Table C.1 presents the distribution of bounds (switching from bounds on proportions to bounds on t_J) for all 250 households consistent with T–CR; figure C.1 plots the bounds along with the naive minimum and maximum per household. Each household has a different naive maximum so the bounds on t_J are mechanically household-specific. Figure C.1 reveals that our approach produces informative and often sharp bounds for a

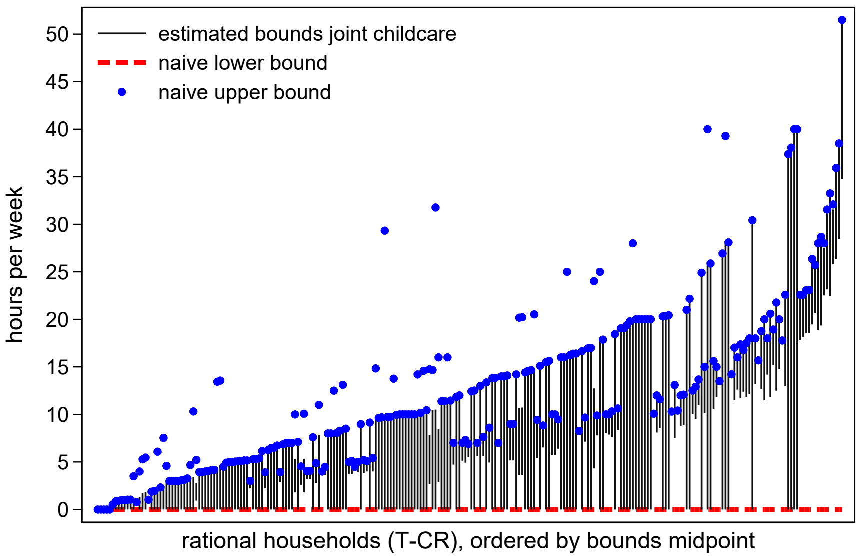


Figure C.1 – Bounds on weekly t_J, after additional restrictions

Notes: The figure plots the estimated bounds on joint childcare t_J (weekly hours) among 250 households consistent with collective rationality with togetherness (T–CR). The naive lower bound is zero; the naive upper bound is $\min\{T_1, T_2\}$.

large fraction of admissible households. Joint childcare takes up a substantial amount of the overall supply of childcare in the family: while on average men spend 13.3 and women 21.7 hours with their children per week (raw data table 1), our average lower bound is 3.8 hours and our average upper bound is 11.9 hours. These results suggest that joint childcare -like joint leisure- is an important component of household time use.

C.4 Mixed integer linear programming problem

In this appendix we show how we formulate our revealed preference conditions as a mixed integer linear programming problem. We define binary variables $x_m^{(s),(v)}$ and $x_K^{(s),(v)}$, with $x_m^{(s),(v)}$ equal to 1 if spouse m 'revealed prefers' bundle s over v and 0 otherwise, and $x_K^{(s),(v)}$ equal to 1 if child bundle s is 'revealed preferred' over v and 0 otherwise.

A set of observations V in D_{LISS} is consistent with collective rationality with togetherness (Proposition 2) if and only if the following problem is feasible (i.e. it has a solution). For all observations $s, s1, v \in V$ in which $h_1^R > h_2^R$ (without loss of generality), there exist

- binary variables $x_1^{(s),(v)}$, $x_2^{(s),(v)}$, $x_K^{(s),(v)}$, all $\in \{0, 1\}$,
- shadow prices associated with joint leisure $\delta_{1,l_J}^{(v)}$, $\delta_{2,l_J}^{(v)} \in R_+$,
- shadow prices associated with parental consumption $\delta_{1,C_P}^{(v)}$, $\delta_{2,C_P}^{(v)} \in R_+$,
- wage premium and price of market childcare $p_m^{(v)}$, $w_K^{(v)}$ within their respective grids,
- levels of private and joint childcare $t_1^{(v)}$, $t_2^{(v)}$, $t_J^{(v)} \in R_+$,

such that

$$\begin{aligned}
\delta_{1,l_J}^{(v)} + \delta_{2,l_J}^{(v)} &= w_1^{(v)} + w_2^{(v)} \\
\delta_{1,C_P}^{(v)} + \delta_{2,C_P}^{(v)} &= 1 \\
t_m^{(v)} + t_J^{(v)} &= T_m^{(v)}
\end{aligned}$$

$$\begin{aligned}
x_1^{(s),(v)} M &> (w_1^{(s)} - p_2^{(s)})(l_1^{(s)} - l_1^{(v)}) + \delta_{1,l_J}^{(s)}(l_J^{(s)} - l_J^{(v)}) + \delta_{1,C_P}^{(s)}(C_P^{(s)} - C_P^{(v)}) \\
x_1^{(s),(s1)} + x_1^{(s1),(v)} &\leq 1 + x_1^{(s),(v)} \\
(1 - x_1^{(s),(v)})M &\geq (w_1^{(v)} - p_2^{(v)})(l_1^{(v)} - l_1^{(s)}) + \delta_{1,l_J}^{(v)}(l_J^{(v)} - l_J^{(s)}) + \delta_{1,C_P}^{(v)}(C_P^{(v)} - C_P^{(s)})
\end{aligned}$$

$$\begin{aligned}
x_2^{(s),(v)} M &> w_2^{(s)}(l_2^{(s)} - l_2^{(v)}) + \delta_{2,l_J}^{(s)}(l_J^{(s)} - l_J^{(v)}) + \delta_{2,C_P}^{(s)}(C_P^{(s)} - C_P^{(v)}) \\
x_2^{(s),(s1)} + x_2^{(s1),(v)} &\leq 1 + x_2^{(s),(v)} \\
(1 - x_2^{(s),(v)})M &\geq w_2^{(v)}(l_2^{(v)} - l_2^{(s)}) + \delta_{2,l_J}^{(v)}(l_J^{(v)} - l_J^{(s)}) + \delta_{2,C_P}^{(v)}(C_P^{(v)} - C_P^{(s)})
\end{aligned}$$

$$\begin{aligned}
x_K^{(s),(v)} M &> (w_1^{(s)} - p_2^{(s)} - w_K^{(s)})(t_1^{(s)} - t_1^{(v)}) + (w_2^{(s)} - w_K^{(s)})(t_2^{(s)} - t_2^{(v)}) \\
&\quad + (w_1^{(s)} + w_2^{(s)} - w_K^{(s)})(t_J^{(s)} - t_J^{(v)}) + (C_K^{(s)} - C_K^{(v)})
\end{aligned}$$

$$
\begin{aligned}
x_K^{(s),(s1)} + x_K^{(s1),(v)} &\leq 1 + x_K^{(s),(v)} \\
(1 - x_K^{(s),(v)})M &\geq (w_1^{(v)} - p_2^{(v)} - w_K^{(v)})(t_1^{(v)} - t_1^{(s)}) + (w_2^{(v)} - w_K^{(v)})(t_2^{(v)} - t_2^{(s)}) \\
&\quad + (w_1^{(v)} + w_2^{(v)} - w_K^{(v)})(t_J^{(v)} - t_J^{(s)}) + (C_K^{(v)} - C_K^{(s)})
\end{aligned}
$$

with M arbitrarily large. We have substituted the shadow prices of private leisure and childcare with their equivalent expressions from Proposition 2.

The first two conditions implement the Lindahl-Bowen-Samuelson condition for the optimal provision of public goods (joint leisure and consumption, respectively). The third condition is the childcare identity (2). Conditions 4 through 6 impose GARP on commodities (l_1, l_J, C_P) of spouse 1 in bundles s and v. The opening condition states that spouse 1 'revealed prefers' s over v ($x_1^{(s),(v)} = 1$) if $(w_1^{(s)} - p_2^{(s)})(l_1^{(s)} - l_1^{(v)}) + \delta_{1,l_J}^{(s)}(l_J^{(s)} - l_J^{(v)}) + \delta_{1,C_P}^{(s)}(C_P^{(s)} - C_P^{(v)}) \geq 0$. Bundle v was affordable given the prices and personalized budget at observation s, but spouse 1 chose s. The middle condition imposes transitivity on revealed preferences. The closing condition requires that if spouse 1 'revealed prefers' s over v ($x_1^{(s),(v)} = 1$) then bundle v must be less expensive than s at prices v. Otherwise, it was not rational to choose v. The other conditions are analogous; they impose GARP on spouse 2's commodities (l_2, l_J, C_P) and child inputs (t_1, t_2, t_J, C_K) respectively.

Linearity. Binary variables $x^{(s),(v)}$ transform the `if/then` GARP condition of Definition 1 into inequalities that are linear in $x^{(s),(v)}$, the δ's, and the t's. Of course, linearity in childcare times t is only possible because we have not kept wage premium p_m and childcare price w_K as unknowns. As we explain in section 4.4, we try values for p_m and w_K over a grid such that $p_m = \{0;\ 0.25;\ 0.5\} \times m$'s regular wage and $w_K = \{0;\ 0.33\} \times$ household's lowest wage. If, for a given p_m and w_K, the program has a solution, then the data are consistent with collective rationality with togetherness. We can search for a solution using standard software, such as `intlinprog` on Matlab. If there are multiple (p_m, w_K) tuples for which the program has a solution, we select the lowest possible prices.

Recovery of joint childcare. To recover bounds on t_J, we minimize and maximize (lower and upper bounds respectively) an appropriate function of t_J subject to all aforementioned conditions. Technically, this is the same mixed integer linear programming problem wrapped around an optimization routine. The appropriate function of t_J is $\alpha = (\sum_v 1)^{-1} \sum_v \alpha^{(v)}$, where $\alpha^{(v)} = t_J^{(v)} / \min\{T_1^{(v)}, T_2^{(v)}\}$ is the fraction that t_J at prices v is of its theoretical maximum value. This has the advantage of incorporating information on the theoretical maximum value of joint childcare (see appendix C.3.1).

Recovery of sharing rule. To recover bounds on the sharing rule, we minimize and maximize $\eta_m / \sum_m \eta_m$ subject to all aforementioned conditions. This is the same mixed integer linear programming problem, again wrapped around an optimization routine.

Measurement error. To address measurement error in individual total childcare without losing linearity, we rewrite the childcare identity (2) as $t_m^{(v)} + t_J^{(v)} = (1 + \varepsilon_m^{(v)}) \times T_m^{(v)}$. Measurement error $\varepsilon_m^{(v)}$ captures the relative deviation of observed individual total childcare $T_m^{(v)}$ from its true level $\widetilde{T}_m^{(v)} = (1 + \varepsilon_m^{(v)}) \times T_m^{(v)}$.

D Extensions

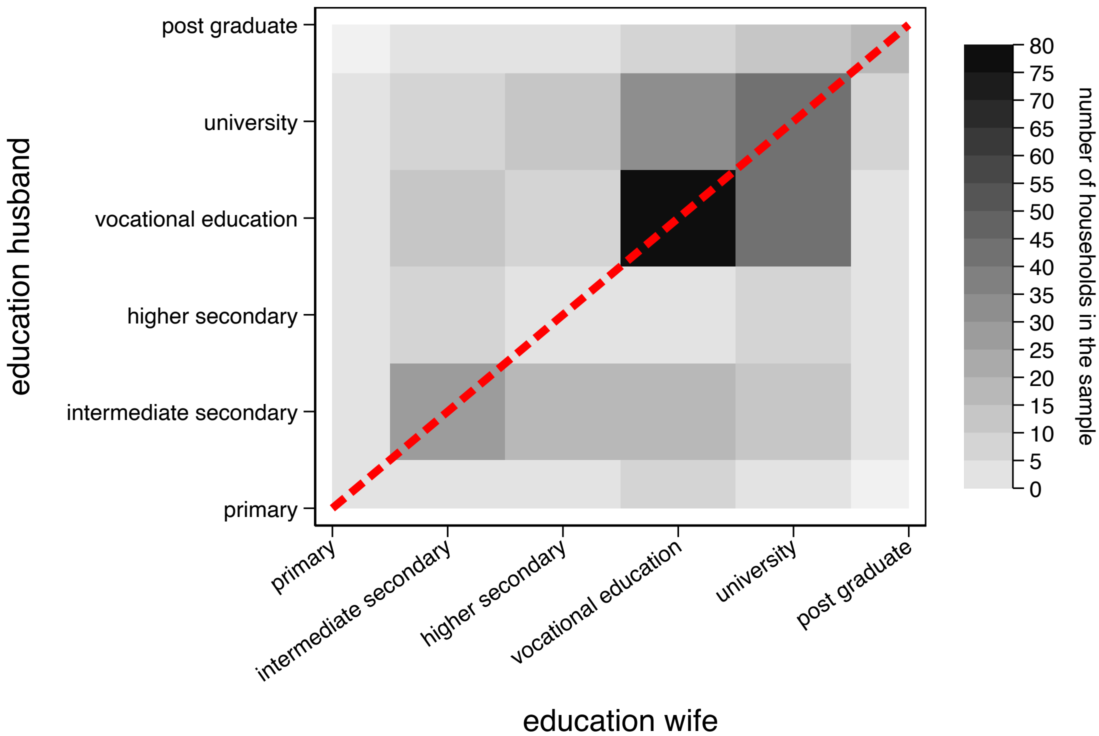
Table D.1 – Extensions: fit and value of togetherness

		Children and assortative matching scenarios				
	Baseline	(1) 1-2 children	(2) 3+ children	(3) Child-less	(4) Positive matching	(5) Negative matching
Testing (pass rates, power, predictive success)						
pass CR	0.25	0.06	0.61	0.67	0.39	0.22
pass T–CR	0.67	0.5	0.89	0.71	0.83	0.44
pass # groups	24	9	16	17	15	8
pass # HHs	250	137	94	233	133	99
power CR	0.82	0.97	0.49	0.62	0.71	0.91
power T–CR	0.62	0.87	0.26	0.59	0.44	0.73
pred. succ. CR	0.07	0.02	0.10	0.29	0.10	0.13
pred. succ. T–CR	0.29	0.37	0.15	0.30	0.27	0.18
Value of joint leisure $\sum_m \delta_{m,l_J} - \sum_m \delta_{m,l_m}$						
first quartile	0	0	0	0	0	0
median	0	0	0	0	0	0
mean	1.22	0.90	0.75	0.14	0.37	0.90
third quartile	2.60	2.18	1.52	0	0	2.17
max	9.46	6.34	8.45	6.25	6.34	5.90
Value of joint childcare $\delta_{K,t_J} - \sum_m \delta_{K,t_m}$						
first quartile	0	2.78	0	0	0	0
median	2.44	3.30	0	0	2.22	0
mean	2.08	3.59	0.97	0.14	1.88	1.83
third quartile	3.30	4.09	1.94	0	3.43	3.23
max	9.46	10.37	8.45	6.25	6.34	9.57

Notes: The table reports pass rates, power, and predictive success of collective rationality (CR) and collective rationality with togetherness (T–CR), and the distribution of the value of joint leisure and joint childcare among households consistent with T–CR. We introduce the following scenarios: (1) sample families with 1-2 children (288 households; 18 groups); (2) sample families with 3+ children (110; 18); (3) a sample of childless families (346; 24) briefly presented in appendix A.2; (4) sample families in which the partners have the same education (166; 18); (5) sample families in which the partners differ in the level of schooling (232; 18). These samples have fewer observations than the baseline so we introduce a smaller number of household groups. Households in scenarios (1)-(2) and (4)-(5) are assigned to 18 homogeneous groups, with the assignment determined by calendar year (for 2009, 2010, 2012), age of youngest child (for 0-5, 6-12), average age of parents (for 25-34, 35-44, 45-60), but not parental education. Households in scenario (3) are assigned to 24 homogeneous groups according to calendar year (for 2009, 2010, 2012), average age of the partners (for 25-35, 36-51, 52-55, 56-60), and education (2 values, cutoff at intermediate vocational education). Unlike the baseline assignment, we have introduced here a fourth bracket for the age of parents given that our sample size in this case (346 households) justifies a finer split. In all cases our assignment to groups is such that the resulting groups have as equal size between them as possible.

Figure D.1 – Assortative matching: occupational sector and education

(a) Occupational sector heatmap



(b) Education heatmap

Notes: The figure plots occupational sector and education husband-wife heatmaps. The heatmaps show the number of sample households in each cell. The red dotted diagonal indicates the cells in which positive matching occurs.

References

Afriat, S. N. (1967). The construction of utility functions from expenditure data. *International Economic Review 8*, 67–77.

Apps, P. and R. Rees (1988). Taxation and the household. *Journal of Public Economics 35*, 355–369.

Barnet-Verzat, C., A. Pailhé, and A. Solaz (2011). Spending time together: the impact of children on couples' leisure synchronization. *Review of Economics of the Household 9*, 465–486.

Beatty, T. K. M. and I. A. Crawford (2011). How demanding is the revealed preference approach to demand. *American Economic Review 101*, 2782–2795.

Becker, G. S. (1962). Irrational behavior and economic theory. *Journal of Political Economy 70*, 1–13.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal 75*, 493–517.

Becker, G. S. (1973). A theory of marriage: Part 1. *Journal of Political Economy 81*, 813–846.

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

Blow, L., M. Browning, and I. Crawford (2020). Nonparametric analysis of time-inconsistent preferences. *Review of Economic Studies, forthcoming*.

Blundell, R., P.-A. Chiappori, and C. Meghir (2005). Collective labor supply with children. *Journal of Political Economy 113*, 1277–1306.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2016). Consumption inequality and family labor supply. *American Economic Review 106*, 387–435.

Bourguignon, F., M. Browning, and P.-A. Chiappori (2009). Efficient intra-household allocations and distribution factors: implications and identification. *Review of Economic Studies 76*, 503–528.

Bronars, S. G. (1987). The power of nonparametric tests of preference maximization. *Econometrica 55*, 693–698.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments in nature and design*, Volume 2. Cambridge, MA: Harvard University Press.

Bronson, M. A. (2014). Degrees are forever: Marriage, educational investment, and lifecycle labor decisions of men and women. Unpublished manuscript.

Brown, K., L. Bradley, H. Lingard, K. Townsend, and S. Ling (2011). Labouring for leisure: Achieving work-life balance through compressed working weeks. *Annals of Leisure Research 14*, 43–59.

Browning, M., O. Donni, and M. Gørtz (2021). Do you have time to take a walk together? Private and joint time within the household. *The Economic Journal 131*(635), 1051–1080.

Browning, M. and M. Gørtz (2012). Spending time and money within the household. *The Scandinavian Journal of Economics 114*, 681–704.

Bryan, M. L. and A. Sevilla (2017). Flexible working in the UK and its impact on couples' time coordination. *Review of Economics of the Household 15*, 1415–1437.

Cano, T., F. Perales, and J. Baxter (2019). A matter of time: Father involvement and child cognitive outcomes. *Journal of Marriage and Family 81*(1), 164–184.

CentERdata (2012). LISS Longitudinal Internet Studies for the Social sciences. https://www.lissdata.nl/Home (accessed February 21, 2019).

Cherchye, L., B. De Rock, and V. Platino (2013). Private versus public consumption within groups: Testing the nature of goods from aggregate data. *Economic Theory 54*(3), 485–500.

Cherchye, L., B. De Rock, and F. Vermeulen (2007). The collective model of household consumption: A nonparametric characterization. *Econometrica 75*, 553–574.

Cherchye, L., B. De Rock, and F. Vermeulen (2009). Opening the black box of intrahousehold decision making: Theory and nonparametric empirical tests of general collective consumption models. *Journal of Political Economy 117*, 1074–1104.

Cherchye, L., B. De Rock, and F. Vermeulen (2011). The revealed preference approach to collective consumption behavior: Testing and sharing rule recovery. *Review of Economic Studies 78*, 176–198.

Cherchye, L., B. De Rock, and F. Vermeulen (2012). Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information. *American Economic Review 102*, 3377–3405.

Cherchye, L. and F. Vermeulen (2008). Nonparametric analysis of household labor supply: Goodness of fit and power of the unitary and the collective model. *The Review of Economics and Statistics 90*, 267–274.

Chiappori, P.-A. (1988). Rational household labor supply. *Econometrica 56*, 63–90.

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of Political Economy 100*, 437–467.

Chiappori, P.-A. (1997). Introducing household production in collective models of labor supply. *Journal of Political Economy 105*, 191–209.

Chiappori, P.-A. and I. Ekeland (2009). The microeconomics of efficient group behavior: Identification. *Econometrica 77*, 763–799.

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

Chiappori, P.-A. and C. Meghir (2014). Intra-household welfare. Working Paper 20189, National Bureau of Economic Research.

Chiappori, P.-A., B. Salanié, and Y. Weiss (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review 107*(8), 2109–67.

Cosaert, S. (2018). Revealed preferences for diamond goods. *American Economic Journal: Microeconomics 10*, 83–117.

Cosaert, S. and V. Hennebel (2019). Parental childcare with process benefits. Unpublished manuscript.

Cousins, C. and N. Tang (2004). Working time and work and family conflict in the Netherlands, Sweden and the UK. *Work, Employment and Society 18*, 531–549.

Craig, L. and A. Powell (2012). Dual-earner parents' work-family time: the effects of atypical work patterns and non-parental childcare. *Journal of Population Research 29*, 229–247.

Crouter, A. C., M. R. Head, S. M. Mchale, and C. J. Tucker (2004). Family time and the psychosocial adjustment of adolescent siblings and their parents. *Journal of Marriage and Family 66*, 147–162.

Cubas, G., C. Juhn, and P. Silos (2019). Coordinated work schedules and the gender wage gap. Working Paper 26548, National Bureau of Economic Research.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*, 883–931.

Del Boca, D., C. Flinn, and M. Wiswall (2014). Household choices and child development. *Review of Economic Studies 81*, 137–185.

Del Boca, D., C. J. Flinn, E. Verriest, and M. J. Wiswall (2019). Actors in the child development process. Working Paper 25596, National Bureau of Economic Research.

Demuynck, T. and C. Seel (2018). Revealed preference with limited consideration. *American Economic Journal: Microeconomics 10*(1), 102–131.

Diewert, W. E. (1973). Afriat and revealed preference theory. *Review of Economic Studies 40*, 419–425.

Donni, O. and P.-A. Chiappori (2011). *Nonunitary Models of Household Behavior: A Survey of the Literature*, pp. 1–40. New York, NY: Springer New York.

Dunbar, G. R., A. Lewbel, and K. Pendakur (2013). Children's resources in collective households: Identification, estimation, and an application to child poverty in Malawi. *American Economic Review 103*, 438–471.

Dutch Belastingdienst (2020). Tax partners. https://www.belastingdienst.nl/wps/wcm/connect/bldcontenten/belastingdienst/individuals/other_subjects/tax_partnership/tax_partners/ (accessed January 2020).

Dutch Coalition for Community Schools (2013). Kinderopvangtoeslagtabel (childcare allowance table) 2007-2012. http://bredeschool-ikc.nl/fileadmin/PDF/2013/2013-06-24__CBS_-_Cijfers_Kinderopvangtoeslag.pdf (accessed March 2019).

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

Famulari, M. (1995). A household-based, nonparametric test of demand theory. *The Review of Economics and Statistics 77*, 372–382.

Flood, S., A. Meier, and K. Musick (2020). Reassessing parents' leisure quality with direct measures of well-being: Do children detract from parents' down time? *Journal of Marriage and Family 82*(4), 1326–1339.

Folbre, N., J. Yoon, K. Finnoff, and A. S. Fuligni (2005). By what measure: Family time devoted to children in the united states. *Demography 42*, 373–390.

Fong, Y.-F. and J. Zhang (2001). The identification of unobservable independent and spousal leisure. *Journal of Political Economy 109*, 191–202.

Golden, L. (2015). Irregular work scheduling and its consequences. Briefing Paper 394, Economic Policy Institute, Washington, DC.

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

Goldin, C. (2020). Journey across a century of women. https://www.nber.org/reporter/2020number3/journey-across-century-women/ (accessed March 2021).

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

Hallberg, D. (2003). Synchronous leisure, jointness and household labor supply. *Labour Economics 10*, 185–203.

Hamermesh, D. (1998). When we work. *American Economic Review P&P 88*, 321–325.

Hamermesh, D. (1999). The timing of work over time. *Economic Journal 109*, 37–66.

Hamermesh, D. (2000). Togetherness: spouses' synchronous leisure, and the impact of children. Working Paper 7455, National Bureau of Economic Research.

Hamermesh, D. and J. Biddle (2018). Taking time use seriously: Income, wages and price discrimination. Working Paper 25308, National Bureau of Economic Research.

Hamermesh, D. and J. Lee (2007). Stressed out on four continents: time crunch or yuppie kvetch? *The Review of Economics and Statistics 89*, 374–383.

Hamermesh, D. and E. Stancanelli (2015). Long workweeks and strange hours. *ILR Review 68*, 1007–1018.

Henly, J. R. and S. J. Lambert (2014). Unpredictable work timing in retail jobs: Implications for employee work-life conflict. *ILR Review 67*(3), 986–1016.

Kok, L., C. Koopmans, C. Berden, and R. Dosker (2011). De waarde van kinderopvang. http://www.seo.nl/uploads/media/2011-33_De_waarde_van_kinderopvang.pdf (accessed March 2019).

Kuiper, N., J. de la Croix, A. Machiels, and A. Houtkoop (2014). Toeslagen bijzondere uren in cao's in 2013. http://cao.minszw.nl/pdf/174/2014/174_2014_13_234500.pdf (accessed March 2019).

Le Forner, H. (2021). Formation of children's cognitive and socio-Emotional skills: Are all parental times equal? Unpublished manuscript.

Lise, J. and K. Yamada (2019). Household sharing and commitment: Evidence from panel data on individual expenditures and time use. *The Review of Economic Studies 86*(5), 2184–2219.

Mas, A. and A. Pallais (2017). Valuing alternative work arrangements. *American Economic Review 107*, 3722–3759.

Milkie, M. A., K. M. Nomaguchi, and K. E. Denny (2015). Does the amount of time mothers spend with children or adolescents matter? *Journal of Marriage and Family 77*, 355–372.

Offer, S. (2013). Family time activities and adolescents' emotional well-being. *Journal of Marriage and Family 75*, 26–41.

Olivetti, C. and B. Petrongolo (2016). The evolution of gender gaps in industrialized countries. *Annual Review of Economics 8*, 405–434.

Presser, H. B. (1994). Employment schedules among dual-earner spouses and the division of household labor by gender. *American Sociological Review 59*, 348–364.

Qi, L., H. Li, and L. Liu (2017). A note on chinese couples' time synchronization. *Review of Economics of the Household 15*, 1–14.

Ruppanner, L. and D. Maume (2016). Shorter work hours and work-to-family interference. *Social Forces 95*, 693—-720.

Selten, R. (1991). Properties of a measure of predictive success. *Mathematical Social Sciences 21*, 153–167.

Stueve, J. L. and J. H. Pleck (2001). 'Parenting voices': Solo parent identity and co-parent identities in married parents' narratives of meaningful parenting experiences. *Journal of Social and Personal Relationships 18*(5), 691–708.

Sullivan, O. (1996). Time co-ordination, the domestic division of labour and affective relations: Time use and the enjoyment of activities within couples. *Sociology 30*, 79–100.

van Klaveren, C. and H. M. van den Brink (2007). Intra-household work time synchronization. *Social Indicators Research 84*, 39–52.

Varian, H. R. (1982). The nonparametric approach to demand analysis. *Econometrica 50*, 945–974.

Voorpostel, M., T. van der Lippe, and J. Gershuny (2010). Spending time together: Changes over four decades in leisure time spent with a spouse. *Journal of Leisure Research 42*, 243–265.