UTMD Working Paper
The University of Tokyo
Market Design Center

UTMD-118

Measuring Spousal Cooperation:

A Biform Game Approach to Intra-Household Allocation

Takahiro Moriya
The University of Tokyo

February 17, 2026

Measuring Spousal Cooperation:
A Biform Game Approach to Intra-Household Allocation*

Takahiro Moriya†
The University of Tokyo

February 17, 2026

Abstract

Do households cooperate or not? Existing models force a binary choice between full cooperation and non-cooperation, yet reality likely lies in between. I develop an empirical biform game framework that identifies the *degree* of spousal cooperation from data. The model treats enforceable decisions (labor supply) cooperatively and non-enforceable decisions (childcare) non-cooperatively, with a caring parameter nesting both extremes. Equilibrium uniqueness and monotonicity properties deliver a novel identification strategy: the cooperation parameter is identified from slope moments—how childcare responds to wages and labor supply—rather than level moments. Applying the framework to households with disabled children, I find 33% lower cooperation, reduced paternal childcare efficiency, and shifted happiness benchmarks. Counterfactual analysis shows that even large Child SSI increases cannot realistically close the welfare gap. This highlights the fundamental limits of cash transfers and the importance of policies supporting spousal cooperation.

Keywords: Spousal cooperation, Biform game, Intra-household allocation, Structural estimation, Disabled children, Subjective well-being
JEL Classification: D13, J13, J22, I12

1 Introduction

A central question in family economics is whether households behave cooperatively or non-cooperatively. The answer has first-order consequences for welfare analysis, policy evaluation, and understanding intra-

*I am grateful to Steven Stern for invaluable guidance and support. I thank Kenzo Imamura, Anning Xie and participants of seminars at Stony Brook (2024) for helpful comments. This work was supported by JST, ERATO Grant Number JPMJER2301. All errors are mine.

†Email: takahiro.moriya@e.u-tokyo.ac.jp.

household inequality. Yet existing models force a binary choice. The collective model (e.g., Chiappori, 1992) assumes Pareto efficiency for all household decisions, while non-cooperative models (e.g., Del Boca and Flinn, 2012; Flinn et al., 2018) assume simultaneous Nash play throughout. Reality likely lies between these extremes, but the literature has lacked a framework that can identify the *degree* of spousal cooperation from data.

This paper develops an empirical biform game framework that nests cooperative and non-cooperative behavior within a single model. The key insight is that household decisions differ in their enforceability: labor supply, governed by employment contracts, involves credible commitment, while daily childcare allocation is difficult to monitor and enforce. The biform structure—cooperative at Stage 1 (labor supply), non-cooperative at Stage 2 (childcare)—captures this asymmetry naturally. A caring parameter $\lambda \in [0, 1]$ governs the degree of spousal altruism in Stage 2 (childcare allocation), with $\lambda = 0$ corresponding to purely egoistic play and $\lambda = 1$ to fully internalized spousal welfare. The main methodological contribution is proving equilibrium uniqueness and establishing monotonicity properties of equilibrium outcomes under standard functional form assumptions. These results resolve the multiple equilibria problem that has limited non-cooperative models, while maintaining weaker assumptions than the collective model. Crucially, because λ is estimable, the degree of spousal cooperation becomes falsifiable—a dimension that the prior literature, by assuming the game structure a priori, could not test. The monotonicity properties also deliver a novel identification strategy: the cooperation parameter λ can be identified from *slope* moments—specifically, how the childcare-labor slope and cross-wage effects vary with λ—rather than relying solely on level moments.

I apply this framework to a setting where variation in cooperation is both substantively important and empirically salient: households with disabled children. These households present a striking puzzle. Parents of disabled children report nearly identical life satisfaction to other parents—a Kolmogorov-Smirnov test fails to reject equality ($p = 0.47$). Yet their behavior differs markedly: fathers spend about 3 fewer hours per week with their children, mothers spend about 2.6 more hours, and mothers are nearly 5 percentage points less likely to be employed. Three channels could explain this puzzle—reduced childcare efficiency, lower spousal cooperation, or shifted happiness perception—and they are observationally intertwined. Therefore, disabled-child households thus provide an ideal testing ground for the biform game framework, which can separate these channels through the structure of the model and the slope-based identification strategy.

Beyond this application, the framework applies generally to any household setting where formal and informal decisions coexist, including elderly care allocation among siblings, division of household chores, and resource sharing in multi-generational households.

The identification strategy proceeds as follows. The cooperation parameter cannot be directly observed;

it must be inferred from behavior through the lens of economic theory. The monotonicity properties established in the theoretical analysis guarantee that increases in λ generate predictable directional changes in the slope of childcare time with respect to labor supply and in cross-wage effects. Combined with subjective well-being data that identify the happiness perception channel, all three channels can be separately identified using publicly available datasets (SIPP, PSID, PSID-CDS). Because households with disabled children constitute a small fraction of the population, the analysis combines these datasets and uses an indirect inference approach that matches key moments from the data to those simulated from the model.

Estimation reveals the following. Households with disabled children exhibit a cooperation parameter that is approximately 33 percent lower than households without disabled children—placing them close to the fully non-cooperative benchmark studied in Gobbi (2018) and Del Boca and Flinn (2012). The father's childcare efficiency parameter is substantially lower for households with disabled children. The analysis also finds structural evidence of response shift in subjective well-being: households with disabled children report equivalent satisfaction at lower underlying utility levels, indicating that these households have recalibrated their happiness benchmarks. These results demonstrate that the biform game framework can successfully disentangle channels that are observationally equivalent under conventional approaches.

Counterfactual analysis yields important policy implications. Under the structural model, the welfare gap between households with and without disabled children cannot realistically be closed through Child SSI benefit increases alone—the required increase would be economically infeasible. Even fully restoring fathers' childcare efficiency to the non-disabled level closes only about one-third of the welfare gap. In contrast, equalizing the cooperation parameter closes approximately 76%. These findings underscore a key insight: spousal cooperation matters more for household welfare than either cash transfers or caregiving technology. Although the cooperation parameter is difficult to influence directly, the results suggest that policies supporting spousal coordination—such as respite care programs and family support services—may play an important role alongside income support.

The remainder of this paper proceeds as follows. Section 2 documents motivating facts. Section 3 develops the theoretical model. Sections 4–7 describe the data, identification strategy, econometric specification, and estimation. Section 8 reports results, and Section 9 presents counterfactual policy analyses.

2 Motivating Facts

2.1 The Happiness Puzzle

Figure 1 presents the distribution of life satisfaction scores from the PSID 2009, 2011, 2013, and 2015, comparing households with and without disabled children. The survey measures life satisfaction on a 5-point scale, where 1 indicates "Completely Satisfied" and 5 indicates "Not at all Satisfied." For estimation purposes, the scale is reverse-coded so that higher values indicate higher satisfaction (i.e., 5 = "Completely Satisfied"); all threshold estimates reported below use this reverse-coded scale.

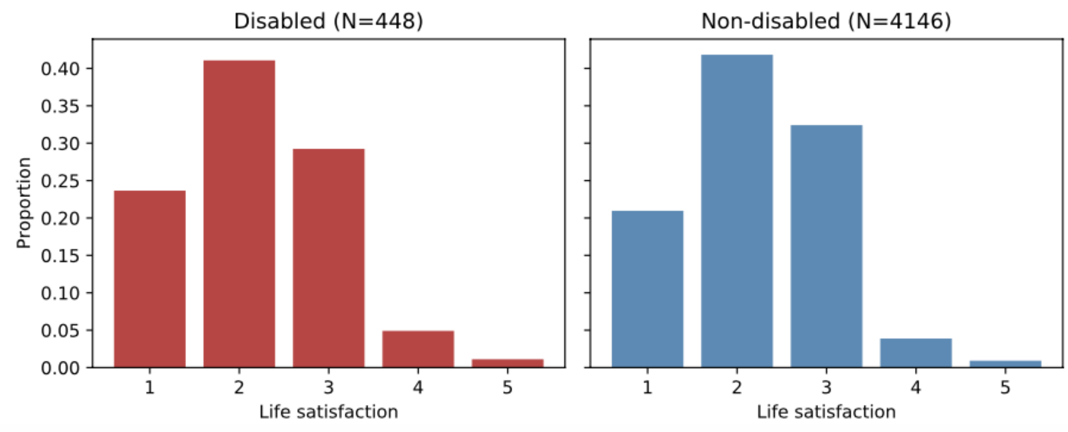

Figure 1: Distribution of life satisfaction by disability status. Data from PSID. Sample restricted to one-child households with spouse/partner present. Disability is defined broadly to include special education receipt and activity limitations.

The distributions are statistically indistinguishable. A Kolmogorov-Smirnov test fails to reject equality ($p = 0.47$), and the modal response for both groups is "Very Satisfied" (category 2), at approximately 41%. If anything, parents of disabled children report slightly *higher* rates of complete satisfaction (23.66% vs. 20.96%).

2.2 Behavioral Differences Despite Similar Happiness

The similarity in reported happiness is puzzling because observable behaviors differ substantially. Table 1 summarizes key differences in parental time allocation and labor supply.

Conditional on working, parents of disabled children work similar hours (fathers: 44.6 vs. 44.5; mothers: 37.8 vs. 37.9). However, time allocation with children differs markedly: fathers spend 3.2 fewer weekly hours with disabled children (15.6 vs. 18.8), while mothers spend 2.6 more hours (20.8 vs. 18.2). The share of non-working mothers is about 5 percentage points higher in the disabled group (34.8% vs. 30.0%).

The puzzle: How can reported happiness be nearly identical when behavior differs so substantially? The structural model developed below provides a framework for decomposing this puzzle into its underlying channels.

Table 1: Parental Time Allocation and Labor Supply by Child Disability Status

	Disabled Child		Non-Disabled Child	
Variable	Mean	(S.E.)	Mean	(S.E.)
Panel A: Labor Supply (SIPP, workers only)				
Father's weekly work hours	44.59	(0.50)	44.49	(0.13)
Mother's weekly work hours	37.75	(0.55)	37.89	(0.14)
Mother not employed (%)	34.8		30.0	
Panel B: Time with Child (PSID-CDS, weekly hours)				
Father's weekly hours with child	15.63	(0.83)	18.79	(0.57)
Mother's weekly hours with child	20.78	(0.78)	18.16	(0.45)
N (Panel A)	1,975		18,268	
N (Panel B)	399		879	

Notes: Panel A: Weekly work hours conditional on positive hours; from pooled SIPP 2004 and 2008 (household-wave observations). Panel B: Weekly hours with child from time diaries; pooled PSID-CDS 2007 and 2014, restricted to one-child households (the full PSID-CDS sample in Appendix I includes multi-child households).

3 Model

This section develops a static model of intra-household allocation in which a husband and wife jointly determine labor supply and choose childcare time independently. The model incorporates two key features central to understanding the welfare of households with disabled children: (i) a childcare efficiency parameter that captures how effectively each parent's time translates into child quality, which may differ depending on whether the child has a disability, and (ii) a caring parameter that captures the degree of spousal altruism in the non-cooperative childcare stage through interdependent utility.

The model builds on the household literature but makes a distinct theoretical contribution. Existing approaches span two extremes. Fully cooperative collective models (e.g., Chiappori, 1992) assume Pareto efficiency for all decisions, while fully non-cooperative approaches (e.g., Del Boca and Flinn, 2012; Flinn et al., 2018) model all choices as simultaneous Nash equilibria. Neither extreme captures a key institutional feature: household decisions differ in their enforceability. Labor supply, governed by employment contracts, involves credible commitment; childcare allocation, occurring daily without monitoring, does not.

The empirical model in this paper adopts a biform game structure (Brandenburger and Stuart, 2007; d'Aspremont and Jacquemin, 1988) similar to Gobbi (2018), treating labor supply as cooperative (Stage 1) and childcare allocation as non-cooperative (Stage 2). The model incorporates the caring parameter from Friedberg and Stern (2014), enabling the framework to nest cooperative and non-cooperative behavior as special cases. The key methodological contribution is establishing equilibrium uniqueness and proving monotonicity properties of equilibrium outcomes under standard functional forms—properties not established in prior work. These results resolve the multiple equilibria problem that plagues non-cooperative models while maintaining weaker assumptions than collective models (Table 2). The monotonicity properties also prove useful for identification, as discussed in Section 5.

Beyond tractability, the biform structure offers a deeper conceptual advantage. Existing household models impose the degree of spousal commitment a priori—full cooperation or full non-cooperation—and this maintained assumption is not itself testable. By introducing a continuously valued caring parameter (defined in Section 3.1) that governs altruism in the non-cooperative childcare stage, the framework nests both extremes as special cases and makes the degree of spousal cooperation estimable and falsifiable—a dimension of household behavior that the prior literature could not test.

3.1 Environment

Players and Household Structure. I consider a game played by two players within a household, indexed by $i \in \{h, w\}$, representing the husband and wife, respectively. Let j denote the household index. I confine attention to households with a single child.[1] Let $z_j \in \{0, 1\}$ be an indicator for whether household j has a disabled child: $z_j = 1$ if the household has a disabled child, and $z_j = 0$ otherwise.

Time Allocation. Each individual i in household j allocates a total time endowment T (set to 112 hours per week, following Del Boca and Flinn 2012) among three activities: work l_{ij}^w, leisure l_{ij}^ℓ, and childcare l_{ij}^c. The time constraint is given by

$$l_{ij}^w + l_{ij}^\ell + l_{ij}^c = T. \tag{1}$$

For notational convenience, I suppress the household subscript j hereafter and simply write l_i^w, l_i^ℓ, and l_i^c. Let $d_j \in \{0, 1\}$ indicate whether household j receives Child SSI (Supplemental Security Income).

Budget Constraint. Let w_i denote individual i's wage and y_i denote i's non-labor income; both are assumed exogenously given and observable from data. I define household consumption x as,

$$x = w_h l_h^w + w_w l_w^w + y_h + y_w + \text{Child SSI}_j \cdot \mathbb{I}[z_j = 1, d_j = 1]. \tag{2}$$

I assume there is no saving, so consumption equals total income within the period. This is justified by the fact that the SIPP and PSID samples used in this study primarily consist of low- to middle-income

[1] As noted in Sauer and Taber (2021), the number of children matters for household behavior. Previous literature, such as Gobbi (2018) and Del Boca et al. (2014), focuses on multiple children. Here, even at the cost of losing some observations, I ignore this point. I focus on households with one child for simplicity.

households, for whom savings rates are low.[2] Wages are determined by wage equations specified in Section 6.

Child Quality Production. Let t denote the age of the child. Child quality k is produced according to the following Cobb–Douglas production function,

$$k = \left(g_h(z_j : t)l_h^c + 1\right)^{0.5} \cdot \left(g_w(z_j : t)l_w^c + 1\right)^{0.5}. \tag{3}$$

Conceptually, k is a household production good that enters both spouses' utility as a public good within the household. Following the child quality literature (e.g., Del Boca and Flinn, 2012; Del Boca et al., 2014), child quality is determined by both parents' time inputs, l_h^c and l_w^c, and the functional form is chosen to derive analytical solutions.

The term $g_i(z_j : t) > 0$ for $i \in \{h, w\}$ captures the efficiency of childcare time—reflecting factors such as conversion efficiency, communication quality, and technology of care. This efficiency may differ between households with and without disabled children. For notational simplicity, I sometimes write $g_i(z_j)$ when the dependence on child age t is clear from context.

I model $g_i(z_j : t)$ as

$$g_i(z_j : t) = \exp(\beta_z^i z_j + \beta_t t), \tag{4}$$

where β_t is symmetric across spouses but β_z^i depends on i. This exponential specification ensures positivity and is used in Del Boca et al. (2014) and Flinn et al. (2018).

Several remarks on the production function are in order. First, while more general CES specifications are possible, the Cobb–Douglas form is restrictive but necessary for identification. As Del Boca and Flinn (2005, 2012, 2014) and Flinn et al. (2018) discuss, I observe four optimal behaviors (labor supply and childcare time for each spouse). So the number of unknown preference parameters which I can back up from four optimal behaviors is at most four ($\eta_h, \mu_h, \eta_w, \mu_w$), which I will introduce below. A CES specification would introduce an additional parameter that cannot be uniquely backed up without restrictions.[3]

Second, the 0.5 Cobb–Douglas specification is empirically supported by Del Boca et al. (2014) and used in Gobbi (2018). As shown in Section 8, my in-sample fit is very good with these restrictions. Importantly, Theorems 1, 2, and 3 below are robust to the specific value of 0.5; what matters is the closed form solutions from the Cobb–Douglas structure itself, which is used in nearly all structural estimation

[2]The determination of Child SSI in equation (2) is discussed in detail in Section 6. Briefly, properly accounting for Child SSI matters because setting it to zero would mechanically attribute all behavioral differences between households with and without disabled children to the childcare technology or the cooperation parameter, rather than to resource differences.

[3]For the same reason, the assumption that each spouse contributes equally (with exponent 0.5) to child quality may appear restrictive, but treating this as a free parameter would also preclude identification.

papers on intra-household allocation.[4]

Under this specification, the marginal return to childcare $\partial \ln k/\partial l_h^c$ is increasing in efficiency $g_h(z_j)$. Consequently, when a child's disability reduces $g_h(z_j)$, the father's optimal childcare time decreases.[5]

Third, following Gobbi (2018), I allow corner solutions: l_h^c and l_w^c can be zero. The "+1" term in equation (3) prevents the marginal product from going to infinity when inputs are zero, addressing Inada-type issues with Cobb–Douglas functions.[6]

Fourth, the specification of k implicitly assumes I do not model monetary investment in children. The implication is discussed in Del Boca et al. (2014) and Gobbi (2018). Brown et al. (2025) also do not include monetary investment.

Preferences: Felicity Utility. Let $u_i(x, l_i^\ell, k)$ for $i \in \{h, w\}$ denote the felicity utility of the husband and wife. Following Del Boca and Flinn (2012) and Del Boca et al. (2014), I specify felicity utility in Cobb–Douglas form,[7]

$$u_i(x, l_i^\ell, k) = (1 - \mu_i - \eta_i) \ln x + \mu_i \ln l_i^\ell + \eta_i \ln(k - \underline{k}_i), \tag{5}$$

where $\mu_i \in (0, 1)$ is the weight on leisure, $\eta_i \in (0, 1)$ is the weight on child quality for individual i, and $\underline{k}_i \geq 0$ is a subsistence level of child quality. Following Del Boca and Flinn (2005), I normalize $\underline{k}_i = 0$ without empirical loss of generality. The joint distributions of μ_i and η_i, denoted $F^z_{\mu,\eta}$, may depend on z_j. This is natural because how much spouses care about their child may differ depending on whether the child has a disability.[8]

The parameters η_i and $g_i(z_j)$ in equations (5) and (3) affect individual observable choices, especially labor supply and childcare time. Therefore, these observable choices can identify the means of η_i and $g_i(z_j : t)$ for $i \in \{h, w\}$. Conditional on all else being equal, the mean difference in optimal behavior

[4]The key conditions are that the production function is log-additive and that utility is linear in $\ln k$ (constant marginal utility with respect to log child quality). These assumptions are standard in the literature. By contrast, Byrne et al. (2009) assume linear production technology.

[5]This formulation is consistent with, for example, Cunha and Heckman (2007), who emphasize that parents invest more in children with higher productivity. Empirically, fathers spend more time with non-disabled children; see Section 8.

[6]That both l_h^c and l_w^c can be 0 in the functional form for k implies that l_h^c and l_w^c represent *active* parenting time: it would be dangerous if neither parent took care of their child, especially when the child is very young. See Del Boca et al. (2014) for the difference between active and passive parenting time. My leisure terms implicitly include passive childcare time. Housework time, which is also important for a household in the literature, is likewise included in leisure.

[7]As with the child quality production function, where adding a free parameter would preclude identification, using more general functional forms such as CES would preclude parameter identification also. Again, since I observe only labor supply and childcare time for each spouse—two observables per person—I cannot have more than two unknown parameters per person.

[8]This is a simplification to obtain analytical solutions. Several papers in this literature (for example, Del Boca and Flinn 2005; Del Boca and Flinn 2012; Flinn et al. 2018) do not include childcare time directly in the utility function. However, for disabled-child care, the direct utility or disutility may be substantial. Recall that any direct effect can be captured through shifts in the distribution of η_i between $z = 0$ and $z = 1$.

between households with and without disabled children identifies these parameters. The key is that I have $4 \times 2 = 8$ observable outcomes (labor supply and childcare time for each spouse, under $z = 0$ and $z = 1$). Using $z = 0$ as the baseline, I identify 4 preference parameters (the means of μ_i and η_i for $i = h, w$). The remaining degrees of freedom identify β_z^i for $i = h, w$ and allow the distribution of η_i to depend on z. The identification argument follows Del Boca and Flinn (2012, 2014).

In this static framework, child quality is better understood as a form of public consumption rather than investment. Time with children, l_i^c, is purely an input to child quality production, with no direct effect on utility, as in Del Boca et al. (2014).

Super Utility and the Caring Parameter. Let $\lambda(z_j) \in [0, 1]$ be a *caring parameter* (restricted to $(0, 1)$ in the econometric specification; see Section 6.1), which governs how much individual i internalizes $-i$'s welfare when choosing childcare in Stage 2. Let cost_j denote the participation cost for receiving Child SSI, and recall that $d_j \in \{0, 1\}$ indicates whether household j receives Child SSI.

The payoff function—corresponding to the "super utility" in Friedberg and Stern (2014)—for individual i is

$$U_i(x, l_i^\ell, k) = u_i(x, l_i^\ell, k) + \lambda(z_j) u_{-i}(x, l_{-i}^\ell, k) - \frac{\text{cost}_j}{2} \cdot \mathbb{I}[z_j = 1, d_j = 1] \tag{6}$$

for both husband and wife.[9, 10, 11] The parameter $\lambda(z_j)$ governs the degree of altruism in Stage 2 (childcare allocation). When $\lambda(z_j) = 1$, each spouse fully internalizes the other's welfare, so Stage-2 childcare choices coincide with the social planner's solution. When $\lambda(z_j) = 0$, the payoff function (6) reduces to an egoistic utility function, yielding purely non-cooperative childcare allocation as in Gobbi (2018). In both cases, Stage 1 labor supply remains cooperative. The identification of $\lambda(z_j)$ is discussed in detail in Section 5.

Participation Cost for Child SSI. The term cost_j represents the participation cost for Child SSI; the empirical specification is given in Section 6.4. As Friedrichsen et al. (2018) discuss, disentangling actual costs associated with participation for Child SSI from observable data is difficult because all types of costs (administration cost, phycological costs) moves together when a recipient takes benefits. I therefore interpret cost_j as a participation cost following the literature.[12]

Since I cannot use credible micro moments for Child SSI take-up, I normalize $\text{cost}_j/2 = 0$. This normalization is partially justified because participation costs in the literature typically represent "welfare

[9]Theoretically, this specification is restrictive and not fully general; see Chiappori and Mazzocco (2017). However, it is commonly used in the family economics literature. Practically, functional form assumptions are necessary, just as recursive utility is assumed in dynamic models.

[10]Love causes people to derive happiness from seeing their partner happy, which motivates the super utility specification.

[11]As Keane and Wolpin (2001) note, such costs can be modeled either in the budget constraint or in utility. While applying for benefits involves dynamic aspects and belief formation (Moffitt, 1983), traditionally (e.g., Blundell and Shephard, 2012) a static model is typically used in the literature.

[12]As Fang and Silverman (2009) note, this is standard practice since Moffitt (1983).

stigma" for the person taking up benefits; here, however, the child is the recipient, so it is unclear whether parents would feel such stigma.[13] The application for Child SSI is made at the household level (a joint decision). Because the participation cost enters super utility additively, it does not alter the first-order conditions for within-period allocations, so the take-up decision can be separated from the allocation problem and estimated outside the main model.[14] Child SSI itself, once received, does affect allocations through income effects on consumption (equation (2)), which is why it must be properly accounted for in estimation rather than set to zero.

My primary objective is to identify the disability-dependent primitives governing within-period allocations and the subjective well-being decomposition—specifically, the childcare technology shifters $g_i(z_j)$ for $i \in \{h, w\}$ and the caring parameter $\lambda(z_j)$. Structurally modeling the application/learning/diagnosis process for d_j is a secondary goal.[15]

Borrowing the idea of indirect inference, to prevent mechanical mismatch between simulated and observed data due to unobserved or misreported d_j, I assign d_j probabilistically in simulations using externally informed take-up rates $P(d_j = 1 \mid z_j)$, reproducing the same mixture distribution as in observed data. This reduces measurement-driven discrepancies while keeping focus on identifying $g_i(z_j)$ and $\lambda(z_j)$.

This simplification does not claim that participation costs are irrelevant; rather, it defines a baseline clarifying identification under data limitations. The framework accommodates counterfactual analyses with nonzero participation costs (e.g., changes in application frictions). To assess sensitivity to the externally imposed take-up rate, I conducted robustness checks varying $P(d_j = 1 \mid z_j = 1)$ around the benchmark value (e.g., 0.7) and verified that main results for $\lambda(z_j)$ and $g_i(z_j)$ remain stable.

3.2 Game Structure

This subsection specifies the timing of decisions and the equilibrium concept. I adopt a two-stage ("biform") game structure that delivers a unique subgame perfect equilibrium, which is essential for well-defined comparative statics and identification. Table 2 summarizes the position of this paper relative to

[13]Importantly, there is no work requirement for Child SSI receipt, so this decision can be handled outside the main model. In contrast, SSI for adults imposes work requirements on the recipient, directly affecting the recipient's (i.e., the parent's) labor supply decision and requiring joint modeling.

[14]Specifically, the participation cost enters super utility additively as a constant,

$$-\frac{\text{cost}_j}{2} \cdot \mathbb{I}[z_j = 1, d_j = 1],$$

and hence does not alter the first-order conditions (FOCs) characterizing within-period allocations. Consequently, it does not directly affect the main identifying moments for $g_i(z_j)$ and $\lambda(z_j)$ (FOC invariance).

[15]Low and Pistaferri (2010) and French and Jones (2011) use reduced-form approaches, and Low and Pistaferri (2020) emphasize the difficulty of explicitly modeling disability insurance. However, in principle, with reliable Child SSI data, cost_j could be identified: one could compare $\widehat{\text{cost}}_j \gtreqless \widehat{\text{Child SSI}}_j$, where $\widehat{\text{Child SSI}}_j$ is imputed from institutional parameters (Federal Benefit Rate by state and year) and observed income as an exogenous policy shifter (see Abrahams et al., 2025).

existing approaches. The key distinction across these approaches is whether and to what extent cooperative behavior is permitted within the model.

Table 2: Comparison of Game-Theoretic Assumptions in Household Models

Model	Assumption	Restrictiveness
Collective model (Chiappori)	Pareto efficiency in all decisions	Most restrictive
Biform game (this paper)	Partial cooperation	Intermediate
Non-cooperative	No cooperation	Least restrictive

Timing. Players are assumed to play a two-stage game within a single period. In Stage 1, each household member chooses labor supply in a cooperative manner. In Stage 2, time allocation for childcare is decided non-cooperatively.

I focus on a static model because (i) I do not model child development dynamics, and (ii) I wish to avoid multiple equilibrium issues that typically arise in dynamic games. Static models are common in this literature.[16]

Why a Two-Stage Structure? Commitment over labor supply is reasonable because employment contracts are typically signed, so employees cannot easily change their work arrangements, at least in the short term. In contrast, for childcare, partners cannot assume commitment because there is generally no written agreement and monitoring is difficult.[17] Therefore, I model Stage 2 as a non-cooperative game.[18] This justification follows Gobbi (2018).

Without dividing the game into stages, with players selecting all choices simultaneously, equilibrium multiplicity would almost certainly arise.[19] The biform structure avoids this problem by isolating childcare into a separate stage. Since this is a finite-stage game, the equilibrium concept is subgame perfection with backward induction, and existence follows immediately.

[16] Gobbi (2018), Friedberg and Stern (2014), Del Boca and Flinn (2012), Flinn et al. (2018), and Gayle and Shephard (2019) all use static models. As Chiappori et al. (2018) note, family economics typically employs static models without investment dynamics. This is appropriate for short-term analysis, though it may introduce bias for long-term outcomes. Incorporating learning, applications, human capital, and matching after diagnosis is beyond this paper's scope. More importantly, as Del Boca and Flinn (2012) and Flinn et al. (2018) emphasize, even with very restricted Cobb–Douglas functional forms in a limited static model, parameters cannot be nonparametrically identified. Adding dynamic aspects would introduce more parameters requiring additional arbitrary assumptions without restrictions.

[17] Even if couples could contract over childcare time, there would always be incentive to deviate due to lack of monitoring—a typical issue with cooperative game assumptions, often termed "lack of monitoring."

[18] I implicitly assume employment contracts precede daily childcare decisions.

[19] See Quint and Shubik (1997) for a general discussion of how the dimension of the action space relates to equilibrium multiplicity.

Equilibrium Uniqueness. Let l_h^{c*} and l_w^{c*} denote equilibrium childcare time for husband and wife. Throughout, superscript $*$ denotes equilibrium values.

Theorem 1 (Uniqueness of Stage 2 Equilibrium)**.** *Given l_h^w and l_w^w, and $\lambda(z_j)$, there is a unique Nash equilibrium (l_h^{c*}, l_w^{c*}) at Stage 2.*

The key to uniqueness is that the Cobb–Douglas child quality function is additively separable in logs, so each parent's first-order condition depends only on their own childcare choice. Stage 2 therefore reduces to a pair of independent one-dimensional problems, each strictly concave, guaranteeing a unique optimum for each player and hence a unique Nash equilibrium.

Proof of Theorem 1. See Appendix A for the full derivation. □

Remark 1. The key conditions for Theorem 1 are:

(A) The production function is Cobb–Douglas (standard in the literature):

$$\ln k = \frac{1}{2}\ln(a_h l_h^c + 1) + \frac{1}{2}\ln(a_w l_w^c + 1).$$

(B) Utility is linear in $\ln k$ (constant marginal utility with respect to $\ln k$), which is satisfied by the Cobb–Douglas specification adopted here:

$$U_i = \cdots + \eta_i \ln k \quad \Rightarrow \quad \frac{\partial U_i}{\partial \ln k} = \eta_i \text{ (constant).}$$

Remark 2. Although Stage 2 FOCs are strategically independent under the Cobb–Douglas specification, Stage 2 remains a non-cooperative game: each parent's childcare choice affects the other's payoff through child quality. Because each parent maximizes their own objective without internalizing the spillover on their spouse's utility, the Stage 2 outcome generates an efficiency loss relative to the joint optimum.

Let l_h^{w*} and l_w^{w*} denote equilibrium labor supply, and let k^* denote equilibrium child quality mapped from l_h^{c*} and l_w^{c*}.

Theorem 2 (Uniqueness of Stage 1 Equilibrium)**.** *Given $\lambda(z_j)$, equilibrium labor supply (l_h^{w*}, l_w^{w*}) is uniquely determined by*

$$\{l_h^{w*}, l_w^{w*}\} = \underset{l_h^w, l_w^w}{\operatorname{argmax}} \left[U_h(x, T - l_h^w - l_h^{c*}, k^*) + U_w(x, T - l_w^w - l_w^{c*}, k^*) \right].$$

Note that x is also a function of l_h^w and l_w^w: $x = w_h l_h^w + w_w l_w^w + y_h + y_w + \mathbb{I}[d_j = 1, z_j = 1] \cdot \text{Child SSI}_j$. The Child SSI term is exogenously given and does not affect this proof. Because players solve backwards, childcare times are equilibrium values l_h^{c*} and l_w^{c*} at this stage.

Proof of Theorem 2. Stage 1 is a cooperative game where the couple jointly decides allocations. The objective $U_h + U_w = (1 + \lambda)(u_h + u_w)$ is strictly concave in (l_h^w, l_w^w) on the (convex) constraint set: under the Cobb–Douglas specification, the Stage-2 equilibrium childcare l_i^{c*} is affine in l_i^w (see Appendix D), so substituting it into the log-linear felicity utilities preserves strict concavity. A strictly concave function on a convex set has at most one maximizer, so (l_h^{w*}, l_w^{w*}) is unique. □

Implication of Theorems 1–2. Together, these theorems guarantee that the subgame perfect equilibrium of the entire biform game is unique. This uniqueness is essential for two reasons: (i) comparative statics with respect to parameters (e.g., λ, wages) are well-defined, and (ii) the identification strategy in Section 5 relies on the monotonicity of equilibrium outcomes, which requires uniqueness.

Theorem 3 (Monotonicity of Indirect Utility in λ)**.** *Define $V(\lambda(z_j))$ at equilibrium as*

$$V(\lambda(z_j)) \;:=\; u_h(x^*, l_h^{\ell*}, k^*) + u_w(x^*, l_w^{\ell*}, k^*).$$

Then $V(\lambda(z_j))$ is weakly increasing in $\lambda(z_j)$. If, in addition, $V(\lambda(z_j)) > 0$, then the joint indirect utility $U_h + U_w = (1 + \lambda(z_j))\,V(\lambda(z_j))$ is also weakly increasing in $\lambda(z_j)$.

Proof. The proof exploits the supermodular game structure at Stage 2. The key steps are: (i) verifying increasing differences of payoffs in (l_i^c, λ), (ii) applying Tarski's fixed-point theorem together with the uniqueness result of Theorem 1, and (iii) using the fact that the max operator preserves monotonicity. The result holds under Cobb–Douglas preferences and child quality technology without requiring the specific 0.5 exponent or symmetry of λ between spouses. See Appendix B for the full proof. □

The condition $V(\lambda(z_j)) > 0$ is natural in my setting: as equation (5) shows, each argument of the felicity utility—consumption x, leisure l_i^ℓ, and child quality k—exceeds 1 under the units adopted here (weekly hours with $T = 112$, weekly dollar consumption, and child quality with the "+1" term in equation (3)), so $\ln(\cdot) > 0$ for each term, $u_i > 0$ for each spouse, and hence $V(\lambda(z_j)) > 0$.

Theorem 3 is nontrivial: while high λ might increase $U_h + U_w$ because it scales up indirect utility (shifting the whole function upward), the felicity utilities $u_h(x^*, l_h^{\ell*}, k^*)$ and $u_w(x^*, l_w^{\ell*}, k^*)$ at equilibrium could decrease when λ is larger. The direction is not obvious a priori. Note that $U_h(l_h^{w*}, l_h^{c*}) + U_w(l_w^{w*}, l_w^{c*})$ is increasing in $\lambda(z_j)$, but U_h or U_w individually need not be.

While Theorem 3 establishes that improvements in λ lead to higher indirect utility for married couples, the *quantitative* magnitude of this improvement remains unknown from theory alone. To assess the welfare implications quantitatively, I employ structural estimation (Section 7) to recover the model primitives, and then conduct counterfactual analysis (Section 9) to measure how much welfare would improve under alternative cooperation levels.

The first and second moments of μ_i and η_i can be identified using standard arguments from the intra-household allocation literature, since I observe labor supply and childcare time. The parameter $\lambda(z_j)$ is identified through the monotonicity of slopes with respect to wage changes (Proposition 4 below).

3.3 Monotonicity of Equilibrium Outcomes

Theorems 1 and 2 guarantee uniqueness of equilibrium, so derivatives are well-defined and comparative statics can be summarized as follows.

Proposition 4 (Monotone Comparative Statics with Respect to $\lambda(z_j)$)**.** *Consider the cross-wage effect evaluated at within-period equilibrium,*

$$\psi_{i,-i}(\lambda(z_j)) \equiv \frac{\partial l_i^{c*}}{\partial w_{-i}} \qquad (i \in \{h, w\}),$$

the childcare-labor marginal rate of substitution (slope) for each spouse:

$$\kappa_i(\lambda(z_j)) \equiv \frac{\partial l_i^{c*}}{\partial l_i^{w}} \qquad (i \in \{h, w\}),$$

and the equilibrium (household) indirect utility:

$$\mathcal{W}(\lambda(z_j)) \equiv U_h(l_h^{w*}, l_h^{c*}) + U_w(l_w^{w*}, l_w^{c*}).$$

Provided the equilibrium is interior and differentiable, the following hold:

1. ***Monotonicity of cross-wage effect:*** *$\psi_{i,-i}(\lambda(z_j))$ is strictly increasing in $\lambda(z_j)$ for each $i \in \{h, w\}$.*
2. ***Monotonicity of childcare-labor slope:*** *$\kappa_i(\lambda(z_j))$ is monotone in $\lambda(z_j)$ (decreasing in this model).*
3. ***Monotonicity of (household) indirect utility:*** *If $V(\lambda(z_j)) > 0$, then $\mathcal{W}(\lambda(z_j))$ is increasing in $\lambda(z_j)$.*

I emphasize that Cobb–Douglas preferences, Cobb–Douglas child quality technology, and the biform game structure are sufficient conditions for these results. Proposition 4 plays an important role in the identification strategy of Section 5.

Proof. See Appendix C. □

Monotonicity of Equilibrium Variables and Sign Reversal. By Proposition 4, these are continuous and monotone on $\lambda \in (0, 1)$,

$$\frac{\partial \psi_{i,-i}(\lambda)}{\partial \lambda} > 0, \qquad \frac{\partial \kappa_i(\lambda)}{\partial \lambda} < 0 \quad \text{for all } \lambda \in (0,1),\ i \in \{h, w\}.$$

Proposition 5 (Threshold Sign Reversal under Monotonicity)**.** *Let $f(\lambda)$ be a continuous, monotone function of λ. If there exist $\underline{\lambda}, \overline{\lambda} \in (0,1)$ such that $f(\underline{\lambda})$ and $f(\overline{\lambda})$ have opposite signs, then by the intermediate value theorem, there exists $\lambda^* \in (\underline{\lambda}, \overline{\lambda})$ with $f(\lambda^*) = 0$.*

Application to $\psi_{i,-i}$: *Since $\psi_{i,-i}(\lambda)$ is monotonically increasing for each $i \in \{h, w\}$ (Proposition 4), if $\psi_{i,-i}(\underline{\lambda}) < 0$ and $\psi_{i,-i}(\overline{\lambda}) > 0$, there exists λ^*_ψ such that*

$$\psi_{i,-i}(\lambda) < 0 \textit{ if } \lambda < \lambda^*_\psi, \qquad \psi_{i,-i}(\lambda) > 0 \textit{ if } \lambda > \lambda^*_\psi.$$

Application to κ_i: *Since $\kappa_i(\lambda)$ is monotonically decreasing (Proposition 4), if $\kappa_i(\overline{\lambda}) < 0$ and $\kappa_i(\underline{\lambda}) > 0$, there exists λ^*_κ such that*

$$\kappa_i(\lambda) > 0 \textit{ if } \lambda < \lambda^*_\kappa, \qquad \kappa_i(\lambda) < 0 \textit{ if } \lambda > \lambda^*_\kappa.$$

Implications.

Although $\psi_{i,-i}(\lambda)$ and $\kappa_i(\lambda)$ are monotone in λ, their signs can reverse when crossing zero. Therefore, the *sign* of the observed (reduced-form) slope provides information about whether λ lies above or below the threshold λ^*. Accordingly, for identifying λ, the key information comes from the *magnitude* of these slopes (measured in absolute value) rather than sign consistency. Consequently, sign inconsistency across different slope moments does not pose a problem for identification (see Section 8.3 for details).

Remark 3 (Dependence on Cobb–Douglas). Consider replacing the Cobb–Douglas child quality function with a CES specification $k = \left[\frac{1}{2}(a_h l_h^c + 1)^\rho + \frac{1}{2}(a_w l_w^c + 1)^\rho\right]^{1/\rho}$, where $\rho \to 0$ recovers Cobb–Douglas. For $\rho \le 0$ (complementary inputs), the cross-partial

$$\frac{\partial^2 \ln k}{\partial l_h^c \, \partial l_w^c} = \frac{-\rho \, a_h a_w \, (a_h l_h^c + 1)^{\rho-1}(a_w l_w^c + 1)^{\rho-1}}{\left[(a_h l_h^c + 1)^\rho + (a_w l_w^c + 1)^\rho\right]^2} \;\ge\; 0,$$

so the Stage-2 game remains supermodular. Combined with $\partial^2 \pi_i / \partial l_i^c \, \partial\lambda = \eta_{-i} \cdot \partial \ln k / \partial l_i^c > 0$ (which holds for any production function in which childcare is productive), Topkis's theorem guarantees that equilibrium childcare $l_i^{c*}(\lambda)$ is weakly increasing in λ—i.e., equilibrium *level* monotonicity is preserved.

The identification strategy in Section 5, however, relies not on the *level* of equilibrium childcare but on the monotonicity of equilibrium *slopes*—the childcare–labor tradeoff $\kappa_i(\lambda) \equiv \partial l_i^{c*} / \partial l_i^w$ and the cross-wage effect $\psi_{i,-i}(\lambda) \equiv \partial l_i^{c*} / \partial w_{-i}$—with respect to λ. These are second-order comparative statics that Topkis does not deliver; the analytical proof (Proposition 4) exploits the closed-form solutions available under Cobb–Douglas ($\rho = 0$). A natural concern is therefore whether the identification result is knife-edge—valid only at the single point $\rho = 0$. It is not. Because the equilibrium is interior and the Jacobian of the first-order conditions is non-singular at $\rho = 0$, the implicit function theorem guarantees that $l_i^{c*}(\lambda; \rho)$ is smooth in ρ in a neighborhood of $\rho = 0$. The signs of $\partial\kappa_i/\partial\lambda$ and $\partial\psi_{i,-i}/\partial\lambda$, being continuous functions of ρ, are therefore preserved whenever $|\rho| < \varepsilon$ for some $\varepsilon > 0$. In other words, the slope monotonicity that underpins identification holds not only under Cobb–Douglas but extends to CES specifications with ρ sufficiently close to zero—that is, production technologies that remain sufficiently close to Cobb–Douglas. The identification strategy is therefore not a knife-edge result tied to a single functional form, but is locally robust within this neighborhood.

3.4 Equilibrium Concept

The equilibrium concept is subgame perfect equilibrium with a biform game modification. For a given household j, at each subgame, each individual behaves rationally and has no incentive to deviate from equilibrium. From the perspective of the entire game, no player has an incentive to deviate. Existence and uniqueness of the equilibrium follow from Theorems 1 and 2.

4 Data

This section describes the data sources used for estimation. Following Del Boca and Flinn (2012) and Flinn et al. (2018), I condition on already-married couples. The key parameters to identify are the distribution of $F^{z}_{\mu,\eta}$, the childcare efficiency $g_i(z_j)$ (defined in the production function), and the cooperation parameter $\lambda(z_j)$ (in the super utility).

The approach to combining multiple data sources follows the methodology developed by Friedberg and Stern (2014), Keane and Wasi (2013) and Del Boca et al. (2026), with a twist suited to the present context.

To address this data limitation, I employ three different types of datasets to identify $g_i(z_j : t)$, $\lambda(z_j)$ and happiness thresholds separately, following a conceptually two-step estimation approach. This strategy circumvents two challenges: the scarcity of observations of households with disabled children, and the age-dependent nature of intra-household behavior (as described in Section 3, household behavior changes substantially depending on child age).

The analysis draws on three complementary data sources: the Survey of Income and Program Participation (SIPP), the Panel Study of Income Dynamics Child Development Supplement (PSID-CDS), and the PSID Main Interview. The specific waves used are PSID-CDS III (2007/08) and CDS 2014, SIPP 2004 Waves 4 and 8 (February 2005–May 2005 and June 2006–September 2006), and SIPP 2008 Panel Waves 5 and 8 (January–April 2010 and January–April 2011), along with PSID Main waves from 2009, 2011, 2013, and 2015. The pooled sample covers the period 2005–2014. By pooling these datasets, I implicitly assume that the economy is stationary within this period.[20]

[20]The primary concern is child age effects rather than calendar time effects; in family economics, spousal ages are generally

Each dataset offers distinct advantages and disadvantages. SIPP provides information on household time use and child disability status with a large sample size, and is available annually. However, SIPP lacks information on subjective well-being, which is needed to identify how the presence of a disabled child affects perceived happiness (see Section 5).Additionally, the quality of childcare time data in SIPP is relatively low. PSID-CDS provides high-quality childcare time data for each spouse, which is essential for identifying the childcare production function. However, its sample size alone is insufficient, and it does not contain subjective well-being data. PSID Main includes subjective well-being data for the responding spouse, enabling identification of happiness thresholds. However, it lacks childcare time information and is collected biennially.

An important consideration is that disability definitions differ between SIPP and PSID-CDS; the survey questions are phrased differently, though they capture the same underlying concepts. Prior literature has successfully combined multiple datasets with differing variable definitions (Goeree, 2008; Del Boca et al., 2026), suggesting this approach is feasible for the present study as well. The detailed correspondence between disability variables across datasets is documented in Table K1 in Appendix K.

Following the structural estimation approach in industrial organization, I pool these datasets and treat observations across waves as repeated cross-sectional data. This pooling strategy is standard in the IO literature; for example, Berry (1994), Berry et al. (1995), and Nevo (2000) employ similar approaches. Nevo (2001) absorbs time variation using year dummies while maintaining parameter stationarity.

Specifically, the deep structural parameters $(\mu_i, \eta_i, g_i, \lambda)$ are assumed to be time-invariant, while $g_i(z_j)$ and $\lambda(z_j)$ incorporate child age effects. Differences in wages and SSI benefit amounts across waves are absorbed into the model through the observed income variables w_i, y_i, and Child SSI_j, which are taken directly from data. The primary concern is controlling for age effects rather than trend or year effects in the literature.[21] As a robustness check, one could allow for wave-specific shifts (hyperparameters) in the specification. The rationale for pooling is twofold: to increase event counts by pooling a stationary one-shot stage game across three waves, and to address the rarity of households with disabled children.

PSID has included subjective well-being measures (life satisfaction) since 2009, though these measures are only available for interview respondents (primarily Head and Spouse/Partner). Subjective well-being is used to identify happiness thresholds. The identification approach following Friedberg and Stern (2014) and Byrne et al. (2009) does not require both spouses' happiness data simultaneously—one can run ordered response models separately for each spouse. Additionally, K6 measures serve as auxiliary statistics for

less important than child age. Even with a relatively small available sample, structural estimation remains feasible: the Del Boca et al. series works with sample sizes around 120 observations. See Appendix J for a discussion of missing data in the PSID-CDS time diary.

[21]More broadly, parameter stationarity is a common assumption in this literature. Del Boca et al. (2024) assume that the discount rate does not vary with age. Eckstein et al. (2019) assume that preferences are stable even across different cohorts—in essence, conditioning on enough observable characteristics.

indirect inference, capturing information that is not orthogonal to the indirect utility function.[22] Following Byrne et al. (2009), randomness is introduced through the econometric specification (see Section 6). Even if using non-revealed preference data is somehow controversial, it is worth noting that using non-revealed-preference data such as job satisfaction and collapsing it to a single index through regression is a well-used approach, for example, employed by French and Jones (2011).

Descriptive statistics from the SIPP and PSID-CDS samples are reported in Appendix I. A notable pattern emerges: mothers spend more time with disabled children, while fathers spend less time, suggesting asymmetric responses to child disability within the household. Detailed moment statistics used for structural estimation are presented in Appendix F.

5 Identification

The identification of the technology parameter $g_i(z_j : t)$ and the preference distribution $F^z_{\mu,\eta}$ has been discussed in Section 3; this section therefore focuses on identifying the remaining structural objects. In particular, I exploit moments associated with comparative statics that are derived from the theoretical monotonicity properties of the equilibrium outcomes established in Section 3.3.

5.1 Identification of Caring Parameter $\lambda(z_j)$: Revealed Preference Approach via Indirect Inference

Because I employ indirect inference, the auxiliary statistics need not be correctly specified in the usual sense. What matters is that the auxiliary statistics computed from actual data and those computed from simulated data share the same data-generating process. By using the relationship between wages and individuals' optimal behavior as auxiliary statistics, my approach—unlike Flinn et al. (2018) and related work—allows for correlation between unobserved preference heterogeneity and wage heterogeneity.[23]

The identification of the caring parameter $\lambda(z_j)$ constitutes one of the main contributions of this paper. Given that time-allocation data identify the technology parameter $g_i(z_j)$ and the preference distribution $F^z_{\mu,\eta}$, I identify the caring parameter $\lambda(z_j)$ by matching model-implied slopes to their empirical counterparts;

[22]In addition, using PSID-CDS III (2007/08) and CDS 2014 with additional subjective well-being information (K6). K6 psychological distress measures are available for both parents in PSID since 2001. The K6 asks respondents (primarily Head and Spouse/Partner) about psychological distress over the past 30 days and has been included since at least 2001. While studies such as Flinn et al. (2018) and Todd and Zhang (2020) construct Big Five personality measures in a relatively straightforward manner, similar approaches can be used to construct a single well-being index from K6. Life satisfaction should be closer to the concept of indirect utility; K6 serves primarily as a robustness check.

[23]Consistency of indirect inference requires that the actual data-generating process and the model-implied process be sufficiently aligned; even if the auxiliary model is misspecified, estimates remain consistent provided both simulated and actual data are subjected to the same auxiliary regression.

see Section 3.3 for the underlying monotonicity results.

This identification strategy enables me to disentangle whether the observed “happiness puzzle” arises from changes in preferences ($F^{z}_{\mu,\eta}$), changes in technological constraints ($g_i(z_j)$), or changes in spousal cooperation ($\lambda(z_j)$).

5.2 Auxiliary Regressions: Childcare–Work Slope

This subsection constructs moments based on the childcare–work slope introduced in Section 3.3,

$$\kappa_i(\lambda) \equiv \left. \frac{\partial\, l_i^{c*}(l_i^w, \lambda)}{\partial\, l_i^w} \right|_{l_i^w = l_i^{w*}},$$

which captures how individual i's equilibrium childcare time in Stage 2 responds to i's (predetermined) market-work time chosen in Stage 1. By Theorem 1, the Stage-2 equilibrium is unique, so derivatives are well-defined and $\kappa_i(\lambda)$ exists. As established in Section 3.3, $\kappa_i(\lambda)$ varies monotonically with the cooperation parameter $\lambda(z_j)$, which allows us to use empirical estimates of $\kappa_i(\lambda)$ as identifying moments for $\lambda(z_j)$.

Because

$$\frac{\partial}{\partial \lambda}\left(\frac{\partial l_i^{c*}}{\partial l_i^w}\right) = -\frac{2\mu_i \eta_{-i}}{\left(2\mu_i + \eta_i + \lambda(z_j)\eta_{-i}\right)^2} < 0,$$

so a larger λ makes the childcare–work slope $\kappa_i(\lambda)$ smaller (more negative or less positive, depending on the sign convention in the solution).

Indirect inference via auxiliary regression. To implement indirect inference, using actual data, I estimate the following auxiliary regression for individual i's childcare time in household j,

$$l_{i,j}^c = \alpha + \kappa_i\, l_{i,j}^w + \kappa_{-i}\, l_{-i,j}^w + X_j'\pi + \varepsilon_{ij}, \qquad \text{(Aux-}\kappa\text{)}$$

where X_j includes controls (demographics, number and ages of children, etc.), π is the coefficient vector for controls, and ε_{ij} is the error term. Alternatively, one may use a simplified specification,

$$l_{i,j}^c = \alpha + \kappa_i\, l_{i,j}^w + \varepsilon_{ij}. \qquad \text{(Aux-}\kappa\text{:simple)}$$

In the main results, I use the above simpler one. The moment condition targets the OLS coefficient itself. Let $\kappa_i^{\text{model}}(\lambda;\theta)$ denote the OLS slope obtained by running the same auxiliary regression (Aux-κ:simple) on simulated data generated under θ. I treat $\widehat{\kappa_i}$ as an auxiliary statistic and match it to $\kappa_i^{\text{model}}(\lambda;\theta)$

$$m_1(\theta) \equiv \widehat{\kappa_i} - \kappa_i^{\text{model}}(\lambda;\theta) = 0.$$

5.3 Auxiliary Regressions: Cross-Wage Effect on Childcare

This subsection constructs auxiliary moments based on the cross-wage response of parent i's equilibrium childcare time. Let j index households, let i denote one spouse/parent, and let $-i$ denote the other.

Following Section 3.3, I define the level-derivative

$$\psi_{i,-i}(\lambda) \;\equiv\; \frac{\partial l_{i,j}^{c*}}{\partial w_{-i,j}} = \frac{1}{w_{-i,j}} \frac{\partial l_{i,j}^{c*}}{\partial \ln w_{-i,j}}. \tag{7}$$

The second equality follows from the chain rule; the existence of a closed-form expression for $l_{i,j}^{c*}$—and hence for $\psi_{i,-i}$—relies on the Cobb–Douglas specification of preferences and child quality production. By Theorems 1 and 2, the equilibrium is unique, so $\psi_{i,-i}(\lambda)$ is well-defined. As discussed in Section 3.3, the model-implied mapping $\lambda \mapsto \psi_{i,-i}(\lambda)$ is monotone.

To exploit this monotonicity, I estimate the auxiliary regression

$$l_{i,j}^{c} = \alpha + \psi_{i,-i} w_{-i,j} + \psi_{i,i} w_{i,j} + X_j'\pi + u_j, \tag{Aux-A}$$

where X_j denotes controls (demographics, number and ages of children, year fixed effects, etc.). As before, one may use a simplified specification, and I use the following one in the main results,

$$l_{i,j}^{c} = \alpha + \psi_{i,-i} w_{-i,j} + u_j. \tag{Aux-A: Simple}$$

To incorporate this information in estimation, I treat $\widehat{\psi}_{i,-i}$ as an auxiliary statistic and match it to the model-implied counterpart

$$m_2(\theta) \equiv \widehat{\psi}_{i,-i} - \psi_{i,-i}^{\text{model}}(\lambda;\theta) = 0,$$

where $\psi_{i,-i}^{\text{model}}(\lambda;\theta)$ is the OLS slope from running the same auxiliary regression (Aux-A: Simple) on simulated data under θ. As with κ, the target is the OLS coefficient, not the structural derivative.

5.4 Identification of Happiness Sensitivity (Thresholds)

I follow a similar approach to Friedberg and Stern (2014) by matching subjective well-being.[24] My framework distinguishes between changes in structural parameters and changes in the "happiness reporting function." Specifically, by estimating the thresholds $\{\tau_k(z)\}$ in the ordered response model separately for households with a disabled child ($z = 1$) and those without ($z = 0$), I capture differences in how spouses perceive and report happiness.

Parents of children with disabilities may shift their satisfaction standards (reference points) and derive greater joy from "trivial things" or small daily achievements. In the model, such behavioral adaptation is represented as a shift in the thresholds that map latent utility onto Life Satisfaction response categories (e.g., "very satisfied").

As a result, the model structurally separates two distinct mechanisms:

- **Technology and Cooperation Effect:** An increase in underlying utility driven by changes in $g_i(z_j)$ or $\lambda(z_j)$.
- **Response Shift Effect:** A change in the way happiness is perceived and reported, captured by shifts in $\tau_k(z)$.

This distinction provides a complete interpretation of the puzzle discussed in Section 2.1: despite the considerable difficulties associated with raising a child with disabilities, reported happiness remains comparable to that of households without disabled children.

5.4.1 Subjective well-being threshold

At this stage, all structural parameters θ have been identified from behavioral data. The term ε_i in the equations below represents the reporting error commonly discussed when using subjective data.[25]

Denoting the latent satisfaction as $S_i^* = U_i(\hat{\theta}, \varepsilon_i)$ with the scale normalization $\text{Var}(\varepsilon_i) = 1$, I estimate thresholds separately for each group $z \in \{0, 1\}$:

$$S_i = k \quad \Longleftrightarrow \quad \hat{\tau}_{k-1,z} < S_i^* \leq \hat{\tau}_{k,z}, \qquad k = 1, \ldots, K.$$

This procedure reveals how thresholds differ across z—that is, whether individuals become more "lenient" or "strict" in their happiness reporting (i.e., whether they tend to report higher satisfaction from small achievements).[26]

[24] Robustness checks using K6 (psychological distress scale) are conducted in supplementary analyses.

[25] See, for example, French and Jones (2011) and Low and Pistaferri (2015).

[26] Because $U_i(\hat{\theta}, \varepsilon_i)$ already pins down the location of the latent index, no location normalization is required. However, the variance must be normalized because threshold values scale with the variance of the latent utility distribution. Without normalization, changes in utility scale and changes in thresholds would be observationally equivalent.

The use of subjective data for identification is well-established; see French and Jones (2011), Friedberg and Stern (2014), Byrne et al. (2009), Chiappori et al. (2018), and Wiswall and Zafar (2018). Low and Pistaferri (2020) also note that while discrete measures pose challenges, they are practically useful, as in Low and Pistaferri (2015).

To implement this ordered regression in practice, I use PSID waves from 2009, 2011, 2013, and 2015, as only PSID contains subjective well-being measures. However, PSID lacks detailed childcare time data needed to construct indirect utility. Therefore, in this step, I simulate optimal individual behaviors using $R = 200$ draws for each household, compute simulated indirect utility, and then average across draws. The ordered regression is then estimated using this simulated utility. The construction of simulated indirect utility follows Friedberg and Stern (2014). Conceptually, this is similar to CCP-based methods that first estimate conditional choice probabilities and then simulate the value function. Friedberg and Stern (2014) explicitly include the discrete subjective happiness measure in the likelihood function, thereby targeting it as a moment. For the precise econometric specification, see Section 6.5.

Sequential Identification of Utility and Reporting Thresholds. A key methodological contribution is resolving the classical non-identification problem in comparing subjective well-being across groups. In standard ordered response models, utility levels U_i and reporting thresholds τ_k are jointly estimated from subjective data, making it impossible to distinguish whether group differences reflect true utility gaps or differences in reporting scales. My approach circumvents this problem by first identifying all structural parameters—including the caring parameter $\lambda(z_j)$ and technology $g_i(z_j)$—from behavioral data (labor supply, childcare time, wages) via indirect inference. With $U_i(\theta)$ pinned down from revealed preference, subjective well-being data are used *solely* to identify thresholds $\tau_k(z)$, enabling a clean decomposition of the happiness puzzle into genuine welfare effects versus response shift.

As an alternative approach, one may jointly estimate all parameters following Friedberg and Stern (2014), augmenting the moment conditions with residuals from the ordered response model.[27]

5.5 Wage equation, measurement error and second moments

Wage equation parameters are identified from observed wages and individual characteristics following standard Mincer regression approaches. As described in more detail in Section 6.2, the observed log wage is assumed to be decomposed as $\ln w_i^{\text{obs}} = \ln w_i + \varepsilon_i^{\text{me}}$, where w_i is the true wage determining labor supply and childcare decisions and $\varepsilon_i^{\text{me}} \sim N(0, \sigma^2_{\text{me},i})$ is classical measurement error independent of all structural shocks and behavioral outcomes (equation (9)). The variance of true wages, $\text{Var}(\ln w)$, is identified through behavioral moments—covariances between labor hours and log wages, cross-spouse

[27] Friedberg and Stern (2014) augmented moments of quite different types in a loss-function-like form.

wage covariances $\mathrm{Cov}(\ln w_h, \ln w_w)$, and cross-wage childcare slopes—because these moments depend on the true wage but are orthogonal to $\varepsilon^{\mathrm{me}}$ by the independence assumption.[28] Given $\mathrm{Var}(\ln w)$, the measurement error variance is identified as the residual: $\sigma^2_{\mathrm{me}} = \mathrm{Var}(\ln w^{\mathrm{obs}}) - \mathrm{Var}(\ln w)$.

6 Econometrics

This section details the econometric specification for the structural model introduced in Section 3. The preference parameters are specified in Section 6.1, and the wage equation follows in Section 6.2. These specifications determine the functional forms used in the equilibrium conditions (Theorems 1–2) and the identification strategy (Section 5).

6.1 Preference Specification

Household heterogeneity is captured by the preference parameters μ_i and η_i for $i \in \{h, w\}$, which are drawn from a joint distribution $F^z_{\mu,\eta}$ that depends on the child's disability status z. Both the mean vector and the covariance matrix of this distribution are allowed to vary with z. For each value of z, there are four preference parameters $(\mu_h, \eta_h, \mu_w, \eta_w)$ and four corresponding endogenous choice variables $(l^{w*}_h, l^{w*}_w, l^{c*}_h, l^{c*}_w)$, representing the optimal labor supply and childcare time for the husband and wife respectively.

Following the approach of Flinn et al. (2018), this paper treats labor force participation as endogenously determined. Consequently, I must specify a distributional assumption for $(\mu_h, \eta_h, \mu_w, \eta_w) \mid z$. This means the model requires parametric functional forms in nature. An important consideration, as discussed in Del Boca and Flinn (2012), is that when corner solutions are permitted, the magnitude of preference parameters below the participation threshold cannot be backed up from observed data—if an individual's preference for leisure exceeds a certain level, they supply zero labor, but the data do not reveal how far above that threshold their latent preference lies.

To parameterize the distribution, let $\boldsymbol{\phi}^z = (\phi_1, \phi^z_2, \phi_3, \phi^z_4)$ denote the vector of means of $(\mu_h, \eta_h, \mu_w, \eta_w)$. The dependence of certain mean components on z allows the model to capture systematic differences in preferences across households with and without a disabled child. Let $\boldsymbol{\Sigma}^z$ denote the covariance matrix of $(\mu_h, \eta_h, \mu_w, \eta_w) \mid z$. Both $\boldsymbol{\phi}^z$ and $\boldsymbol{\Sigma}^z$ (and hence the full distribution $F^z_{\mu,\eta}$) are permitted to vary across disability status z.

[28]They are identified through generalized residuals constructed from first-stage estimates, following Gourieroux et al. (1987), Goeree (2008), and Friedberg and Stern (2014).

6.2 Wage Equation and Non-labor Income

When a reduced-form wage process is sufficient. In many applications, the counterfactual experiments primarily target *intra-household allocation*, *labor supply*, and the *household cooperation/bargaining regime*, while the distribution of wage offers is treated as exogenous (i.e., the policy does not directly shift the offer distribution on the firm side). In this case, the goal of the model is not to micro-found wage formation itself, but to *identify household behavior* from observed choices and outcomes. Accordingly, it is sufficient to approximate the wage process as a parsimonious *conditional offer distribution* given the relevant state variables (e.g., education, experience, persistent heterogeneity, and aggregate conditions), rather than imposing a fully structural wage-setting mechanism.

Building on this motivation, I now specify the wage equation. Following Del Boca et al. (2014), the wage specification includes education level, age, and age squared as explanatory variables.The wage equations also capture a component of assortative mating on unobserved determinants by allowing correlation in the error terms across spouses, as described in Del Boca et al. (2014). Unlike Gobbi (2018), who treats non-labor income as negligible, this study incorporates non-labor income into the household budget constraint.

In household j, the wage offer processes are assumed to have the following structure,

$$\begin{bmatrix} \ln w_h \\ \ln w_w \end{bmatrix} = \begin{bmatrix} \nu_h \\ \nu_w \end{bmatrix} + \begin{bmatrix} \zeta_h \\ \zeta_w \end{bmatrix}, \qquad \begin{bmatrix} \zeta_h \\ \zeta_w \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \sigma_{hh} & \sigma_{wh} \\ \sigma_{hw} & \sigma_{ww} \end{bmatrix}\right). \tag{8}$$

Here w_i denotes the true wage that enters the household's optimization problem; agents observe their own w_i when making labor supply and childcare decisions. Survey-reported wages, however, are subject to reporting error (rounding, recall bias, imputation). The econometrician observes

$$\ln w_i^{\text{obs}} = \ln w_i + \varepsilon_i^{\text{me}}, \qquad \varepsilon_i^{\text{me}} \sim N(0, \sigma^2_{\text{me},i}), \qquad i \in \{h, w\}, \tag{9}$$

where $\varepsilon_i^{\text{me}}$ is independent of all structural shocks $(\mu_h, \eta_h, \mu_w, \eta_w, \zeta_h, \zeta_w)$ and of all behavioral outcomes. I include measurement error in the wage equation because, as discussed in Flinn et al. (2025), survey data on wages typically include measurement error.[29] In estimation, simulated moments are computed from w_i^{obs} to match the data moments constructed under the same measurement conditions. Let $\boldsymbol{\sigma} = (\sigma_{hh}, \sigma_{hw}, \sigma_{wh}, \sigma_{ww})$ collect the wage shock variance-covariance parameters. The terms ν_h and ν_w are the conditional means of the log wage draws for parent h and parent w, respectively. In my empirical work, I

[29] Bound et al. (1994) find that measurement error in survey-reported earnings accounts for 50%–60% of the variance. My estimates of σ_{me}, reported in Section 8.1, are consistent with this range.

assume a Mincer equation specification,

$$\nu_{i,t} = \nu_i^0 + \nu_i^1 age_i + \nu_i^2 age_i^2 + \nu_i^3 s_i, \qquad i \in \{h, w\}, \tag{10}$$

where s_i is the completed schooling level (which is time-invariant) of parent i. The coefficients ν_i^1 and ν_i^2 capture age effects in the wage offer, while ν_i^3 captures the labor market "return to schooling" for each parent.[30]

The disturbances in the parental wage equations are allowed to be correlated, which could arise through assortative mating on unobservable determinants of wages and through spouses sharing the same local labor market. Importantly, the distribution of wage shocks F_ζ and the distribution of preferences $F_{\mu,\eta}^z$ are correlated. The specific parameterization of this joint covariance structure is detailed in Section 6.3.1.

Parametric Specification of the Caring Parameter. In Section 3, the caring parameter $\lambda(z_j)$ was introduced as depending solely on the child's disability status z_j. For estimation, I generalize this to allow dependence on the child's age t,

$$\lambda(z_j, t) = \frac{1}{1 + \exp\big(-(\alpha_0 + \alpha_1 z_j + \alpha_2 t + \alpha_3 z_j \cdot t)\big)}, \tag{11}$$

which ensures $\lambda \in (0, 1)$. The parameter α_0 governs the baseline level of Stage-2 altruism, α_1 captures the main effect of having a disabled child on Stage-2 altruism, α_2 captures how this altruism evolves as the child ages (common to all households), and α_3 captures the interaction between disability status and child age. When $\alpha_2 \neq 0$ or $\alpha_3 \neq 0$, the caring parameter varies with the child's age, allowing for the possibility that the caregiving burden associated with a disabled child may evolve over time—for instance, through caregiver burnout ($\alpha_3 < 0$) or adaptation ($\alpha_3 > 0$).

6.3 Implementation of Unobserved Heterogeneity

6.3.1 Restrictions on the Covariance Structure

Let $\boldsymbol{\Delta}_j^z = (\mu_h, \eta_h, \mu_w, \eta_w, \zeta_h, \zeta_w)'$ denote the vector of unobserved heterogeneity for household j with disability status z. The unrestricted covariance matrix contains 21 free parameters. I impose six zero restrictions on cross-spouse different-type correlations (e.g., $\text{Cov}^z(\mu_h, \eta_w) = 0$), retaining all within-person correlations and cross-spouse same-type correlations (leisure-leisure, childcare-childcare, wage-wage).

[30] The choice of covariates in the Mincer equation varies across the literature. For example, Flinn et al. (2018) does not include year of birth, while Del Boca and Flinn (2014) includes it. Some specifications omit the quadratic age term. These variations do not substantially affect the main results.

This reduces the parameter count from 21 to 15. The restrictions capture assortative mating through same-type correlations while maintaining computational tractability.[31]

Positive definiteness is ensured via a restricted Cholesky factorization $\Sigma^z = U^{z\prime}U^z$ with 15 free parameters. The latent normal draws $\boldsymbol{q} \sim \mathcal{N}((\boldsymbol{\phi}^z, 0, 0), \Sigma^z)$ are then transformed into preference parameters via a softmax mapping following Del Boca and Flinn (2012)

$$\mu_i = \frac{\exp(q_{\mu,i})}{1 + \exp(q_{\mu,i}) + \exp(q_{\eta,i})}, \qquad \eta_i = \frac{\exp(q_{\eta,i})}{1 + \exp(q_{\mu,i}) + \exp(q_{\eta,i})}, \qquad \zeta_i = q_{\zeta,i}, \tag{12}$$

for $i \in \{h, w\}$. This ensures $\mu_i, \eta_i \in (0, 1)$ with $\mu_i + \eta_i < 1$. The full list of zero restrictions, the explicit Cholesky factor, and the derivation of constrained elements are provided in Appendix H.

6.4 Latent Type of Child SSI Receipt

Following the approach of Sullivan (2006), French (2005), and Iskhakov (2010), I do not use the self-reported d_j (Child SSI receipt) directly, but instead impute it in estimation.[32] This approach is motivated by the observation that the self-reported Child SSI receipt rate in the data is approximately 10 percent, which appears too low relative to external institutional and macro-level data.[33]

From external institutional and macro-level data, I know that $P(d_j = 1 \mid z_j = 1) = 0.7$, meaning that approximately 70 percent of households with a disabled child receive Child SSI. Since individual-level receipt data are unreliable, neither the variance nor the mean (i.e., the cost parameter $cost_j$) of the receipt decision can be identified directly from the micro data. I apply an indirect inference approach to address this limitation.

Specifically, when generating simulated data in estimation, I randomly assign $d_j = 1$ with probability 0.7 and $d_j = 0$ with probability 0.3 for households with $z_j = 1$. This ensures that the simulated data match the population take-up rate implied by external information, thereby eliminating the discrepancy between actual and simulated data that would arise from misreporting. This approach is a direct application of indirect inference. The details of how this assignment is implemented in the simulation are described in Section 7.2.

[31]More flexible covariance structures, as in Del Boca and Flinn (2012) and Flinn et al. (2018), could be accommodated but would substantially increase the computational burden.

[32]A limitation of using self-reported disability status is that the reporting may not be reliable.

[33]As documented by Celhay et al. (2025), survey reports of program participation often suffer from measurement error. Therefore, it may be preferable not to rely directly on self-reported receipt status.

6.5 SWB Measurement Equation

This subsection specifies the measurement equation linking the structural model's utility to observed subjective well-being (SWB) data. Let $U_i(\hat{\theta})$ denote the (latent) subjective welfare level for household (individual) i implied by the structural model, and let $S_i \in \{1, \ldots, K\}$ denote the observed life satisfaction response (e.g., on a 5-point scale). I treat this categorical response as an *ordered response*,

$$\begin{aligned} S_i^* &= U_i(\hat{\theta}) + \varepsilon_i, \\ S_i = k \quad &\Longleftrightarrow \quad \tau_{k-1} < S_i^* \le \tau_k, \qquad k = 1, \ldots, K, \end{aligned} \tag{13}$$

where S_i^* is a continuous latent index and $\{\tau_k\}_{k=0}^{K}$ are threshold parameters (cutpoints) satisfying $\tau_0 = -\infty$ and $\tau_K = +\infty$.

Role of the Error Term ε_i. The error term ε_i in equation (13) represents *measurement noise* that causes observed SWB responses to vary even when the underlying latent welfare $U_i(\hat{\theta})$ is the same. This noise encompasses several sources: (i) transitory mood, fatigue, or variation in question interpretation (response noise); (ii) determinants of SWB not explicitly included in the model (health shocks, family events, etc.); and (iii) coarsening or quantization that arises from mapping a continuous quantity onto K discrete categories. Consequently, the mapping from $U_i(\hat{\theta})$ to S_i is probabilistic rather than deterministic, and ε_i is an essential component of the observation equation for treating SWB data consistently.[34]

Cross-Group Comparability of Thresholds ($z = 0$ vs. $z = 1$). In this paper, the indirect utility obtained from the structural model (via simulation) is denoted $U_i(\hat{\theta})$, and the SWB categorical response is denoted $S_i \in \{1, \ldots, K\}$ (with latent index S_i^*). The measurement (reporting) equation is

$$\begin{aligned} S_i^* &= U_i(\hat{\theta}) + \varepsilon_i, \qquad \varepsilon_i \sim \mathcal{N}(0, 1), \\ S_i = k \iff &\tau_{k-1,z} < S_i^* \le \tau_{k,z}, \qquad k = 1, \ldots, K. \end{aligned} \tag{14}$$

The thresholds $\{\tau_{k,z}\}_{k=1}^{K-1}$ are allowed to depend on the child's disability status $z \in \{0, 1\}$.

The key point is that in equation (14), I *define* S_i^* in the same units as $U_i(\hat{\theta})$. This definition is equivalent to normalizing to one the coefficient that would otherwise multiply $U_i(\hat{\theta})$ in the measurement equation. Furthermore, by imposing the standard ordered probit scale normalization $\text{Var}(\varepsilon_i) = 1$ *commonly across* both values of z, the scale of the latent index S_i^* is pinned down uniformly across groups. Consequently,

[34] For discussions of noise in subjective data, see French and Jones (2011) and Low and Pistaferri (2015).

$\tau_{k,0}$ and $\tau_{k,1}$ are both estimated on the same U-scale, and

$$\Delta\tau_k \equiv \tau_{k,1} - \tau_{k,0}$$

is comparable across groups. This difference summarizes the response shift in reporting cutpoints—that is, the shift in how households with disabled children ($z = 1$) map the same latent welfare level into categorical responses, relative to households without disabled children ($z = 0$).

7 Estimation

The estimation combines data from multiple sources—the Survey of Income and Program Participation (SIPP), the Panel Study of Income Dynamics (PSID), and the PSID Child Development Supplement (CDS)—to construct moments that discipline different aspects of the structural model. Table 3 summarizes the mapping between moment categories and their corresponding data sources. Appendix I presents descriptive statistics for these moments.

Table 3: Moments and Data Sources

Label	Moment Description	Dataset
(M1)	Labor supply hours for husband and wife (conditional on employment)	SIPP 2004 W4 W8, SIPP 2008 W5 W8, PSID 2009 2011 2013 2015, PSID CDS 2007, 2014
(M2)	Time spent on childcare for husband and wife	PSID CDS 2007, 2014
(M3)	Household head's subjective well-being, K6	PSID 2009 2011 2013 2015
(M4)	Wage information (conditional on employment)	SIPP, PSID (all waves)
(M5)	Child SSI receipt	SIPP 2004, 2008

[a] Although SIPP contains information on time spent with children, the measurement error and inconsistent definitions are relatively large, making it difficult to construct childcare time at the same quality level as PSID CDS. Therefore, this study does not use childcare time from SIPP (for either fathers or mothers), and instead relies exclusively on PSID CDS for childcare time measures.

The moments can be organized into several categories based on their role in the estimation. *Level moments* (L1, L2) target mean childcare time and labor supply by disability status z_j, disciplining the technology shifter $g_i(z_j)$ and the resource channel. *Slope moments* (S1, S2) capture how childcare responds to changes in own labor supply and spouse's wage, providing the key variation for identifying $\lambda(z_j)$ through the monotonicity properties established in Proposition 4. *Wage moments* (W1) discipline the level and scale of the wage distribution. *Second-moment conditions* (A1–A3, B1–B3, C1–C3) target variances and covariances of residual wages, labor supply, and childcare time—both within individuals and across spouses—to pin down the joint distribution of unobserved heterogeneity (μ_i, η_i, ζ_i). Finally, *SWB threshold moments* (T1, T2) calibrate the mapping from model-implied utility to observed well-being categories, allowing for response shift across z. Appendix E provides the detailed construction, model counterparts, and identification role of each moment.

7.1 Estimation Method

I estimate the model using indirect inference. Given individual exogenous characteristics and candidate parameter values, the estimation procedure computes the distance between moments from simulated data and their empirical counterparts.

Let R denote the number of simulation draws, set to $R = 100$ for the main structural estimation.[35] Define $\hat{\bar{\Xi}}(\theta)$ as the vector of auxiliary statistics computed from R simulated datasets generated under parameter vector θ, and let $\bar{\Xi}$ denote the corresponding vector of auxiliary statistics computed from the actual data. The SMM estimator $\hat{\theta}$ minimizes the weighted distance between simulated and empirical moments

$$\hat{\theta} = \arg\min_{\theta} \left[\left(\hat{\bar{\Xi}}(\theta) - \bar{\Xi} \right)' W \left(\hat{\bar{\Xi}}(\theta) - \bar{\Xi} \right) \right],$$

where W is a weighting matrix.

Following much of the literature (e.g., Altonji and Segal, 1996), I do not use an optimal weighting matrix. Instead, W is specified as a diagonal matrix whose baseline elements are the inverse variance of each moment with scale adjustments applied as needed.[36]

More concretely, because the estimation relies on indirect inference, the researcher has discretion over which auxiliary statistics to include and how many moments to draw from each block—a choice that itself implicitly determines each block's relative influence on the estimates. However, inverse-variance normalization alone does not equalize the influence of different moment blocks on the objective function. Covariance moments, despite having large standard errors, also have raw magnitudes that far exceed those of slope moments such as κ and ψ. Without further adjustment, the noisy covariance block would dominate the objective function and swamp the well-identified slope moments that are central to identifying λ and g_i. I therefore apply additional block-level scale weights so that each block of moments—first moments, slope moments (κ, ψ), standard deviations, covariances, and wage moments—contributes approximately equally to the objective function (see Appendix G for details).[37]

[35] For the ordered probit threshold estimation (Section 8.2), I use $R = 200$ draws to reduce simulation noise in the latent utility index, since the threshold estimates are sensitive to the precision of simulated indirect utility.

[36] See Sauer and Taber (2021) and Jakobsen et al. (2024) for discussions of weighting in indirect inference, and Flinn et al. (2018) for discussion of moment selection. Additional weight adjustments are also used in Gayle and Shephard (2019) and Sauer and Taber (2021).

[37] As a robustness check, I also estimated the model using an identity weighting matrix (i.e., equal weight on every moment

The structural model ensures uniqueness of the game's solution and continuity of all choice variables, which permits the use of derivative-based optimization methods. In practice, I employ both the Nelder–Mead algorithm and derivative-based methods for optimization. I compute standard errors via nonparametric bootstrap with 200 replications, clustering at the household level.

7.2 Structural Model Simulation

Given a set of parameter values and exogenous characteristics, I can simulate the model. The simulation proceeds as follows. First, I extract households from the data (SIPP 2004 Waves 4 and 8, SIPP 2008 Waves 5 and 8, PSID 2009/2011/2013/2015, and PSID-CDS 2007/2014) that have the exogenous characteristics required for simulation: gender, each spouse's education level, each spouse's age (used in the wage equation), each spouse's non-labor income, whether the household has a disabled child, and the child's age. I construct the distribution of exogenous characteristics to generate simulated households.[38]

Second, for each simulated household j from the distribution constructed above, given their exogenous variables (gender, education, age), I specify the wage equation parameters $\nu_i^0, \nu_i^1, \nu_i^2, \nu_i^3, \sigma^2_{\text{me},i}$ for $i = h, w$, the mean vector of preferences $\boldsymbol{\phi}^z = (\phi_1, \phi_2^z, \phi_3, \phi_4^z)$ which may vary with z, and the joint distribution $F^z_{\mu,\eta,\zeta}$ of preference and wage equation error terms (parameterized by the covariance matrix Σ^z).

Third, I specify the parameters of the child quality production function $g_i(z_j; t)$, including the disability dummy coefficient β_z^i for $i = h, w$ and the child's age coefficient β_t (see Section 3.1). Given these parameters and $\lambda(z_j)$, and given the randomly assigned $d_j = 1$ (as described in Section 6.4), I compute the simulated individual choices, $O = \{\hat{w}_i, \hat{w}_i^{\text{obs}}, \hat{l}_i^{c*}, \hat{l}_i^{w*}, \hat{d}_j, \widehat{\text{Child SSI}}_j\}_{i=h,w}$ for all households. The computational details for solving the optimal labor supply and childcare allocation are provided in Appendix D.

8 Results

This section presents the estimation results for the structural model developed in Section 3. The identification strategy follows Section 5, and the moments used for estimation are detailed in Section 7.

8.1 Model Estimation Results

Table 4 reports the structural parameter estimates. The model includes 51 parameters. Panel A of Table 4 presents the preference parameters after transformation. As discussed in Section 3, the only constraint required for identification is that leisure preferences remain invariant across $z = 0$ and $z = 1$. This

without inverse-variance normalization). The key qualitative result—$\lambda_1 < \lambda_0$ and the relative magnitudes of g_h and g_w—is robust to this alternative weighting scheme.

[38] The sample is restricted to households with exactly one child.

Table 4: Structural Model Parameter Estimates (51 Parameters)

	Parameter	$z = 0$	$z = 1$	Notes
Panel A: Preference Parameters				
Leisure Pref.	μ_h	0.60** (0.004)		Invariant in z
	μ_w	0.45** (0.003)		Invariant in z
Childcare Pref.	η_h	0.192** (0.005)	0.191** (0.017)	$\Delta \approx 0$ (0.017)
	η_w	0.100** (0.004)	0.132** (0.011)	$\Delta = +0.032$ (0.010)
Panel B: Std. Deviations of Heterogeneity				
Leisure	$\sigma_{\mu,h}$	1.20** (0.060)	1.21** (0.131)	$\Delta = +0.01$ (0.117)
	$\sigma_{\mu,w}$	0.80** (0.035)	0.87** (0.059)	$\Delta = +0.07$ (0.048)
Childcare	$\sigma_{\eta,h}$	1.90** (0.066)	1.90** (0.140)	$\Delta \approx 0$ (0.123)
	$\sigma_{\eta,w}$	1.21** (0.026)	1.37** (0.045)	$\Delta = +0.16$ (0.037)
Wage	$\sigma_{\zeta,h}$	0.55** (0.012)		Invariant in z
	$\sigma_{\zeta,w}$	0.32** (0.010)		Invariant in z
Panel C: Wage Equation				
Husband	$\nu_{0,h}$	0.76** (0.003)		Intercept
	$\nu_{1,h}$	0.057** (<0.001)		Age effect
	$\nu_{2,h}$	−0.00062** (<0.001)		Age^2 effect
	$\nu_{3,h}$	0.078** (<0.001)		College premium
Wife	$\nu_{0,w}$	0.71** (0.001)		Intercept
	$\nu_{1,w}$	0.048** (<0.001)		Age effect
	$\nu_{2,w}$	−0.00049** (<0.001)		Age^2 effect
	$\nu_{3,w}$	0.065** (<0.001)		College premium
Panel D: Childcare Eff., Cooperation, ME				
Childcare Eff.	$\beta_{z,h}$	—	−1.88** (0.350)	$g_h(z=1) = 0.15$
	$\beta_{z,w}$	—	−0.07 (0.326)	$g_w(z=1) = 0.93$
	β_t	0.09** (0.003)		Child age effect
Cooperation	$\lambda(\bar{t})$	0.31** (0.040)	0.21** (0.070)	↓33%
	$\alpha_{\lambda,0}$	−0.849** (0.182)		Intercept
	$\alpha_{\lambda,1}$	−0.481 (0.348)		Disability effect
	$\alpha_{\lambda,2}$	0.005 (0.007)		Child age effect
	$\alpha_{\lambda,3}$	−0.007 (0.008)		Disability × age
Meas. Error	$\sigma_{me,h}$	0.45** (0.038)		Husband wage ME
	$\sigma_{me,w}$	0.50** (0.063)		Wife wage ME

Parameter	$z = 0$	$z = 1$	Notes
Panel E: Within-Indiv. Corr. (Husband)			
$\rho(\mu_h, \eta_h)$	0.27** (0.016)	0.19** (0.045)	$\Delta = -0.08^{**}$ (0.042)
$\rho(\mu_h, \zeta_h)$	−0.30** (0.047)	−0.59** (0.076)	$\Delta = -0.30$ (0.060)
$\rho(\eta_h, \zeta_h)$	−0.39** (0.012)	−0.38** (0.055)	$\Delta = +0.01$ (0.054)
Panel F: Within-Indiv. Corr. (Wife)			
$\rho(\mu_w, \eta_w)$	−0.12 (0.050)	0.11 (0.070)	$\Delta = +0.23^{**}$ (0.049)
$\rho(\mu_w, \zeta_w)$	−0.33** (0.030)	−0.45** (0.064)	$\Delta = -0.12$ (0.056)
$\rho(\eta_w, \zeta_w)$	−0.04 (0.030)	−0.46** (0.064)	$\Delta = -0.42$ (0.057)
Panel G: Cross-Spouse Correlations			
$\rho(\mu_h, \mu_w)$	0.53** (0.091)	0.66** (0.118)	$\Delta = +0.13$ (0.075)
$\rho(\mu_h, \eta_w)$	0.00	0.00	Constraint: fixed
$\rho(\eta_h, \mu_w)$	0.00	0.00	Constraint: fixed
$\rho(\eta_h, \eta_w)$	−0.55** (0.037)	−0.87** (0.069)	$\Delta = -0.32$ (0.058)
$\rho(\zeta_h, \zeta_w)$	0.39** (0.023)	0.92** (0.109)	$\Delta = +0.53$ (0.107)

Notes: 51 parameters estimated by SMM. $z = 0$: non-disabled; $z = 1$: disabled child households. Bootstrap standard errors in parentheses (200 replications, clustering at household level; Panel A via delta method). ** Statistically significant at the 5% level (bootstrap 95% CI excludes zero). Panel A: softmax-transformed: $\mu = e^{q_\mu}/(1 + e^{q_\mu} + e^{q_\eta})$, $\eta = e^{q_\eta}/(1 + e^{q_\mu} + e^{q_\eta})$. $\lambda(z, t) = 1/\big(1 + \exp(-\alpha_{\lambda,0} - \alpha_{\lambda,1} z - \alpha_{\lambda,2} t - \alpha_{\lambda,3} z \cdot t)\big)$; $\lambda(\bar{t})$ at $\bar{t} = 9.5$. σ_{me}: wage measurement error std. dev.

constraint ensures that changes in time allocation can be attributed to differences in childcare technology and taste for childcare.

Panel D contains the key parameters capturing differences in childcare technology between disabled and non-disabled child households. The childcare efficiency parameter for fathers is strongly negative ($\beta_{z,h} = -1.88$), implying that fathers become substantially less efficient at caring for disabled children ($g_h(z = 1) = 0.15$ compared to the normalized $g_h(z = 0) = 1$). While g_h itself drops substantially, the impact on child quality is more moderate because of the "+1" term in the production function (3). At average childcare hours, the father's input to child quality is $(g_h l_h^c + 1)^{0.5}$: approximately $(1.0 \times 18 + 1)^{0.5} \approx 4.4$ for $z = 0$ and $(0.15 \times 16 + 1)^{0.5} \approx 1.8$ for $z = 1$—a reduction of about 59%, not an 85% decline. In contrast, mothers' childcare efficiency shows only a modest, statistically insignificant decline ($\beta_{z,w} = -0.07$, $g_w(z = 1) = 0.93$). Mothers, who typically serve as primary caregivers, adapt to disability-specific care demands, maintaining most of their effectiveness. This stark asymmetry in childcare efficiency provides a technological foundation for the observed patterns of specialization in disabled child households, where mothers tend to increase their childcare time while fathers decrease theirs.

Although $\beta_{z,w}$ is theoretically identified, it is weakly identified in my empirical setting: the bootstrap standard error (0.326) is large relative to the point estimate (-0.07), and the 95% confidence interval spans $[-1.13, 0.05]$. To assess whether this imprecision affects the main conclusions, I re-estimate the model under two alternative fixed values, $\beta_{z,w} = 0$ and $\beta_{z,w} = -0.5$, holding all other parameters free. The key estimates—$\beta_{z,h}$, $\lambda(z=0)$, and $\lambda(z=1)$—remain stable across these specifications, confirming that the paper's central findings on the cooperation channel and the father's childcare efficiency are robust to the value of $\beta_{z,w}$.

The cooperation parameter λ is specified as a parametric function of disability status z and child age t: $\lambda(z, t) = \frac{1}{1+\exp(-\alpha_{\lambda,0}-\alpha_{\lambda,1}z-\alpha_{\lambda,2}t-\alpha_{\lambda,3}z\cdot t)}$. The individual coefficients $\alpha_{\lambda,1}$ through $\alpha_{\lambda,3}$ have relatively large bootstrap standard errors. This is a well-known property of logistic index models: the logistic function $\sigma(\cdot)$ is flat in the tails, so when the index $\alpha' x$ places the probability away from 0.5—as is the case here, with $\lambda \in [0.21, 0.31]$—the marginal effect $\partial\lambda/\partial\alpha_k = \lambda(1-\lambda)x_k$ is attenuated.[39] However, the data can pin down λ itself precisely even when individual α coefficients are imprecisely estimated. What matters for the economic question is not the precision of each α separately, but whether the composite objects—$\lambda(z=0, \bar{t})$ and $\lambda(z=1, \bar{t})$—are precisely estimated. At the mean child age ($\bar{t} = 9.5$), the bootstrap confirms that both are: $\lambda(z=0) = 0.31$ (SE = 0.040) and $\lambda(z=1) = 0.21$ (SE = 0.070), representing approximately a 33% reduction.

Importantly, while these estimates may appear low, they fall squarely within the range considered in prior literature: my model nests Gobbi (2018) as the special case $\lambda = 0$ (fully non-cooperative childcare

[39] At $\lambda = 0.31$, $\lambda(1-\lambda) = 0.21$; at $\lambda = 0.21$, $\lambda(1-\lambda) = 0.17$. A unit perturbation in the index therefore shifts λ by only 0.17–0.21.

allocation), and further converges to Del Boca and Flinn (2012) if the cooperative assumption at Stage 1 is also relaxed. Thus, my estimates of $\lambda(z=1)$ and $\lambda(z=0)$ are, to some degree, consistent with previous studies.

Several correlations change substantially between $z = 0$ and $z = 1$. I highlight two cross-spouse correlations. First, the wage shock correlation ($\rho(\zeta_h, \zeta_w)$) increases from 0.39 to 0.92. This high correlation reflects the common labor market constraints faced by disabled-child households: both spouses may need to find employment near specialized care facilities, require flexible schedules, or face similar employer responses to caregiving demands, generating strongly correlated wage shocks. Second, the correlation between spouses' childcare preferences ($\rho(\eta_h, \eta_w)$) shifts from −0.55 to −0.87, indicating pronounced caregiving specialization patterns in disabled-child households—when one parent takes on the primary caregiver role, the other shifts toward the labor market.

The measurement error standard deviations are estimated at $\sigma_{me,h} = 0.45$ and $\sigma_{me,w} = 0.50$. These magnitudes are consistent with the well-documented finding that hourly wages constructed from survey data on earnings and hours contain substantial measurement error. Bound et al. (1994) report that measurement error accounts for 50–60% of the variance in hourly wages, and Flinn et al. (2025) emphasize that this noise is particularly relevant for structural models of labor supply. My estimates fall squarely within this range. Note that the auxiliary regression for the cross-wage slope $\psi_{i,-i}$ (Aux-A) is estimated using observed wages $w^{\text{obs}}_{-i,j}$ in both data and simulations.

8.2 Threshold Estimation

Table 5: Ordered Probit Threshold Estimates by Group

Group	N	τ_1	τ_2	τ_3	τ_4
Husband, $z = 0$	3,437	−0.824	0.308	1.543	2.128
Husband, $z = 1$	962	−0.843	0.188	1.465	2.088
Wife, $z = 0$	1,702	−0.667	0.839	2.277	2.845
Wife, $z = 1$	482	−0.682	0.775	2.034	2.650

Notes: Thresholds estimated via maximum likelihood. τ_k represents the cutoff between life satisfaction categories k and $k+1$. The latent utility S^* is the 200-draw average of simulated indirect utility, defined on a common structural utility scale across groups.

Turning to the threshold estimates, Table 5 presents the ordered probit thresholds that map latent utility S^* into observed life satisfaction categories. I estimate the thresholds separately for each group defined by gender and disability status.

Table 5 reveals that parents of disabled children have lower thresholds for reporting higher life satisfaction categories, suggesting they become more sensitive to increments in well-being. For example, among husbands, the threshold τ_2 decreases from 0.308 for $z = 0$ to 0.188 for $z = 1$. Similarly, among wives, τ_3 decreases from 2.277 to 2.034 when comparing $z = 0$ to $z = 1$. This pattern—lower thresholds for parents of disabled children—helps explain the "happiness puzzle" documented in Section 2.1: despite facing objectively more challenging circumstances, parents of disabled children do not report substantially lower life satisfaction. The threshold estimates suggest that these parents adjust their reporting scale, requiring less latent utility to report a given satisfaction level.

8.3 In-Sample Fit

I now evaluate how well the estimated model replicates the main data moments described in Section 7. Table 6 compares the simulated moments from the model with their data counterparts.

Table 6: Comparison of Data and Simulated Moments

Panel A: Basic Moments ($z = 0$)		
Moment	Data	Simulated
Pr(employed$_h$)	0.870	0.872
Pr(employed$_w$)	0.700	0.697
$E[l_h^w \mid l_h^w > 0]$	44.49	44.63
$E[l_w^w \mid l_w^w > 0]$	37.89	37.88
$E[l_h^c]$	18.79	17.66
$E[l_w^c]$	18.16	18.41
Panel B: Basic Moments ($z = 1$)		
Moment	Data	Simulated
Pr(employed$_h$)	0.826	0.831
Pr(employed$_w$)	0.652	0.660
$E[l_h^w \mid l_h^w > 0]$	44.59	44.50
$E[l_w^w \mid l_w^w > 0]$	37.75	37.26
$E[l_h^c]$	15.63	15.61
$E[l_w^c]$	20.78	20.87

Panel C: κ (Childcare–Labor Slope)		
Moment	Data	Simulated
κ_h ($z = 0$)	−0.025	−0.056
κ_h ($z = 1$)	+0.082	+0.074
κ_w ($z = 0$)	−0.015	−0.009
κ_w ($z = 1$)	−0.022	−0.058
Panel D: ψ (Cross-Wage, ×1000)		
Moment	Data	Simulated
ψ_{hw} ($z = 0$)	+0.060	+0.056
ψ_{hw} ($z = 1$)	−0.013	+0.012
ψ_{wh} ($z = 0$)	+0.013	+0.039
ψ_{wh} ($z = 1$)	+0.034	+0.050
Panel E: Std. Dev. (Labor Hours)		
Moment	Data	Simulated
SD(l_h^w) ($z = 0$)	12.21	21.12
SD(l_w^w) ($z = 0$)	11.42	20.36
SD(l_h^w) ($z = 1$)	12.89	22.24
SD(l_w^w) ($z = 1$)	12.39	20.82
Panel F: Std. Dev. (Childcare Hours)		
Moment	Data	Simulated
SD(l_h^c) ($z = 0$)	16.32	16.70
SD(l_w^c) ($z = 0$)	13.83	13.21
SD(l_h^c) ($z = 1$)	15.52	15.84
SD(l_w^c) ($z = 1$)	15.24	15.98

Notes: Data vs. simulated moments from the structural model ($N = 20{,}000$, 51 parameters). Estimation uses Powell optimization with wage measurement error.

The model closely reproduces the level moments: the asymmetric childcare response—fathers reduce and mothers increase childcare time when $z = 1$—is matched in both direction and magnitude. The model also fits the key identifying slope moments well: the standardized differences for κ and ψ are all below one in absolute value (see Appendix G). Although ψ_{hw} under $z = 1$ exhibits a sign discrepancy between data (−0.013) and model (+0.012), this is not problematic. By Proposition 4, $\psi_{hw}(\lambda)$ is continuous and monotone in λ and crosses zero at some threshold λ^*; when the true λ lies near λ^*, the realized sign is sensitive to small perturbations and is not the object of interest. What matters for identification is the *monotone relationship* between λ and the slope, not the sign at any particular point. The small standardized differences across all κ and ψ moments support reliable identification of $\lambda(z_j)$.

An internal consistency check further supports the model's heterogeneity structure. In the model, both labor hours and childcare hours are chosen through continuous optimization over the same underlying preference heterogeneity (μ_i, η_i) and wage shocks (ζ_i). The simulated data reveal a striking asymmetry: the model over-predicts the standard deviation of labor hours (simulated ≈ 21 vs. data ≈ 12), yet accurately reproduces the standard deviation of childcare hours (simulated ≈ 13–17 vs. data ≈ 14–16). The key difference between these two outcomes lies not in the model but in the data-generating environment. In the labor market, institutional constraints—standard full-time contracts clustered around 40 hours per week and part-time norms around 20 hours—compress the observed distribution of work hours into a bimodal shape with narrow peaks.[40] The model, which allows a continuous hours choice without such institutional frictions, naturally generates a wider distribution. Childcare hours, by contrast, are not subject to any comparable institutional bunching: there is no "standard childcare week," and parents genuinely range from zero to over 40 hours. The model's continuous optimization therefore maps preference heterogeneity into childcare dispersion at roughly the correct scale. This selective pattern—where only the institutionally constrained outcome exhibits excess dispersion—provides evidence that the variance parameters are correctly estimated and that the labor-hour misfit reflects an omitted institutional feature rather than a fundamental misspecification of the preference or technology structure. Incorporating discrete hours choices, bunching, (e.g., 0/20/40-hour options) to model institutional bunching would eliminate the closed-form solutions on which the equilibrium uniqueness proofs (Theorems 1–2) and the monotonicity results (Proposition 4) depend. Since these theoretical properties constitute a central contribution of the paper, sacrificing them to improve the fit of a secondary moment would be disproportionate.[41]

Appendix G also reports large standardized differences in some covariances, including sign reversals. Covariance moments have large raw scale, and accordingly the estimation assigns them very small per-moment weights to equalize each block's contribution to the objective function (Section 6). The estimation

[40] In the literature, this is called bunching. See, for example, Keane and Wasi (2013).

[41] One can see the similar pattern in, for example, Del Boca et al. (2014). As discussed in Kaplan (2012), even if the structure of second moments is modified to seek for better fit, the estimators of λ and the childcare efficiency parameters remain approximately robust, because, as shown in Section 5, these parameters are primarily identified by moments associated with first moments.

therefore does not target these moments; the large standardized differences reflect this design choice. The key structural parameters—λ and g_i—are identified by first moments that the model matches closely: λ by slope moments (κ, ψ) and g_i by level moments of childcare time (L1).[42]

Following Flinn et al. (2018) and Del Boca et al. (2014), the present paper adopts a very parsimonious specification. These fitting patterns are comparable to those in the existing literature on structural household models (for example, Del Boca et al. (2014), Chiappori et al. (2018) and Gayle and Shephard (2019)), where fitting all of first and second moments precisely is challenging due to the prevalence of corner solutions and selection effects.[43] See Appendix G for fits of all moments. The present model's simpler structure is chosen to maintain computational tractability while still capturing the key behavioral differences between disabled and non-disabled child households. More flexible specifications are left for future work.

9 Counterfactuals

Using the estimated structural parameters from Section 8 (Table 4), I conduct counterfactual analyses to understand the sources of welfare differences between households with and without disabled children. The key parameters that vary in my experiments are the amount of Child SSI and the childcare efficiency parameters $g_i(z_j)$ defined in Section 3.1.

These two parameters have distinct interpretations, as can be seen from the child quality production function (3) and the super utility function (6). The childcare efficiency parameter $g_i(z_j)$ is related to technology and thus can be influenced by policy intervention. In contrast, the caring parameter $\lambda(z_j)$ governs the degree of altruism in Stage-2 childcare allocation and is more closely related to parental attitudes.

9.1 Mom and Dad? If I Were Not Disabled, Would You Be Happier?

While existing research has documented that subjective well-being shows relatively little difference between households with and without disabled children, my structural model allows us to compare households in terms of their indirect utility function values. The mean value of $U_h + U_w$ is 12.09 for households with disabled children ($z = 1$) and 13.56 for households without disabled children ($z = 0$). At first glance, it may appear that households with disabled children do not achieve substantially lower indirect utility compared to those without disabled children.

However, a closer examination reveals a more nuanced picture. I ask whether Child SSI subsidies

[42]Fitting covariances is also a well-known structural challenge in household models. Comparable models—Del Boca et al. (2014), Flinn et al. (2018), Chiappori et al. (2018), Gayle and Shephard (2019)—face similar fitting problems.

[43]See Arellano and Bonhomme (2017).

alone can be helpful in closing this welfare gap. To investigate this question, I randomly generate 1,000 households with disabled children ($z = 1$) and compute their average indirect utility. I then fix the realizations of random variables for each household and hypothetically change their status to $z = 0$. The counterfactual experiment asks: how much Child SSI would be required for a household in the $z = 1$ state to achieve the utility level it would have attained in the hypothetical $z = 0$ state?

Strikingly, this exercise reveals that the required monthly payment would be economically infeasible—far exceeding any realistic policy intervention.

Household welfare is defined as the sum of super utilities from equation (6)

$$\mathcal{W} \equiv U_h + U_w = (1 + \lambda(z_j)) \times (u_h + u_w) \tag{15}$$

where $\lambda(z_j)$ is the caring parameter and $u_h + u_w$ is the sum of the spouses' felicity utilities from equation (5).

Using the estimated values from Table 4, simulation yields the following welfare levels:

$$z = 0: \quad \mathcal{W}_0 = (1 + \lambda_0) \times (u_h + u_w)_{z=0} \approx 1.310 \times 10.35 \approx 13.56 \tag{16}$$

$$z = 1: \quad \mathcal{W}_1 = (1 + \lambda_1) \times (u_h + u_w)_{z=1} \approx 1.206 \times 10.03 \approx 12.09 \tag{17}$$

The welfare difference is therefore $\Delta\mathcal{W} = 13.56 - 12.09 = 1.47$.

Decomposing the sources of $\Delta\mathcal{W}$ reveals two contributing factors. First, the coefficient $(1 + \lambda)$ differs between the two groups: 1.310 versus 1.206, representing an 8.6% difference. Second, the felicity utility levels differ: 10.35 versus 10.03, representing a 3.1% difference. Thus, the vast majority of the welfare gap originates from the difference in $\lambda(z_j)$. This finding is consistent with Theorem 3, which establishes that the caring parameter has a substantial impact on household welfare.

Increasing Child SSI raises consumption x through the budget constraint, which in turn increases felicity utility $(u_h + u_w)$. However, the coefficient $(1 + \lambda_1) = 1.206$ magnifies any shortfall in felicity utility.

To equalize the welfare of a $z = 1$ household to the $z = 0$ level of 13.56, I require

$$1.206 \times (u_h + u_w)^{\text{new}}_{z=1} = 13.56 \quad \Rightarrow \quad (u_h + u_w)^{\text{new}}_{z=1} = 11.24 \tag{18}$$

Since the current felicity utility is $(u_h + u_w)_{z=1} = 10.03$, an increase of $\Delta u = +1.21$ is required.

From the felicity utility specification in equation (5), the household's total felicity utility takes the form

$$u_h + u_w = \underbrace{\left[(1 - \mu_h - \eta_h) + (1 - \mu_w - \eta_w)\right]}_{\approx 0.63} \times \ln(x) + (\text{leisure and child quality terms}) \tag{19}$$

where I have used the estimated values from Table 4. To achieve $\Delta u = 1.21$ through consumption alone,

$$0.63 \times \Delta \ln(x) = 1.21 \tag{20}$$

$$\Delta \ln(x) = 1.92 \tag{21}$$

$$\frac{x^{\text{new}}}{x^{\text{old}}} = e^{1.92} \approx 6.8 \tag{22}$$

That is, consumption would need to increase by approximately 6.8-fold.

The composition of consumption follows from the budget constraint

$$x = w_h l_h^w + w_w l_w^w + Y + \text{Child SSI} \tag{23}$$

For the average $z = 1$ household in my simulation,

$$x = \$2{,}032 + \$109 \approx \$2{,}142/\text{week} \tag{24}$$

where Child SSI (currently \$625/month per recipient × 70% take-up ≈ \$109/week on average) constitutes approximately 5.1% of consumption.

To achieve 6.8-fold consumption (\$14,566/week),

$$\text{Required additional consumption} = \$14{,}566 - \$2{,}142 = \$12{,}424/\text{week} \tag{25}$$

$$\text{New Child SSI} = \$109 + \$12{,}424 = \$12{,}533/\text{week} \tag{26}$$

This corresponds to approximately 115 times the current Child SSI level.[44]

The simple calculation above suggests that a Child SSI increase of orders of magnitude would be necessary. However, the actual simulation results show that welfare equalizes at a comparable level, as behavioral responses to the Child SSI increase partially offset the required transfer.

The estimation results in Table 4 reveal a striking asymmetry: fathers' childcare efficiency drops substantially when caring for a disabled child ($g_h(z=1) = 0.15$), while mothers' efficiency shows only a modest decline ($g_w(z=1) = 0.93$). A natural policy question is whether programs that restore fathers' caregiving skills—such as parent training, specialized respite support, or assistive technology—could close the welfare gap.

[44]The specific multiplier depends on the logarithmic specification of felicity utility. Under a more general strictly concave utility function, the exact figure would differ. However, the qualitative conclusion—that the required Child SSI increase is economically infeasible—is robust to the choice of any strictly concave utility, because the welfare gap originates primarily from the $(1+\lambda)$ multiplier rather than from consumption levels, and any strictly concave function exhibits diminishing marginal utility that makes it increasingly costly to close the gap through consumption alone.

To investigate this, I conduct a counterfactual in which the father's childcare efficiency parameter is set to its non-disabled level, $\beta_{z,h} = 0$ (i.e., $g_h(z = 1) = 1.0$), while holding all other parameters—including $\lambda(z_j)$, $g_w(z_j)$, preferences, and the current Child SSI level—fixed. I also consider a combined scenario in which both g_h and λ are simultaneously restored to their $z = 0$ levels. Table 7 reports the results.

Table 7: Welfare Effects of Improving Paternal Childcare Efficiency

Scenario	$\mathcal{W}$	$\Delta\mathcal{W}$	vs. SSI×2	Gap closed
Baseline ($z = 1$, current params)	12.09	—	—	—
Child SSI ×2 ($1,250/month)	12.14	+0.06	1.0×	3.8%
Child SSI ×5 ($3,125/month)	12.29	+0.20	3.6×	13.7%
Father training (g_h: 0.15 → 1.0)	12.59	+0.50	9.0×	34.1%
λ equalized (0.206 → 0.310)	13.21	+1.12	20.1×	75.9%
Both (g_h restored + λ equalized)	13.76	+1.67	30.0×	113.1%
Reference: $z = 0$ (no disability)	13.56	+1.47	26.4×	100.0%

Notes: $N = 10{,}000$ households, same-household comparison. "Father training" sets $\beta_{z,h} = 0$ so that $g_h(z=1) = 1.0 = g_h(z=0)$. "$\lambda$ equalized" sets $\alpha_{\lambda,1} = \alpha_{\lambda,3} = 0$ so that $\lambda(z=1,t) = \lambda(z=0,t)$ for all t. "Both" simultaneously applies both changes. The gap closure exceeds 100% because $z = 1$ households receive Child SSI, which the $z = 0$ reference does not.

Two findings emerge. First, restoring fathers' childcare efficiency to the non-disabled level closes 34.1% of the welfare gap—approximately 9.0 times the effect of doubling Child SSI, but only 0.45 times the effect of equalizing λ. This confirms that *the cooperation channel dominates the technology channel*: even fully eliminating the father's efficiency disadvantage leaves about two-thirds of the welfare gap unresolved. Simultaneously restoring both g_h and λ closes 113.1% of the gap, overshooting the $z = 0$ reference because $z = 1$ households retain Child SSI, which the $z = 0$ reference does not receive.

Second, the behavioral responses to improved paternal efficiency are notable. Table 8 reports the time allocation changes.

Table 8: Behavioral Responses to Paternal Efficiency Improvement

Variable	Baseline ($z = 1$)	Father training	λ equalized	Both	Reference ($z = 0$)
Father labor (h/week)	37.4	36.3	37.4	36.3	38.8
Mother labor (h/week)	23.6	24.1	23.6	24.1	26.5
Father childcare (h/week)	15.9	17.8	16.9	18.7	17.6
Mother childcare (h/week)	21.8	21.6	23.8	23.6	18.2
Child quality (k)	20.2	51.3	22.7	57.5	43.3

Notes: All time variables in hours per week. Child quality $k = (g_h l_h^c + 1)^{0.5}(g_w l_w^c + 1)^{0.5}$. In all scenarios, $z = 1$ households retain their Child SSI benefits (consistent with Table 7).

Improving paternal efficiency from $g_h = 0.15$ to $g_h = 1.0$ increases child quality by 154% (from 20.2 to 51.3), yet fathers' childcare time increases only modestly (from 15.9 to 17.8 hours per week). This is because higher efficiency means each hour of paternal care produces substantially more child quality, so the father achieves a large quality improvement with only a small time increase. In contrast, equalizing λ increases fathers' childcare time by a comparable amount (to 16.9 hours) but raises child quality by only 12%, because the cooperation channel operates through time reallocation rather than through productivity gains.

The preceding counterfactual analyses yield three policy implications. First, cash transfers alone are insufficient. Compensating for the lower $(1+\lambda)$ coefficient requires a 12.1% increase in felicity utility (from 10.03 to 11.24). Due to the diminishing marginal utility inherent in logarithmic preferences, achieving this increase through consumption alone necessitates increasing consumption to approximately 6.8 times its current level. Since Child SSI constitutes only 5.1% of household consumption, this would require a Child SSI increase of orders of magnitude—a clearly infeasible policy.

Second, improving spousal cooperation is the most effective single intervention. Equalizing λ to the non-disabled level closes 75.9% of the welfare gap, while fully restoring paternal childcare efficiency closes only 34.1%. This asymmetry reflects the fact that λ affects welfare both through the $(1 + \lambda)$ scaling of felicity utility and through equilibrium reallocation of childcare time in Stage 2, whereas g_h operates through child quality alone.

A reverse counterfactual reinforces this conclusion. When $z = 0$ households are assigned the $z = 1$ cooperation level—$\lambda(z = 0, t) \rightarrow \lambda(z = 1, t)$ for all t, while retaining their non-disabled childcare efficiency—the welfare gap shrinks by 75.3%, from $\Delta\mathcal{W} = 1.47$ to 0.36. In other words, imposing only the cooperation breakdown observed in disabled-child households on otherwise identical non-disabled households reproduces three-quarters of the welfare gap. This symmetry—whether one raises λ_1 to λ_0 or lowers λ_0 to λ_1, approximately 75% of the gap is accounted for—confirms that the welfare cost of having a disabled child is driven primarily not by the disability-specific caregiving burden (g_i), but by the erosion of spousal cooperation (λ) that accompanies it.

Third, combining both channels closes the entire welfare gap (113.1%), overshooting the $z = 0$ reference because $z = 1$ households retain Child SSI. Programs that simultaneously improve fathers' caregiving skills and support spousal cooperation—such as structured parent training with a couples-based component, or workplace flexibility policies that enable shared caregiving—are therefore the most effective policy approach. Since each channel alone leaves a substantial portion of the gap unresolved, policies targeting only one channel are insufficient.

10 Conclusion

This paper develops an empirical biform game framework for intra-household allocation and applies it to understand the welfare of households with disabled children. My approach makes three main contributions to the literature.

First, I introduce the biform game structure to family economics, providing a middle ground between the fully cooperative collective model and fully non-cooperative approaches. By treating labor supply as cooperative (enforceable via employment contracts) and childcare allocation as non-cooperative (difficult to monitor), the framework captures the realistic asymmetry in household decision-making. I prove equilibrium uniqueness under standard Cobb–Douglas assumptions, resolving the multiple equilibria problem that has plagued non-cooperative household models while maintaining weaker assumptions than the collective model. This framework is not limited to disabled-child households; it applies broadly to any setting where some household decisions are enforceable while others are not—including elderly care among adult children, division of household chores, and intergenerational resource sharing.

Second, I develop a novel identification strategy based on slope moments rather than level moments. The monotonicity properties established in Proposition 4—that the childcare-labor slope κ and cross-wage effect ψ vary monotonically with the cooperation parameter λ—allow me to identify λ from behavioral responses without relying on exclusive instruments or unique datasets. This slope-based identification approach may prove useful in other structural estimation contexts where a key parameter affects the *responsiveness* of behavior rather than its level.

Third, I provide the first structural decomposition of the "happiness puzzle" in disabled-child households. Despite reporting nearly identical life satisfaction, these households exhibit 33% lower spousal cooperation and substantially reduced paternal childcare efficiency. My counterfactual analysis reveals that closing the welfare gap through Child SSI alone would require benefit increases of orders of magnitude, highlighting the fundamental limits of cash transfers. Even fully restoring fathers' childcare efficiency to the non-disabled level closes only 34.1% of the welfare gap, whereas equalizing spousal cooperation closes 75.9%. These findings underscore the primacy of the cooperation channel and the importance of policies that support spousal coordination—such as respite care programs, family counseling, and workplace flexibility—as complements to income support and parent training.

Several limitations warrant discussion. The static nature of my model precludes analysis of how cooperation evolves over time or how child disability affects long-run outcomes such as child development and marital stability. Extending the model to a dynamic setting is a nontrivial task: it requires modeling joint spousal search over labor markets, and, in addition, the best-response mappings in a dynamic household game are highly nonlinear, making multiple fixed points difficult to rule out. The Cobb–Douglas functional form, while standard in the literature and necessary for analytical tractability, imposes restrictions on substitution patterns. Additionally, my identification of λ relies on the maintained assumption that the biform game structure correctly describes household decision-making; alternative game structures could yield different conclusions.

These limitations point to directions for future research. Extending the biform framework to a dynamic setting would allow analysis of how cooperation responds to shocks and how it affects child development trajectories. Incorporating richer heterogeneity in the cooperation parameter—allowing it to depend on observable characteristics such as education or marriage duration—could shed light on which households are most vulnerable to cooperation breakdown. Finally, combining the structural approach with policy variation (e.g., changes in Child SSI eligibility rules or respite care availability) would strengthen identification and enable direct evaluation of policy interventions.

More broadly, the biform game framework opens new avenues for understanding household behavior. The recognition that different household decisions have different enforceability—and that this asymmetry shapes equilibrium outcomes—has implications beyond disabled-child households. As populations age and caregiving responsibilities grow, understanding how families coordinate care provision becomes increasingly important. The tools developed here—biform games, slope-based identification, and structural decomposition of well-being—provide a foundation for addressing these questions.

References

Abrahams, S., L. Flabbi, and J. Mabli (2025). Child SSI and family welfare. *Journal of Labor Economics*. Forthcoming.

Altonji, J. G. and L. M. Segal (1996). Small-sample bias in GMM estimation of covariance structures. *Journal of Business & Economic Statistics 14*(3), 353–366.

Arellano, M. and S. Bonhomme (2017). Nonlinear panel data methods for dynamic heterogeneous agent models. *Annual Review of Economics 9*, 471–496.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics 25*(2), 242–262.

Blundell, R. and A. Shephard (2012). Employment, hours of work and the optimal taxation of low-income families. *Review of Economic Studies 79*(2), 481–510.

Bound, J., C. Brown, G. J. Duncan, and W. L. Rodgers (1994). Evidence on the validity of cross-sectional and longitudinal labor market data. *Journal of Labor Economics 12*(3), 345–368.

Brandenburger, A. and H. W. Stuart (2007). Biform games. *Management Science 53*(4), 537–549.

Brown, M., C. J. Flinn, and J. Mullins (2025). Family law effects on divorce, fertility and child investment.

International Economic Review. Forthcoming.

Byrne, D., M. S. Goeree, B. Hiedemann, and S. Stern (2009). Formal home health care, informal care, and family decision making. *International Economic Review 50*(4), 1205–1242.

Celhay, P. A., B. D. Meyer, and N. Mittag (2025). Errors in reporting and imputation of government benefits and their implications. *Journal of Human Resources 60*(1), 1–35.

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of Political Economy 100*(3), 437–467.

Chiappori, P.-A., M. Costa Dias, and C. Meghir (2018). The marriage market, labor supply, and education choice. *Journal of Political Economy 126*(S1), S26–S72.

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

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

Cunha, F. and J. J. Heckman (2007). The technology of skill formation. *American Economic Review 97*(2), 31–47.

d'Aspremont, C. and A. Jacquemin (1988). Cooperative and noncooperative R&D in duopoly with spillovers. *American Economic Review 78*(5), 1133–1137.

Del Boca, D. and C. Flinn (2012). Endogenous household interaction. *Journal of Econometrics 166*(1), 49–65.

Del Boca, D., C. Flinn, E. Verriest, and M. Wiswall (2026). Parenting with patience: Parental incentives and child development. *Journal of Political Economy 134*(1), 210–284.

Del Boca, D., C. Flinn, and M. Wiswall (2014). Household choices and child development. *Review of Economic Studies 81*(1), 137–185.

Del Boca, D., C. Flinn, M. Wiswall, and B. Zafar (2024). Parents' beliefs and investments in children. *Journal of Political Economy 132*(12), 4077–4123.

Del Boca, D. and C. J. Flinn (2005). Household time allocation and modes of behavior: A theory of sorts. Working Paper, New York University.

Del Boca, D. and C. J. Flinn (2014). Household behavior and the marriage market. *Journal of Economic Theory 150*, 515–550.

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

Fang, H. and D. Silverman (2009). Time-inconsistency and welfare program participation: Evidence from the NLSY. *International Economic Review 50*(4), 1043–1077.

Flinn, C. J., P. E. Todd, and W. Zhang (2018). Personality traits, intra-household allocation and the gender wage gap. *European Economic Review 109*, 191–220.

Flinn, C. J., P. E. Todd, and W. Zhang (2025). Labor market returns to personality. *Journal of Political Economy*. Forthcoming.

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *Review of Economic Studies 72*(2), 395–427.

French, E. and J. B. Jones (2011). The effects of health insurance and self-insurance on retirement behavior.

Econometrica 79(3), 693–732.

Friedberg, L. and S. Stern (2014). Marriage, divorce, and asymmetric information. *International Economic Review 55*(4), 1155–1199.

Friedrichsen, J., T. König, and R. Schmacker (2018). Social image concerns and welfare take-up. *Journal of Public Economics 168*, 174–192.

Gayle, G.-L. and A. Shephard (2019). Optimal taxation, marriage, home production, and family labor supply. *Econometrica 87*(1), 291–326.

Geanakoplos, J. (2003). Nash and walras equilibrium via brouwer. In *Economic Theory*, Volume 21, pp. 585–603. Springer.

Gobbi, P. E. (2018). Childcare and commitment within households. *Journal of Economic Theory 176*, 503–551.

Goeree, M. S. (2008). Limited information and advertising in the U.S. personal computer industry. *Econometrica 76*(5), 1017–1074.

Gourieroux, C., A. Monfort, E. Renault, and A. Trognon (1987). Generalised residuals. *Journal of Econometrics 34*(1-2), 5–32.

Iskhakov, F. (2010). Structural dynamic model of retirement with latent health indicator. *Econometrics Journal 13*(3), S126–S161.

Jakobsen, K., T. H. Jørgensen, and H. Low (2024). Fertility and family labor supply. *American Economic Review 114*(9), 2749–2795.

Kaplan, G. (2012). Moving back home: Insurance against labor market risk. *Journal of Political Economy 120*(3), 446–512.

Keane, M. P. and N. Wasi (2013). The structure of consumer taste heterogeneity in revealed vs. stated preference data. *Quantitative Economics 4*(3), 481–512.

Keane, M. P. and K. I. Wolpin (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review 42*(4), 1051–1103.

Low, H. and L. Pistaferri (2010). Disability risk, disability insurance and life cycle behavior. Working Paper 15962, NBER.

Low, H. and L. Pistaferri (2015). Disability insurance and the dynamics of the incentive insurance trade-off. *American Economic Review 105*(10), 2986–3029.

Low, H. and L. Pistaferri (2020). Disability insurance: Theoretical trade-offs and empirical evidence. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*. Elsevier. Forthcoming.

Moffitt, R. (1983). An economic model of welfare stigma. *American Economic Review 73*(5), 1023–1035.

Nevo, A. (2000). A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of Economics & Management Strategy 9*(4), 513–548.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*(2), 307–342.

Quint, T. and M. Shubik (1997). A model of migration. *Cowles Foundation Discussion Paper* (1088).

Sauer, R. M. and C. Taber (2021). Understanding aggregate human capital. *Annual Review of Economics 13*, 153–178.

Sullivan, J. X. (2006). Welfare reform, saving, and vehicle ownership: Do asset limits and vehicle exemptions matter? *Journal of Human Resources 41*(1), 72–105.

Todd, P. E. and W. Zhang (2020). A dynamic model of personality, schooling, and occupational choice. *Quantitative Economics 11*(1), 231–275.

Wiswall, M. and B. Zafar (2018). Preference for the workplace, investment in human capital, and gender. *Quarterly Journal of Economics 133*(1), 457–507.

Supplementary Appendix (Online Appendix)

Measuring Spousal Cooperation: A Biform Game Approach to Intra-Household Allocation

Takahiro Moriya

The University of Tokyo

takahiro.moriya@e.u-tokyo.ac.jp

This supplementary material is organized as follows. Appendices A–C collect proofs of the theoretical results stated in the main text: uniqueness of the Stage 2 Nash equilibrium (Theorem 1), monotonicity of indirect utility in the caring parameter (Theorem 3), and monotone comparative statics of equilibrium outcomes (Proposition 4). Appendix D derives the closed-form solutions for optimal labor supply and childcare time. Appendix E details the construction of each moment condition used in estimation, and Appendix F reports the corresponding descriptive statistics. Appendix G presents the full model fit tables. Appendix H documents the restricted covariance structure, and Appendix I provides sample descriptive statistics. Finally, Appendices J and K discuss missing data in the PSID-CDS time diary and the correspondence of disability variables across datasets, respectively.

A Proof of Theorem 1

Because the child quality function is Cobb–Douglas, the child-input aggregator is additively separable in logs,

$$\ln k = \frac{1}{2}\ln(a_h l_h^c + 1) + \frac{1}{2}\ln(a_w l_w^c + 1),$$

where $a_i \equiv g_i(z_j : t)$. The marginal contribution of the husband's childcare input satisfies

$$\frac{\partial \ln k}{\partial l_h^c} = \frac{1}{2}\cdot\frac{a_h}{a_h l_h^c + 1},$$

which does not depend on l_w^c. Hence the husband's first-order condition (and thus his best response) depends only on his own choice variable l_h^c given (l_h^w, l_w^w) and parameters; symmetrically for the wife.

Therefore, Stage 2 is effectively a pair of independent one-dimensional problems: each $i \in \{h, w\}$ chooses l_i^c to maximize a strictly concave objective on a convex feasible set $l_i^c \in [0, T - l_i^w]$.

To verify strict concavity, consider the husband's FOC with respect to l_h^c. Fixing l_h^w and treating l_w^c as given,

$$\begin{aligned} 0 &= \mu_h \frac{\partial}{\partial l_h^c}\ln(T - l_h^w - l_h^c) + \eta_h \frac{\partial \ln k}{\partial l_h^c} + \lambda(z_j)\eta_w \frac{\partial \ln k}{\partial l_h^c} \\ &= -\frac{\mu_h}{T - l_h^w - l_h^c} + \big(\eta_h + \lambda(z_j)\eta_w\big)\cdot\frac{1}{2}\cdot\frac{a_h}{a_h l_h^c + 1}. \end{aligned}$$

The second derivative is

$$\frac{\partial^2 U_h}{\partial (l_h^c)^2} = -\frac{\mu_h}{(T - l_h^w - l_h^c)^2} - \big(\eta_h + \lambda(z_j)\eta_w\big)\cdot\frac{1}{2}\cdot\frac{a_h^2}{(a_h l_h^c + 1)^2} < 0,$$

confirming strict concavity.[1] The same argument applies to l_w^{c*}. Strict concavity implies each player's best response is single-valued. Since best responses do not depend on the opponent's choice, the Nash equilibrium is simply the product of the two unique optima, so (l_h^{c*}, l_w^{c*}) is unique. □

B Proof of Theorem 3

Because the structure is symmetric for i and $-i$, I focus on player i. Fix (l_h^w, l_w^w) and consider Stage 2. Let the action set be $A_i = [0, T - l_i^w]$, which is a compact complete lattice.

Define player i's Stage-2 payoff as

$$\pi_i(l_i^c, l_{-i}^c; \lambda) := u_i(\cdot, l_i^\ell, k) + \lambda\, u_{-i}(\cdot, l_{-i}^\ell, k), \qquad l_i^\ell = T - l_i^w - l_i^c.$$

Under the Cobb–Douglas specification, one can verify directly that (i) $\frac{\partial^2 \pi_i}{\partial l_i^c \partial \lambda} > 0$, so π_i has strictly increasing differences in (l_i^c, λ), and (ii) $\frac{\partial^2 \pi_i}{\partial l_i^c \partial l_{-i}^c} = 0$, so π_i has (weakly) increasing differences in (l_i^c, l_{-i}^c). The Stage-2 game is therefore supermodular. Hence by Tarski/Topkis, the set of Stage-2 equilibria is a nonempty complete lattice; by Theorem 1 the equilibrium is unique. Denote it by $l^{c*}(\lambda) = (l_h^{c*}(\lambda), l_w^{c*}(\lambda))$. By monotone comparative statics for supermodular games, each component of $l^{c*}(\lambda)$ is weakly increasing in λ.

Next consider the sum of felicity utilities at Stage 2 for fixed (l_h^w, l_w^w)

$$S(\lambda) := u_h(\cdot, l_h^{\ell*}(\lambda), k^*(\lambda)) + u_w(\cdot, l_w^{\ell*}(\lambda), k^*(\lambda)),$$

where $l_i^{\ell*}(\lambda) = T - l_i^w - l_i^{c*}(\lambda)$ and $k^*(\lambda)$ is induced by $l^{c*}(\lambda)$. Suppose the Stage-2 equilibrium is interior. The first-order conditions are

$$\frac{\partial u_h}{\partial l_h^c} + \lambda \frac{\partial u_w}{\partial l_h^c} = 0, \qquad \frac{\partial u_w}{\partial l_w^c} + \lambda \frac{\partial u_h}{\partial l_w^c} = 0.$$

Differentiating $S(\lambda)$ and using the FOCs yields

$$\frac{dS(\lambda)}{d\lambda} = \left(\frac{\partial u_h}{\partial l_h^c} + \frac{\partial u_w}{\partial l_h^c}\right)\frac{dl_h^{c*}}{d\lambda} + \left(\frac{\partial u_h}{\partial l_w^c} + \frac{\partial u_w}{\partial l_w^c}\right)\frac{dl_w^{c*}}{d\lambda} = (1-\lambda)\left(\frac{\partial u_w}{\partial l_h^c}\frac{dl_h^{c*}}{d\lambda} + \frac{\partial u_h}{\partial l_w^c}\frac{dl_w^{c*}}{d\lambda}\right).$$

Therefore, for $\lambda \le 1$, if each spouse weakly benefits from the other's childcare (i.e., $\partial u_w/\partial l_h^c \ge 0$ and $\partial u_h/\partial l_w^c \ge 0$), and since $dl_i^{c*}/d\lambda \ge 0$ by supermodularity, we obtain $dS(\lambda)/d\lambda \ge 0$. Hence, given (l_h^w, l_w^w), the felicity sum is weakly increasing in λ.

[1] See Geanakoplos (2003).

At Stage 1, define the value function

$$V(\lambda) := \max_{l_h^w, l_w^w} S(\lambda; l_h^w, l_w^w).$$

Because $S(\lambda; l_h^w, l_w^w)$ is weakly increasing in λ pointwise in (l_h^w, l_w^w), the max operator preserves monotonicity, and thus $V(\lambda)$ is weakly increasing in λ.

Finally, if $V(\lambda) > 0$, then $U_h + U_w = (1+\lambda)V(\lambda)$ is a product of two weakly increasing, nonnegative terms, and therefore is itself weakly increasing in λ. This proves Theorem 3.[2]

C Proof of Proposition 4

I use the model setup from Section 3: felicity (5), super utility (6), and child quality production (3) with $\underline{k}_i = 0$. Let $a_i \equiv g_i(z_j : t) > 0$ and $A_i(\lambda) \equiv \eta_i + \lambda\eta_{-i}$ (where $\lambda = \lambda(z_j)$ for notational simplicity).

Part 1: Monotonicity of $\psi_{i,-i}(\lambda)$.

I show the result for ψ_{hw}; the case ψ_{wh} follows by symmetry (swapping $h \leftrightarrow w$).

From Stage 2, the equilibrium childcare rule satisfies

$$l_i^{c*}(l_i^w, \lambda) = \frac{A_i(\lambda)a_i(T - l_i^w) - 2\mu_i}{a_i(2\mu_i + A_i(\lambda))}. \tag{27}$$

Since Stage 2 depends on wages only through Stage-1 labor,

$$\psi_{hw} = \frac{\partial l_h^{c*}}{\partial w_w} = \frac{\partial l_h^{c*}}{\partial l_h^w} \cdot \frac{\partial l_h^{w*}}{\partial w_w}.$$

From (27)

$$\frac{\partial l_h^{c*}}{\partial l_h^w} = -\frac{A_h(\lambda)}{2\mu_h + A_h(\lambda)}.$$

From Stage 1 (maximizing $\mathcal{S} \equiv u_h + u_w$), with $C_h \equiv \frac{\mu_h + \frac{1}{2}(\eta_h + \eta_w)}{\alpha}$ where $\alpha = 2 - \mu_h - \mu_w - \eta_h - \eta_w$,

$$\frac{\partial l_h^{w*}}{\partial w_w} = -\frac{C_h}{w_h} \cdot \frac{T + \frac{1}{a_w}}{1 + C_h + C_w} < 0.$$

Stage-1 derivatives do not depend on λ. To see this, note that $U_h + U_w = (1+\lambda)(u_h + u_w)$, so maximizing $U_h + U_w$ is equivalent to maximizing $\mathcal{S} = u_h + u_w$. Although Stage 2 equilibrium childcare $l_i^{c*}(\lambda)$ depends

[2] Theorem 3 does not require the 0.5 specification nor symmetry of λ between a husband and wife; the Cobb–Douglas assumption for preferences and child quality is sufficient. These are standard in the literature, making the theorem robust.

on λ, substituting it back into $\mathcal{S}$ yields λ-dependent terms that are additive constants with respect to (l_h^w, l_w^w)—specifically, $\ln l_i^{\ell*}$ and $\ln k^*$ contribute terms $\ln(2\mu_i/\mathcal{D}_i(\lambda))$ and $\ln(A_i(\lambda)/\mathcal{D}_i(\lambda))$, which depend on λ but not on l_i^w (see Appendix D for the full derivation). These terms therefore drop out of the first-order conditions $\partial\mathcal{S}/\partial l_i^w = 0$, so C_h, C_w, and all Stage-1 derivatives are λ-free. Thus

$$\psi_{hw}(\lambda) = \frac{A_h(\lambda)}{2\mu_h + A_h(\lambda)} \cdot \underbrace{\frac{C_h}{w_h} \cdot \frac{T + \frac{1}{a_w}}{1 + C_h + C_w}}_{\equiv \Gamma > 0}.$$

Differentiating

$$\frac{\partial \psi_{hw}}{\partial \lambda} = \frac{2\mu_h \eta_w}{(2\mu_h + A_h(\lambda))^2} \cdot \Gamma > 0.$$

Part 2: Monotonicity of $\kappa_i(\lambda)$.

From (27)

$$\kappa_i(\lambda) \equiv \frac{\partial l_i^{c*}}{\partial l_i^w} = -\frac{\eta_i + \lambda\eta_{-i}}{2\mu_i + \eta_i + \lambda\eta_{-i}}.$$

Differentiating

$$\frac{\partial \kappa_i}{\partial \lambda} = -\frac{2\mu_i \eta_{-i}}{(2\mu_i + \eta_i + \lambda\eta_{-i})^2} < 0.$$

Part 3: Monotonicity of $\mathcal{W}(\lambda)$.

This follows directly from Theorem 3.

D Analytical Solutions

This appendix presents the closed-form derivations of optimal labor supply and childcare time allocation in the biform game framework. The analytical solutions provide explicit expressions for equilibrium choices as functions of model parameters and state variables.

D.1 Computation of Optimal Solutions

I now derive the closed-form solutions assuming interior solutions for both l_h^w and l_w^w.

D.2 Stage 1 Labor Supply: Closed-Form Solution

Setup. For each $i \in \{h, w\}$, the time constraint is

$$l_i^w + l_i^\ell + l_i^c = T, \qquad \Rightarrow \qquad l_i^\ell = T - l_i^w - l_i^c.$$

The budget constraint with no savings is

$$x = w_h l_h^w + w_w l_w^w + Y,$$

where

$$Y \equiv y_h + y_w + \mathbb{I}[d_j = 1, z_j = 1] \cdot \widehat{\text{Child SSI}}_j,$$

and $\widehat{\text{Child SSI}}_j$ is determined prior to the first stage.

Child Quality (Public Good): Updated Specification of k. For $i \in \{h, w\}$, define

$$a_i \equiv g_i(z_j; t) > 0 \qquad \text{(with } a_h = a_w \equiv g(z_j; t) \text{ under common efficiency)}.$$

Child quality (public good) is given by

$$k = \sqrt{(a_h l_h^c + 1)(a_w l_w^c + 1)}.$$

Therefore,

$$\ln k = \frac{1}{2} \ln(a_h l_h^c + 1) + \frac{1}{2} \ln(a_w l_w^c + 1), \qquad \frac{\partial \ln k}{\partial l_i^c} = \frac{1}{2} \frac{a_i}{a_i l_i^c + 1}.$$

Felicity Utility (with $\underline{k}_i = 0$ Normalization).

$$u_i(x, l_i^\ell, k) = (1 - \mu_i - \eta_i) \ln x + \mu_i \ln l_i^\ell + \eta_i \ln k.$$

Super Utility with Caring.

$$U_i = u_i + \lambda(z_j)\, u_{-i}.$$

D.3 Stage 2 Childcare: Closed-Form Solution

Key Point of Stage 2. In Stage 2, l_h^w and l_w^w are predetermined from Stage 1, so consumption x is constant with respect to l_i^c. The only terms in U_i that depend on l_i^c are $\mu_i \ln(T - l_i^w - l_i^c)$ and $(\eta_i + \lambda(z_j)\eta_{-i}) \ln k$. Define

$$A_i(\lambda) \equiv \eta_i + \lambda(z_j)\eta_{-i}.$$

First-Order Condition (FOC).

$$
\begin{aligned}
0 = \frac{\partial U_i}{\partial l_i^c} &= \mu_i \frac{\partial}{\partial l_i^c} \ln(T - l_i^w - l_i^c) + A_i(\lambda) \frac{\partial}{\partial l_i^c} \ln k \\
&= \mu_i \left(\frac{-1}{T - l_i^w - l_i^c} \right) + A_i(\lambda) \left(\frac{1}{2} \frac{a_i}{a_i l_i^c + 1} \right).
\end{aligned}
$$

Rearranging,

$$
\begin{aligned}
\frac{A_i(\lambda(z_j))}{2} \frac{a_i}{a_i l_i^c + 1} &= \frac{\mu_i}{T - l_i^w - l_i^c} \\
2\mu_i(a_i l_i^c + 1) &= A_i(\lambda(z_j)) a_i (T - l_i^w - l_i^c) \\
2\mu_i a_i l_i^c + 2\mu_i &= A_i(\lambda(z_j)) a_i (T - l_i^w) - A_i(\lambda(z_j)) a_i l_i^c \\
a_i \big(2\mu_i + A_i(\lambda(z_j))\big) l_i^c &= A_i(\lambda(z_j)) a_i (T - l_i^w) - 2\mu_i.
\end{aligned}
$$

Therefore, the Stage 2 best response is given in closed form by

$$
l_i^{c*}(l_i^w, \lambda(z_j)) = \frac{A_i(\lambda(z_j))\, a_i\, (T - l_i^w) - 2\mu_i}{a_i \big(2\mu_i + A_i(\lambda(z_j))\big)} \qquad (i \in \{h, w\}). \tag{28}
$$

Derived Expressions for Later Use. Define

$$
\mathcal{D}_i(\lambda(z_j)) \equiv 2\mu_i + A_i(\lambda(z_j)) = 2\mu_i + \eta_i + \lambda(z_j)\eta_{-i}.
$$

Then

$$
\begin{aligned}
a_i l_i^{c*} + 1 &= a_i \cdot \frac{A_i(\lambda(z_j))(T - l_i^w) - \dfrac{2\mu_i}{a_i}}{\mathcal{D}_i(\lambda(z_j))} + 1 = \frac{A_i(\lambda(z_j))\big(a_i(T - l_i^w) + 1\big)}{\mathcal{D}_i(\lambda(z_j))}, \\
l_i^{\ell*} = T - l_i^w - l_i^{c*} &= \frac{2\mu_i}{\mathcal{D}_i(\lambda(z_j))} \left(T - l_i^w + \frac{1}{a_i} \right) = \frac{2\mu_i}{\mathcal{D}_i(\lambda(z_j))} \cdot \frac{a_i(T - l_i^w) + 1}{a_i}.
\end{aligned}
$$

D.4 Stage 1 Cooperative Solution

Simplification of Welfare Maximization. Since $U_h = u_h + \lambda(z_j) u_w$ and $U_w = u_w + \lambda(z_j) u_h$, I have

$$
U_h + U_w = (1 + \lambda(z_j))(u_h + u_w).
$$

For $\lambda(z_j) > -1$, I have $(1 + \lambda(z_j)) > 0$, so maximizing $U_h + U_w$ in Stage 1 is equivalent to maximizing $u_h + u_w$.

Stage 1 Objective Function. Define

$$W(l_h^w, l_w^w) \equiv u_h(x, l_h^{\ell *}, k^*) + u_w(x, l_w^{\ell *}, k^*).$$

Then

$$W = \underbrace{\left[(1-\mu_h-\eta_h)+(1-\mu_w-\eta_w)\right]}_{\alpha} \ln x + \mu_h \ln l_h^{\ell *} + \mu_w \ln l_w^{\ell *} + (\eta_h+\eta_w)\ln k^*,$$

where $\alpha \equiv 2-(\mu_h+\eta_h+\mu_w+\eta_w)$.

From the Stage 2 derived expressions,

$$\begin{aligned}
\ln l_i^{\ell *} &= \ln\left(\frac{2\mu_i}{\mathcal{D}_i(\lambda(z_j))}\right) + \ln\left(T - l_i^w + \frac{1}{a_i}\right) \\
&= \ln\left(\frac{2\mu_i}{\mathcal{D}_i(\lambda(z_j))}\right) + \ln\big(a_i(T-l_i^w)+1\big) - \ln a_i, \\
\ln k^* &= \frac{1}{2}\ln(a_h l_h^{c*}+1) + \frac{1}{2}\ln(a_w l_w^{c*}+1) \\
&= \frac{1}{2}\ln\left(\frac{A_h(\lambda(z_j))\big(a_h(T-l_h^w)+1\big)}{\mathcal{D}_h(\lambda(z_j))}\right) + \frac{1}{2}\ln\left(\frac{A_w(\lambda(z_j))\big(a_w(T-l_w^w)+1\big)}{\mathcal{D}_w(\lambda(z_j))}\right).
\end{aligned}$$

Extracting only the terms in W that depend on l_i^w,

$$W = \alpha \ln x + \left(\mu_h + \tfrac{\eta_h+\eta_w}{2}\right)\ln\big(a_h(T-l_h^w)+1\big) + \left(\mu_w + \tfrac{\eta_h+\eta_w}{2}\right)\ln\big(a_w(T-l_w^w)+1\big) + \text{(constants)}.$$

Define

$$C_h \equiv \frac{\mu_h + \frac{\eta_h+\eta_w}{2}}{\alpha}, \qquad C_w \equiv \frac{\mu_w + \frac{\eta_h+\eta_w}{2}}{\alpha}. \tag{29}$$

First-Order Conditions (FOC). Since $x = w_h l_h^w + w_w l_w^w + Y$,

$$\frac{\partial x}{\partial l_h^w} = w_h, \qquad \frac{\partial x}{\partial l_w^w} = w_w, \qquad \frac{\partial}{\partial l_i^w}\ln\big(a_i(T-l_i^w)+1\big) = -\frac{a_i}{a_i(T-l_i^w)+1}.$$

The FOCs are therefore

$$\begin{aligned}
0 = \frac{\partial W}{\partial l_h^w} &= \alpha\frac{w_h}{x} - \left(\mu_h + \tfrac{\eta_h+\eta_w}{2}\right)\frac{a_h}{a_h(T-l_h^w)+1}, \\
0 = \frac{\partial W}{\partial l_w^w} &= \alpha\frac{w_w}{x} - \left(\mu_w + \tfrac{\eta_h+\eta_w}{2}\right)\frac{a_w}{a_w(T-l_w^w)+1}.
\end{aligned}$$

Rearranging,

$$\alpha\frac{w_h}{x} = \left(\mu_h + \tfrac{\eta_h+\eta_w}{2}\right)\frac{a_h}{a_h(T-l_h^w)+1} \iff a_h(T-l_h^w)+1 = a_h C_h \frac{x}{w_h},$$
$$\alpha\frac{w_w}{x} = \left(\mu_w + \tfrac{\eta_h+\eta_w}{2}\right)\frac{a_w}{a_w(T-l_w^w)+1} \iff a_w(T-l_w^w)+1 = a_w C_w \frac{x}{w_w}.$$

Hence,

$$l_h^w = T + \frac{1}{a_h} - \frac{C_h}{w_h}x, \qquad l_w^w = T + \frac{1}{a_w} - \frac{C_w}{w_w}x. \tag{30}$$

Solving for x (Closed Form). Substituting equations (30) into $x = w_h l_h^w + w_w l_w^w + Y$,

$$\begin{aligned} x &= w_h\left(T + \frac{1}{a_h} - \frac{C_h}{w_h}x\right) + w_w\left(T + \frac{1}{a_w} - \frac{C_w}{w_w}x\right) + Y \\ &= w_h\left(T + \frac{1}{a_h}\right) + w_w\left(T + \frac{1}{a_w}\right) + Y - (C_h + C_w)x. \end{aligned}$$

Therefore,

$$x\left(1 + C_h + C_w\right) = w_h\left(T + \frac{1}{a_h}\right) + w_w\left(T + \frac{1}{a_w}\right) + Y,$$

which yields

$$x^* = \frac{w_h\left(T + \frac{1}{a_h}\right) + w_w\left(T + \frac{1}{a_w}\right) + Y}{1 + C_h + C_w}. \tag{31}$$

Stage 1 Labor Supply (Closed Form). Finally,

$$l_h^{w*} = T + \frac{1}{a_h} - \frac{C_h}{w_h}x^* = T + \frac{1}{a_h} - \frac{C_h}{w_h}\cdot\frac{w_h\left(T + \frac{1}{a_h}\right) + w_w\left(T + \frac{1}{a_w}\right) + Y}{1 + C_h + C_w}, \tag{32}$$

$$l_w^{w*} = T + \frac{1}{a_w} - \frac{C_w}{w_w}x^* = T + \frac{1}{a_w} - \frac{C_w}{w_w}\cdot\frac{w_h\left(T + \frac{1}{a_h}\right) + w_w\left(T + \frac{1}{a_w}\right) + Y}{1 + C_h + C_w}. \tag{33}$$

Summary of Constants.

$$\alpha = 2 - (\mu_h + \eta_h + \mu_w + \eta_w), \qquad C_h = \frac{\mu_h + \frac{\eta_h+\eta_w}{2}}{\alpha}, \qquad C_w = \frac{\mu_w + \frac{\eta_h+\eta_w}{2}}{\alpha}, \tag{34}$$

$$Y = y_h + y_w + \text{Child SSI}_j[z_j = 1, d_j = 1].$$

Since corner solutions are permitted, I project onto $[0, T]$

$$l_i^{w*} \leftarrow \min\{T, \max\{0, l_i^{w*}\}\}, \quad i = h, w.$$

D.5 Corner Solution Details

In practice, the Stage 2 solutions l_i^c and Stage 1 solutions l_i^w for $i = h, w$ may involve corner solutions. The closed-form derivations in Sections D.2–D.3 assumed interior solutions.

The solution procedure is as follows. First, assume l_i^c for $i = h, w$ are interior and derive l_i^{w*} for $i = h, w$. Given these l_i^{w*}, check whether the implied l_i^{c*} are indeed interior. If so, the solution is consistent. If not, the candidate is not an equilibrium since there would be an incentive to deviate.

Second, assume l_i^c for $i = h, w$ are at corner values (either 0 or $T - l_i^w$). Given these corner values, solve the Stage 1 optimization problem to obtain l_i^{w*} for $i = h, w$. Then re-solve for l_i^{c*} for $i = h, w$ and verify whether the solution coincides with the assumed corner values. This procedure is repeated for all households j.

E Moment Construction Details

This appendix provides the detailed construction of estimation moments summarized in Section 7. For each moment category, I describe the empirical construction from data, the model counterpart, and the role in identification.

The moments used for identification can be organized into several categories based on their role in the estimation. Table E1 presents the level and slope moments. These moments discipline the key structural parameters governing household time allocation decisions.

Level moments (L1 and L2) capture the average time allocations for childcare and market work, respectively. Specifically, level moment L1 targets the mean childcare time for fathers and mothers, $\mathbb{E}[l_h^c \mid z_j]$ and $\mathbb{E}[l_w^c \mid z_j]$, computed separately by child disability status z_j. These moments primarily discipline the technology shifter $g_i(z_j)$ in the child quality production function (see equation (3)), while also interacting with preference heterogeneity parameters μ_i and η_i. Level moment L2 targets mean market labor supply conditional on employment, $\mathbb{E}[l_h^w \mid z_j, l_i^w > 0]$ and $\mathbb{E}[l_w^w \mid z_j, l_i^w > 0]$. These moments discipline the resource channel in Stage 1 (Section 3.2), including income Y, and help stabilize the interpretation of childcare level differences across disability status.

Slope moments (S1 and S2) capture how time allocations respond to changes in other endogenous variables, providing key variation for identifying the cooperation parameter $\lambda(z_j)$. Slope moment S1 measures the "childcare–labor" slope, defined as $\kappa_i \equiv \partial\, \mathbb{E}[l_i^c \mid X_j]/\partial l_i^w$. This is constructed through an auxiliary regression where, for each $z \in \{0, 1\}$, childcare time $l_{i,j}^c$ is regressed on own labor supply $l_{i,j}^w$, spouse's labor supply $l_{-i,j}^w$, and controls X_j. The coefficient $\widehat{\kappa_i}(z)$ on own labor supply serves as the empirical moment. In the model, the theoretical counterpart $\kappa_i^{\text{model}}(\theta) = \partial l_i^{c*}/\partial l_i^w$ has a closed-form expression under the log-Cobb–Douglas specification, with $\partial \kappa_i^{\text{model}}/\partial \lambda < 0$ as established in Proposition 4. This monotonicity property is central to identifying $\lambda(z_j)$, as it ensures that the Stage 2 cooperation parameter directly affects how childcare responds to changes in labor supply.

Slope moment S2 captures the cross-wage response of childcare, defined as $\psi_{i,-i} \equiv \partial l_i^{c*} / \partial w_{-i}$. This is constructed through an auxiliary regression on the dual-earner sample (both $l_h^w > 0$ and $l_w^w > 0$), where childcare time is regressed on spouse's wage $w_{-i,j}$, own wage $w_{i,j}$, and controls X_j. The coefficient $\widehat{b}_{i,-i}(z)$ on spouse's wage provides additional identification of $\lambda(z_j)$ through the monotonicity property $\lambda \mapsto \psi_{i,-i}^{\text{model}}(\lambda)$ (Proposition 4). This moment disciplines how wage shocks propagate to childcare allocation through Stage 2 cooperation.

Table E1: Moments for Identification: Level, Slope, and Auxiliary Shape Constraints

Category	Observed Moment	Construction from Data	Model Counterpart	Role in Identification/Estimation	Data Source
Level (L1)	Mean childcare time (father, mother): $\mathbb{E}[l_h^c \mid z_j]$, $\mathbb{E}[l_w^c \mid z_j]$	Compute mean by z_j (and age t if needed). Use time-diary-based childcare time when available	$\mathbb{E}[l_h^{c*}(\ell^{w*}, \lambda(z_j), g_i(z_j)) \mid z_j]$, $\mathbb{E}[l_w^{c*}(\ell^{w*}, \lambda(z_j), g_i(z_j)) \mid z_j]$	Primarily disciplines technology shifter $g_i(z_j)$ (see eq. (3)); also interacts with μ_i, η_i	(M2)
Level (L2)	Mean market labor (father, mother): $\mathbb{E}[l_h^w \mid z_j, l_i^w > 0]$, $\mathbb{E}[l_w^w \mid z_j, l_i^w > 0]$	Compute mean by z_j for employed sample ($l_i^w > 0$); use annual hours or weeks×hours; align timing with wage measurement	$\mathbb{E}[l_h^{w*}(w_h, w_w, Y(z_j)) \mid z_j]$, $\mathbb{E}[l_w^{w*}(w_h, w_w, Y(z_j)) \mid z_j]$	Disciplines resource channel in Stage 1 (Section 3.2), including Y; stabilizes interpretation of childcare level differences	(M1)
Slope (S1)	Childcare–labor slope (Stage 2 sensitivity): $\kappa_i \equiv \partial \mathbb{E}[l_i^c \mid X_j]/\partial l_i^w$	Auxiliary regression (Aux-κ): split sample by $z_j \in \{0,1\}$; for each z, regress $l_{i,j}^c$ on $(l_{i,j}^w, l_{-i,j}^w, X_j)$; store coefficient $\widehat{\kappa}_i(z)$ on $l_{i,j}^w$	$\kappa_i^{\text{model}}(\theta) = \partial l_i^{c*}/\partial l_i^w$ (closed form under log-Cobb–Douglas; $\partial \kappa_i^{\text{model}}/\partial \lambda < 0$, Proposition 4)	Core identification of $\lambda(z_j)$: exploits monotonicity $\lambda \mapsto \kappa_i^{\text{model}}(\lambda)$ (sign restriction from Proposition 4) to stabilize estimation; directly captures Stage 2 cooperation	(M2)
Slope (S2)	Cross-wage response of childcare: $\psi_{i,-i} \equiv \partial l_i^{c*}/\partial w_{-i}$	Auxiliary regression (Aux-A): dual-earner sample ($l_h^w > 0$ and $l_w^w > 0$); split by $z_j \in \{0,1\}$; for each z, regress $l_{i,j}^c$ on $(w_{-i,j}, w_{i,j}, X_j)$; store coefficient $\widehat{b}_{i,-i}(z)$ on $w_{-i,j}$	$\psi_{i,-i}^{\text{model}}(\theta)$ (equivalently $b_{i,-i}^{\text{model}}(\theta)$)	Additional $\lambda(z_j)$ identification: exploits monotonicity $\lambda \mapsto \psi_{i,-i}^{\text{model}}(\lambda)$ (Proposition 4); disciplines how wage shocks propagate to childcare through Stage 2 cooperation	(M2)

Note: Simulated data can be generated under given parameter values once household characteristics are specified: gender, education of each spouse, age of each spouse (wage equation), non-labor income, disabled child indicator, and state of residence. Actual data to match are distributed across SIPP, PSID, PSID CDS for each moment.

Building on these behavioral moments, Table E2 presents the wage-related identification moments. The wage level moment (W1) targets the mean log wage for fathers and mothers conditional on employment, $\mathbb{E}[\ln w_h \mid z_j, l_i^w > 0]$ and $\mathbb{E}[\ln w_w \mid z_j, l_i^w > 0]$. These moments are constructed by computing log wages for the employed sample, stratified by z_j and other characteristics such as education and child age. Care is taken to align the timing of wage measurement, apply consistent deflation, and trim outliers uniformly across samples. The model counterpart $\mathbb{E}[\ln w_i(\cdot) \mid z_j]$ combines the wage equation with heterogeneity. These moments discipline the level and scale of the wage distribution, providing the foundation for the price channel in labor supply decisions.

Table E2: Wage Identification Moments: Level, Slope, and Auxiliary Shape Constraints

Category	Observed Moment	Construction from Data	Model Counterpart	Role in Identification/Estimation	Data Source
Level (W1)	Mean log wage (father, mother): $\mathbb{E}[\ln w_h \mid z_j, l_i^w > 0]$, $\mathbb{E}[\ln w_w \mid z_j, l_i^w > 0]$	Compute $\ln w_i$ for employed sample ($l_i^w > 0$); stratify by z_j (education, child age, etc.); align timing, apply deflation, trim outliers consistently	$\mathbb{E}[\ln w_i(\cdot) \mid z_j]$ (wage equation + heterogeneity)	Disciplines level and scale of wage distribution (intercept and variance of wage equation); provides foundation for price channel in labor supply	(M1)

Table E3 presents the moments used to identify the covariance structure of the heterogeneity distribution. These second moments of observed outcomes—variances and covariances of wages, labor supply, and childcare time—provide the variation needed to pin down the joint distribution of unobserved heterogeneity (μ_i, η_i, ζ_i).

The variance moments (A1–A3) discipline the scale of each heterogeneity component. Variance moment A1 targets $\mathrm{Var}(\widetilde{\ln w_i} \mid z_j, l_i^w > 0)$, where $\widetilde{\ln w_i} \equiv \ln w_i - \widehat{\mathbb{E}}[\ln w_i \mid X_i]$ denotes the residual log wage after partialling out observables (e.g., via a Mincer-type regression on X_i). This moment primarily disciplines $\mathrm{Var}(\zeta_i)$, the scale of wage heterogeneity. Variance moment A2 targets $\mathrm{Var}(\widetilde{l_i^w} \mid z_j)$ for residual labor supply, disciplining $\mathrm{Var}(\mu_i)$, the scale of labor supply preference heterogeneity. Variance moment A3 targets $\mathrm{Var}(\widetilde{l_i^c} \mid z_j)$ for residual childcare time, disciplining $\mathrm{Var}(\eta_i)$, the scale of childcare preference heterogeneity. All variance moments are computed separately for $z_j \in \{0, 1\}$.

The within-individual covariance moments (B1–B3) discipline the correlations among heterogeneity components. Covariance moment B1 targets $\mathrm{Cov}(\widetilde{\ln w_i}, \widetilde{l_i^w} \mid z_j, l_i^w > 0)$, which primarily identifies $\mathrm{Cov}(\zeta_i, \mu_i)$—the correlation between wage heterogeneity and labor supply preferences that underlies wage endogeneity. Covariance moment B2 targets $\mathrm{Cov}(\widetilde{\ln w_i}, \widetilde{l_i^c} \mid z_j, l_i^w > 0)$, identifying $\mathrm{Cov}(\zeta_i, \eta_i)$—the correlation between wage heterogeneity and childcare preferences that aids causal interpretation of cross-wage effects. Covariance moment B3 targets $\mathrm{Cov}(\widetilde{l_i^w}, \widetilde{l_i^c} \mid z_j)$, identifying $\mathrm{Cov}(\mu_i, \eta_i)$—the correlation between labor and childcare preferences that reduces the risk of confounding κ with λ in identification.

The within-couple cross-covariance moments (C1–C3) discipline spousal correlations in heterogeneity, capturing assortative matching and common household shocks. Covariance moment C1 targets $\mathrm{Cov}(\widetilde{\ln w_h}, \widetilde{\ln w_w} \mid z_j, l_h^w > 0, l_w^w > 0)$ for dual-earner couples, identifying $\mathrm{Cov}(\zeta_h, \zeta_w)$. Covariance moment C2 targets both $\mathrm{Cov}(\widetilde{l_h^w}, \widetilde{l_w^w} \mid z_j)$ and $\mathrm{Cov}(\widetilde{l_h^c}, \widetilde{l_w^c} \mid z_j)$, identifying $\mathrm{Cov}(\mu_h, \mu_w)$ and $\mathrm{Cov}(\eta_h, \eta_w)$. Covariance moment C3 targets cross-covariances such as $\mathrm{Cov}(\widetilde{\ln w_h}, \widetilde{l_w^w} \mid z_j, l_h^w > 0)$ and $\mathrm{Cov}(\widetilde{\ln w_h}, \widetilde{l_w^c} \mid z_j, l_h^w > 0)$ (and symmetrically for $h \leftrightarrow w$), identifying $\mathrm{Cov}(\zeta_h, \mu_w)$, $\mathrm{Cov}(\zeta_h, \eta_w)$, and related terms. These moments flexibly accommodate common unobservables within couples and absorb spurious correlations that might otherwise contaminate cross-wage effects and time allocation patterns.

Table E3: Moments for Identifying the Covariance Matrix of Heterogeneity Distribution: Second Moments of Observed Outcomes (μ_i, η_i, ζ_i)

Category	Observed Moment	Construction from Data	Model Counterpart	Role in Identification/Estimation	Data Source
(A) Within-individual: Variance of outcomes (scale) (all computed separately for $z_j \in \{0,1\}$)					
Variance (A1)	Var($\ln w_i$ z_j=0, l_i^w>0), Var($\widetilde{\ln w_i}$ z_j=1, l_i^w>0)	Split employed sample ($l_i^w > 0$) by $z_j \in \{0,1\}$. For each z, subtract conditional mean: $\widetilde{\ln w_i} \equiv \ln w_i - \widehat{\mathbb{E}}[\ln w_i \mid X_i]$ (e.g., Mincer-type X_i). Compute variance	Primarily Var(ζ_i) (+ measurement error)	Disciplines scale of wage error (ζ) distribution; enables reading Cov($\ln w, l$) as "wage error × preference correlation"	(M1)
Variance (A2)	Var($\widetilde{l}_i^w \mid z_j$=0), Var($\widetilde{l}_i^w \mid z_j$=1)	Split sample by $z_j \in \{0,1\}$. For each z, subtract conditional mean: $\widetilde{l}_i^w \equiv l_i^w - \widehat{\mathbb{E}}[l_i^w \mid X]$. Compute variance	Primarily Var(μ_i) (+ measurement error)	Disciplines scale of labor supply preference heterogeneity (μ)	(M1)
Variance (A3)	Var($\widetilde{l}_i^c \mid z_j$=0), Var($\widetilde{l}_i^c \mid z_j$=1)	Split sample by $z_j \in \{0,1\}$. For each z, subtract conditional mean: $\widetilde{l}_i^c \equiv l_i^c - \widehat{\mathbb{E}}[l_i^c \mid X]$. Compute variance	Primarily Var(η_i) (+ measurement error)	Disciplines scale of childcare preference heterogeneity (η)	(M2)
(B) Within-individual: Covariance across outcomes (identifies Cov(μ, η), Cov(ζ, μ), Cov(ζ, η)) (all computed separately for $z_j \in \{0,1\}$)					
Covariance (B1)	Cov($\widetilde{\ln w_i}, \widetilde{l}_i^w$ z_j=0, l_i^w>0), Cov($\widetilde{\ln w_i}, \widetilde{l}_i^w$ z_j=1, l_i^w>0)	Split employed sample ($l_i^w > 0$) by $z_j \in \{0,1\}$; for each z, compute Cov($\widetilde{\ln w_i}, \widetilde{l}_i^w$)	Primarily Cov(ζ_i, μ_i)	Disciplines correlation between wage error and labor preference (source of wage endogeneity)	(M1)
Covariance (B2)	Cov($\widetilde{\ln w_i}, \widetilde{l}_i^c$ z_j=0, l_i^w>0), Cov($\widetilde{\ln w_i}, \widetilde{l}_i^c$ z_j=1, l_i^w>0)	Split employed sample ($l_i^w > 0$) by $z_j \in \{0,1\}$; for each z, compute Cov($\widetilde{\ln w_i}, \widetilde{l}_i^c$)	Primarily Cov(ζ_i, η_i)	Disciplines correlation between wage error and childcare preference (aids causal interpretation of cross-wage effects)	(M2)
Covariance (B3)	Cov($\widetilde{l}_i^w, \widetilde{l}_i^c \mid z_j$=0), Cov($\widetilde{l}_i^w, \widetilde{l}_i^c \mid z_j$=1)	Split by $z_j \in \{0,1\}$; for each z, compute Cov($\widetilde{l}_i^w, \widetilde{l}_i^c$)	Primarily Cov(μ_i, η_i)	Disciplines labor–childcare preference correlation (reduces risk of confounding κ with λ)	(M2)
(C) Within-couple: Cross-covariance (matching / common environment) (all computed separately for $z_j \in \{0,1\}$)					
Covariance (C1)	Cov($\widetilde{\ln w_h}, \widetilde{\ln w_w}$ z_j=0, l_h^w>0, l_w^w>0), Cov($\widetilde{\ln w_h}, \widetilde{\ln w_w}$ z_j=1, l_h^w>0, l_w^w>0)	Dual-earner sample ($l_h^w > 0$ and $l_w^w > 0$); split by $z_j \in \{0,1\}$; for each z, compute Cov($\widetilde{\ln w_h}, \widetilde{\ln w_w}$)	Primarily Cov(ζ_h, ζ_w)	Disciplines spousal wage error correlation (assortative matching / common shocks)	(M1)
Covariance (C2)	Cov($\widetilde{l}_h^w, \widetilde{l}_w^w \mid z_j$=0), Cov($\widetilde{l}_h^w, \widetilde{l}_w^w \mid z_j$=1), Cov($\widetilde{l}_h^c, \widetilde{l}_w^c \mid z_j$=0), Cov($\widetilde{l}_h^c, \widetilde{l}_w^c \mid z_j$=1)	Split by $z_j \in \{0,1\}$; for each z, compute Cov($\widetilde{l}_h^w, \widetilde{l}_w^w$), Cov($\widetilde{l}_h^c, \widetilde{l}_w^c$)	Primarily Cov(μ_h, μ_w), Cov(η_h, η_w)	Disciplines spousal preference correlation (matching)	(M1)
Covariance (C3)	Spousal cross: Cov($\widetilde{\ln w_h}, \widetilde{l}_w^w$ z_j=0, l_h^w>0), Cov($\widetilde{\ln w_h}, \widetilde{l}_w^w$ z_j=1, l_h^w>0), Cov($\widetilde{\ln w_h}, \widetilde{l}_w^c$ z_j=0, l_h^w>0), Cov($\widetilde{\ln w_h}, \widetilde{l}_w^c$ z_j=1, l_h^w>0) (symmetric for $h \leftrightarrow w$)	Wage-side employed sample (e.g., $l_h^w > 0$); split by $z_j \in \{0,1\}$; for each z, compute Cov($\widetilde{\ln w_h}, \widetilde{l}_w^w$), Cov($\widetilde{\ln w_h}, \widetilde{l}_w^c$)	Primarily Cov(ζ_h, μ_w), Cov(ζ_h, η_w), etc.	Flexibly accommodates common unobservables within couples; absorbs spurious correlations in cross-wage and allocation patterns	(M1)/(M2)

Note: This study primarily uses pooled repeated cross-sections from multiple years and datasets; the unit of observation is household-year. $\widetilde{\cdot}$ denotes the centered variable after subtracting the conditional mean (e.g., on X). The second moments in this table are computed separately for $z_j \in \{0,1\}$, yielding $\{\widehat{\text{Var}}(\cdot \mid z), \widehat{\text{Cov}}(\cdot,\cdot \mid z)\}_{z=0,1}$, which serve as backup discipline for the model-implied Var(μ_i), Var(η_i), Var(ζ_i) and associated covariances.

Finally, Table E4 presents the moments related to subjective well-being (SWB) thresholds, which address response shift in reporting scales. At this stage, the structural parameters θ have already been estimated (i.e., $\hat{\theta}$ is given), and these moments serve to calibrate the mapping from model-implied utility to observed SWB categories rather than to identify θ itself.

Threshold moment T1 targets the cutpoints $\widehat{\tau}_{k,z=0}$ for $k = 1, \ldots, K-1$ estimated from an ordered probit or ordered logit on the reference group ($z = 0$). These thresholds establish the baseline scale for mapping the utility $U_i(\hat{\theta}, \varepsilon)$ implied by equation (5) to observed SWB response categories. Threshold moment T2 targets the cutpoints $\widehat{\tau}_{k,z=1}$ for the comparison group ($z = 1$). By allowing the thresholds to differ across z, the model accommodates response shift effects—systematic differences in reporting standards between households with and without a disabled child. This separation ensures that "differences in utility levels" and "differences in reporting scales" are not confounded. As a normalization, the ordered probit imposes $\mathrm{Var}(\varepsilon) = 1$ for scale identification.

Table E4: SWB Threshold Moments: Response Shift Effect and Cutpoint Auxiliary Statistics

Category	Observed Moment	Construction from Data	Model Counterpart	Role in Identification/Estimation	Data Source
Auxiliary (T1)	**Reference group thresholds** (Section 5.4): $\widehat{\tau}_{k,\, z=0}$ ($k = 1, \ldots, K-1$)	Estimate ordered probit/logit on z=0 sample; store cutpoints $\widehat{\tau}_{k,z=0}$. (Condition on X if needed)	$\tau^{\text{model}}_{k,\, z=0}$	**Fixes the scale (response shift) baseline**. Disciplines cutpoints for mapping model-implied $U_i(\hat{\theta}, \varepsilon)$ to observed SWB categories	SWB data (PSID, etc.)
Auxiliary (T2)	**Comparison group thresholds**: $\widehat{\tau}_{k,\, z=1}$ ($k = 1, \ldots, K-1$)	Estimate ordered probit/logit on z=1 sample; store $\widehat{\tau}_{k,z=1}$	$\tau^{\text{model}}_{k,\, z=1}$	Recovers **response shift (reporting standard shift)** by z. Allows SWB cutpoints to differ by z, separating utility level differences from reporting scale differences	SWB data (PSID, etc.)

Note: At this stage, the structural parameters θ have already been estimated ($\hat{\theta}$ given); the moments in this table serve to **calibrate the SWB measure (response shift)** rather than to identify θ. As a normalization, the ordered probit imposes $\mathrm{Var}(\varepsilon) = 1$ (scale normalization). The approach of estimating (T1) and (T2) separately by z is not the only option; following Friedberg and Stern (2014), one could integrate these into a single likelihood and estimate jointly with the main model.

F Detailed Moment Statistics

This appendix presents the full set of empirical moments used for structural estimation, organized by type: level moments (L1, L2), slope moments (S1, S2), wage moments (W1, W2), and covariance moments. All moments are stratified by child disability status ($z = 0$ for non-disabled, $z = 1$ for disabled) to facilitate identification of disability-specific parameters.

F.1 Level and Slope Moments

Table F1 reports level moments for childcare time (L1) and labor supply (L2). The childcare time moments reveal the asymmetric response documented above: husbands reduce childcare time by 3.16 hours when the child is disabled, while wives increase it by 2.62 hours. Conditional on employment, labor supply differences are modest and statistically insignificant. Employment rates show a pronounced gap for both spouses: husbands' employment rate declines by 4.4 percentage points (87.0% to 82.6%) and wives' by 4.8 percentage points (70.0% to 65.2%) in SIPP for households with disabled children.[3]

Table F2 presents slope moments that capture behavioral responses central to identification. The coefficient $\kappa_i \equiv \partial l_i^c / \partial l_i^w$ measures how spouse i's childcare time responds to their own labor supply, while $\psi_{i,-i} \equiv \partial l_i^c / \partial w_{-i}$ captures the response to the other spouse's wage. These slopes are derived from auxiliary regressions and serve as key identifying moments for the cooperation parameter $\lambda(z_j)$.

Table F3 reports wage moments for employed couples. Both mean log wages and their variances are similar across disability status, with no statistically significant differences. Households in which either spouse reports an hourly wage above \$100 or below \$1 are excluded from the sample.

Table F4 presents the full covariance structure of the key variables, combining SIPP and PSID-CDS via inverse-variance weighting. Several patterns are noteworthy. First, the between-spouse covariance of labor supply $\text{Cov}(l_h^w, l_w^w)$ decreases from +7.8 to +1.1 when moving from non-disabled to disabled child households, suggesting a shift toward specialization. Second, the between-spouse childcare covariance $\text{Cov}(l_h^c, l_w^c)$ is negative in both groups but less negative for disabled-child households (−43.7 vs. −71.0), indicating reduced substitutability or increased complementarity in parental care.

Table F5 summarizes the implications of the slope moments for identifying changes in λ. Moments with large standard errors in the data receive lower weights in estimation through the optimal weighting matrix, so imprecise moments do not unduly influence the estimates.

An important caveat is that the empirical values of κ and ψ need not exactly match the theoretical predictions from Section 3.3. The auxiliary regression model is deliberately simple and potentially misspecified. Moreover, when comparing $z = 0$ and $z = 1$ households, not only λ but also $g_i(z_j : t)$ and

[3] The employment rate of fathers is lower than that reported in Del Boca et al. (2014).

Table F1: Level Moments (L1, L2)

Variable	$z = 0$		$z = 1$		$\Delta(z1 - z0)$		Data
	Value	SE	Value	SE	Value	SE	
L1: Childcare time $E[l_i^c]$ (weekly hours)							
$E[l_h^c]$ husband	18.79	0.57	15.63	0.83	-3.16	1.01	PSID-CDS
$E[l_w^c]$ wife	18.16	0.45	20.78	0.78	+2.62	0.90	PSID-CDS
L2: Labor supply $E[l_i^w \mid l_i^w > 0]$ (weekly hours, workers only)							
$E[l_h^w \mid l_h^w > 0]$ husband	44.49	0.13	44.59	0.50	+0.10	0.52	SIPP
$E[l_w^w \mid l_w^w > 0]$ wife	37.89	0.14	37.75	0.55	-0.14	0.57	SIPP
$E[l_h^w \mid l_h^w > 0]$ husband	38.97	0.38	38.25	0.76	-0.72	0.85	PSID
$E[l_w^w \mid l_w^w > 0]$ wife	32.16	0.41	32.55	0.79	+0.39	0.89	PSID
L2b: Employment rate $Pr(employed)$ (%)[a]							
$Pr(employed_h)$ husband	87.0	—	82.6	—	-4.4	—	SIPP
$Pr(employed_w)$ wife	70.0	—	65.2	—	-4.8	—	SIPP
$Pr(l_h^w > 0)$ husband	80.6	—	80.9	—	+0.3	—	PSID
$Pr(l_w^w > 0)$ wife	69.0	—	68.3	—	-0.7	—	PSID

[a] SIPP: Non-employed = hours = 0 AND earnings = 0. PSID: Non-employed = hours = 0.

$F^z_{\mu,\eta}$ differ. This point is crucial and is echoed in the discussion of estimation results.

The large standard errors on certain slope moments warrant explanation. For ψ_{hw} (husband's childcare response to wife's wage), the wife's employment rate is relatively low (approximately 65–70% in PSID-CDS). Since wages can only be computed for employed wives, the effective estimation sample is restricted to employed-wife observations. This substantially reduces the sample size from the full PSID-CDS sample ($z = 0$: approximately 4,100 household-wave observations; $z = 1$: approximately 450 household-wave observations) to employed wives only.

In contrast, ψ_{wh} (wife's childcare response to husband's wage) has somewhat smaller standard errors because the husband's employment rate is higher (approximately 80–85% in PSID-CDS), yielding a larger sample of observations with computable wages.

Standard errors are particularly large for the $z = 1$ (disabled child) subsample because the sample size itself is small (approximately 450 in PSID-CDS). Adding the employment condition further reduces estimation precision. With infinite sample size, the estimates would conform to theoretical predictions; the observed deviations reflect finite-sample limitations inherent in studying a relatively rare population.

Table F2: Slope Moments (S1, S2)

Variable	$z = 0$		$z = 1$		$\Delta(z1 - z0)$		Data
	Value	SE	Value	SE	Value	SE	
S1: $\kappa_i = \partial l_i^c / \partial l_i^w$							
κ_h husband	-0.0253	0.0320	+0.0821	0.0434	+0.1074	0.0539	PSID-CDS
κ_w wife	-0.0146	0.0252	-0.0221	0.0451	-0.0075	0.0517	PSID-CDS
S2: $\psi_{i,-i} = \partial l_i^c / \partial w_{-i}$ **(hourly wage,** $\times$**1000)**							
ψ_{hw} husband	+0.060	0.031	-0.013	0.036	-0.073	0.048	PSID-CDS
ψ_{wh} wife	+0.013	0.011	+0.034	0.024	+0.020	0.026	PSID-CDS

Table F3: Wage Moments (W1, W2) — Employed Couples Only

Variable	$z = 0$		$z = 1$		$\Delta(z1 - z0)$		Data
	Value	SE	Value	SE	Value	SE	
W1: Mean log wage $E[\ln(w_i)]$							
$E[\ln(w_h)]$ husband	2.963	0.008	2.983	0.022	+0.020	0.023	SIPP + PSID
$E[\ln(w_w)]$ wife	2.739	0.008	2.779	0.025	+0.040	0.026	SIPP + PSID
W2: Variance of log wage $\text{Var}[\ln(w_i)]$							
$\text{Var}[\ln(w_h)]$ husband	0.356	0.007	0.353	0.025	-0.003	0.026	SIPP + PSID
$\text{Var}[\ln(w_w)]$ wife	0.382	0.008	0.384	0.023	+0.002	0.024	SIPP + PSID

Sample: Employed couples only (both husband and wife have hours $>$ 0), hourly wage \$1–\$100.

G Detailed Model Fit Tables

This appendix presents detailed comparisons of data and simulated moments. Table G1 shows the basic moments, Table G2 shows the behavioral slopes (κ, ψ), Table G3 shows the standard deviations, Table G4 shows the covariances, and Table G5 shows the wage moments. Table G6 summarizes the overall fit. The fit quality is assessed using the standardized difference, defined as (Data − Sim)/SE(Data): ⊛ for excellent fit (|Std. Diff| $<$ 2), ◯ for good fit (2 $\leq$ |Std. Diff| $<$ 5), and △ for moderate fit (|Std. Diff| $\geq$ 5).

Table F4: Covariance Matrix Moments

Variable	$z = 0$		$z = 1$		$\Delta(z1 - z0)$		Data
	Value	SE	Value	SE	Value	SE	
Section A: Variance Var(X)[a]							
Var(l_h^w) husband labor	148.2	5.9	166.7	14.0	+18.5	15.2	SIPP + PSID
Var(l_w^w) wife labor	126.0	4.2	156.0	11.7	+30.0	12.4	SIPP + PSID
Var($\ln w_h$) husband log wage	0.356	0.007	0.353	0.023	-0.003	0.024	SIPP + PSID
Var($\ln w_w$) wife log wage	0.382	0.008	0.384	0.024	+0.002	0.025	SIPP + PSID
Var(l_h^c) husband childcare[b]	266.4	13.2	241.0	24.3	-25.5	27.6	PSID
Var(l_w^c) wife childcare[b]	191.3	10.7	232.3	20.5	+41.1	23.1	PSID
Section B: Within-person covariance Cov(X, Y)							
Cov($l_h^w, \ln w_h$) husband	-0.46	0.12	-0.18	0.41	+0.28	0.43	SIPP + PSID
Cov($l_w^w, \ln w_w$) wife	0.42	0.11	0.49	0.35	+0.07	0.37	SIPP + PSID
Cov(l_h^c, l_h^w) husband	-3.6	10.1	-13.4	16.7	-9.8	19.5	PSID
Cov(l_w^c, l_w^w) wife	-6.2	7.5	+0.9	12.2	+7.1	14.3	PSID
Section C: Between-spouse covariance Cov(X_h, Y_w)							
Cov(l_h^w, l_w^w) labor	+7.8	2.5	+1.1	6.6	−6.7	7.1	SIPP + PSID
Cov($\ln w_h, \ln w_w$) log wage	0.136	0.006	0.092	0.016	-0.044	0.017	SIPP + PSID
Cov($l_h^w, \ln w_w$)	0.02	0.10	0.48	0.33	+0.46	0.34	SIPP + PSID
Cov($l_w^w, \ln w_h$)	−0.27	0.11	−0.83	0.33	−0.56	0.35	SIPP + PSID
Cov(l_h^c, l_w^c) childcare[b]	**-71.0**	6.9	**-43.7**	12.6	+27.3	14.4	PSID

[a] Labor variance/covariance: **employed couples only** (both husband_hours > 0 and wife_hours > 0).
[b] Childcare variance/covariance: all sample (non-missing childcare hours).

Table F5: Implications for λ Identification

Moment	$\Delta(z = 1 - z = 0)$	Theory	Implication for λ
κ_h	+0.1074 ($t = 1.99$)	$\partial\kappa/\partial\lambda < 0$	$\lambda \downarrow$
κ_w	−0.0075 ($t = -0.14$)	$\partial\kappa/\partial\lambda < 0$	$\lambda \uparrow$
ψ_{hw}	−0.0729 ($t = -1.52$)	$\partial\psi/\partial\lambda > 0$	$\lambda \downarrow$
ψ_{wh}	+0.0204 ($t = 0.81$)	$\partial\psi/\partial\lambda > 0$	$\lambda \uparrow$

Conclusion: κ_h is significant at 5% level ($t = 1.99$), suggesting λ is lower for households with disabled children. ψ_{hw} points in the same direction but is not significant.

Table G1: Model Fit: Basic Moments

Moment	Data	Sim	Diff	Std. Diff	Fit
Panel A: $z = 0$ **(Children without disabilities)**					
Employment (Husband)	0.870	0.872	0.002	0.63	⊛
Employment (Wife)	0.700	0.697	−0.003	−0.70	⊛
$E[\ell_w \| emp]$ Husband	44.49	44.63	0.14	1.08	⊛
$E[\ell_w \| emp]$ Wife	37.89	37.88	−0.01	−0.07	⊛
$E[\ell_c]$ Husband	18.79	17.66	−1.13	−1.98	⊛
$E[\ell_c]$ Wife	18.16	18.41	0.25	0.56	⊛
Panel B: $z = 1$ **(Children with disabilities)**					
Employment (Husband)	0.826	0.831	0.005	0.38	⊛
Employment (Wife)	0.652	0.660	0.008	0.52	⊛
$E[\ell_w \| emp]$ Husband	44.59	44.50	−0.09	−0.17	⊛
$E[\ell_w \| emp]$ Wife	37.75	37.26	−0.49	−0.86	⊛
$E[\ell_c]$ Husband	15.63	15.61	−0.02	−0.02	⊛
$E[\ell_c]$ Wife	20.78	20.87	0.09	0.12	⊛

Notes: Std. Diff ≡ (Data − Sim)/SE(Data). ⊛: |Std. Diff| < 2 (excellent),
◯: 2 ≤ |Std. Diff| < 5 (good), △: |Std. Diff| ≥ 5 (moderate).

Table G2: Model Fit: Behavioral Slopes

Moment	Data	Sim	Diff	Std. Diff	Fit
κ**: Caregiver Burden (labor → childcare)**					
κ_h $(z = 0)$	−0.025	−0.056	−0.031	−0.97	⊛
κ_w $(z = 0)$	−0.015	−0.009	0.006	0.22	⊛
κ_h $(z = 1)$	0.082	0.074	−0.008	−0.19	⊛
κ_w $(z = 1)$	−0.022	−0.058	−0.036	−0.80	⊛
ψ**: Childcare–Wage Complementarity** (×1000)					
ψ_{hw} $(z = 0)$	0.060	0.056	−0.004	−0.07	⊛
ψ_{wh} $(z = 0)$	0.013	0.039	0.026	1.12	⊛
ψ_{hw} $(z = 1)$	−0.013	0.012	0.025	0.29	⊛
ψ_{wh} $(z = 1)$	0.034	0.050	0.016	0.31	⊛

Notes: $\kappa_i = \partial l_i^c / \partial l_i^w$; $\psi_{i,-i} = \partial l_i^c / \partial w_{-i}$.

H Covariance Structure Details

This appendix provides the full details of the covariance restrictions and Cholesky parameterization summarized in Section 6.3.1.

Table G3: Model Fit: Standard Deviations

Moment	Data	Sim	Diff	Std. Diff	Fit
SD of Labor Hours (conditional on employment)					
$\sigma(\ell_w)$ Husband ($z = 0$)	12.21	21.12	8.91	36.61	△
$\sigma(\ell_w)$ Wife ($z = 0$)	11.42	20.36	8.94	45.01	△
$\sigma(\ell_w)$ Husband ($z = 1$)	12.89	22.24	9.35	15.76	△
$\sigma(\ell_w)$ Wife ($z = 1$)	12.39	20.82	8.43	20.08	△
SD of Childcare Hours					
$\sigma(\ell_c)$ Husband ($z = 0$)	16.32	16.70	0.38	0.90	⊛
$\sigma(\ell_c)$ Wife ($z = 0$)	13.83	13.21	−0.62	−1.50	⊛
$\sigma(\ell_c)$ Husband ($z = 1$)	15.52	15.84	0.32	0.46	⊛
$\sigma(\ell_c)$ Wife ($z = 1$)	15.24	15.98	0.74	1.07	⊛

Notes: Model over-predicts variance in labor hours, a common challenge in structural household models with corner solutions.

Table G4: Model Fit: Covariances

Moment	Data	Sim	Diff	Std. Diff	Fit
Within-Person: Cov(ℓ_c, ℓ_w)					
Husband ($z = 0$)	−3.6	−34.1	−30.5	−2.86	◯
Wife ($z = 0$)	−6.2	−5.6	0.7	0.10	⊛
Husband ($z = 1$)	−13.4	50.8	64.2	3.53	◯
Wife ($z = 1$)	0.9	−34.9	−35.8	−2.74	◯
Couple: Cov(ℓ_c^h, ℓ_c^w)					
$z = 0$	−71.0	72.1	143.1	17.65	△
$z = 1$	−43.7	2.0	45.7	3.63	◯
Couple: Cov(ℓ_w^h, ℓ_w^w)					
$z = 0$	+7.8	−320.7	−328.5	−147.34	△
$z = 1$	+1.1	−340.9	−342.0	−48.82	△

Notes: ⊛: |Std. Diff| < 2, ◯: 2 ≤ |Std. Diff| < 5, △: |Std. Diff| ≥ 5.

H.1 Zero Restrictions

I allow within-person correlations and cross-spouse same-type correlations:

- Within-person: $\text{Cov}^z(\mu_h, \eta_h)$, $\text{Cov}^z(\mu_h, \zeta_h)$, $\text{Cov}^z(\eta_h, \zeta_h)$ for the husband; $\text{Cov}^z(\mu_w, \eta_w)$, $\text{Cov}^z(\mu_w, \zeta_w)$, $\text{Cov}^z(\eta_w, \zeta_w)$ for the wife.
- Cross-spouse same-type: $\text{Cov}^z(\mu_h, \mu_w)$, $\text{Cov}^z(\eta_h, \eta_w)$, $\text{Cov}^z(\zeta_h, \zeta_w)$.

Table G5: Model Fit: Wage Moments

Moment	Data	Sim	Diff	Std. Diff	Fit
Mean Log Wage: $E[\ln w]$					
Husband ($z = 0$)	2.963	3.063	0.099	13.07	△
Husband ($z = 1$)	2.983	3.142	0.158	6.93	△
Wife ($z = 0$)	2.739	2.820	0.081	9.93	△
Wife ($z = 1$)	2.779	2.842	0.063	2.65	◯
Variance Log Wage: $\mathbf{Var}[\ln w]$					
Husband ($z = 0$)	0.356	0.452	0.096	12.27	△
Husband ($z = 1$)	0.353	0.456	0.101	4.59	◯
Wife ($z = 0$)	0.382	0.382	0.000	0.00	⊛
Wife ($z = 1$)	0.384	0.451	0.067	2.81	◯
Cov(Labor Hours, Log Wage)					
$\text{Cov}(\ell_w, \ln w)$ Husband ($z = 0$)	−0.46	4.73	5.19	42.90	△
$\text{Cov}(\ell_w, \ln w)$ Husband ($z = 1$)	−0.18	5.93	6.11	16.39	△
$\text{Cov}(\ell_w, \ln w)$ Wife ($z = 0$)	0.42	2.29	1.87	16.99	△
$\text{Cov}(\ell_w, \ln w)$ Wife ($z = 1$)	0.49	2.97	2.49	7.34	△
Cross-Spousal Wage Covariances					
$\text{Cov}(\ln w^h, \ln w^w)$ ($z = 0$)	0.136	0.106	−0.030	−4.63	◯
$\text{Cov}(\ln w^h, \ln w^w)$ ($z = 1$)	0.092	0.158	0.066	3.79	◯
$\text{Cov}(\ell_w^h, \ln w^w)$ ($z = 0$)	0.02	−0.22	−0.24	−2.38	◯
$\text{Cov}(\ell_w^h, \ln w^w)$ ($z = 1$)	0.48	1.46	0.98	2.71	◯
$\text{Cov}(\ell_w^w, \ln w^h)$ ($z = 0$)	−0.27	−1.41	−1.14	−11.81	△
$\text{Cov}(\ell_w^w, \ln w^h)$ ($z = 1$)	−0.83	−0.30	0.53	1.57	⊛

Notes: Wage equation: $\ln w_i = \nu_0 + \nu_1 \cdot \text{age} + \nu_2 \cdot \text{age}^2 + \nu_3 \cdot \text{edu} + \xi_i$.
⊛: |Std. Diff| < 2, ◯: 2 ≤ |Std. Diff| < 5, △: |Std. Diff| ≥ 5.

All other cross-spouse correlations are restricted to zero

$$\text{Cov}^z(\mu_h, \eta_w) = \text{Cov}^z(\eta_h, \mu_w) = \text{Cov}^z(\mu_h, \zeta_w) = \text{Cov}^z(\eta_h, \zeta_w) = \text{Cov}^z(\mu_w, \zeta_h) = \text{Cov}^z(\eta_w, \zeta_h) = 0. \quad (35)$$

H.2 Cholesky Parameterization

Let

$$\Sigma^z \equiv \text{Var}^z\Big(q_1^{\mu,\eta,\zeta}, q_2^{\mu,\eta,\zeta}, q_3^{\mu,\eta,\zeta}, q_4^{\mu,\eta,\zeta}, q_5^{\mu,\eta,\zeta}, q_6^{\mu,\eta,\zeta}\Big) = U^{z\prime} U^z, \quad (36)$$

where the ordering is $(\mu_h, \eta_h, \mu_w, \eta_w, \zeta_h, \zeta_w)$. The zero restrictions (35) translate into

$$\Sigma^z_{1,4} = \Sigma^z_{2,3} = \Sigma^z_{1,6} = \Sigma^z_{2,6} = \Sigma^z_{3,5} = \Sigma^z_{4,5} = 0. \quad (37)$$

Table G6: Summary of Moment Fit

Category	⊛ ($\lvert d\rvert < 2$)	◯ ($2 \le \lvert d\rvert < 5$)	△ ($\lvert d\rvert \ge 5$)	Total
Basic Moments ($z = 0$)	6	0	0	6
Basic Moments ($z = 1$)	6	0	0	6
κ (Caregiver Burden)	4	0	0	4
ψ (Childcare–Wage)	4	0	0	4
Standard Deviations	4	0	4	8
Covariances	1	4	3	8
Wage Moments	2	7	9	18
Total	**27**	**11**	**16**	**54**

$d \equiv (\text{Data} - \text{Sim})/\text{SE}(\text{Data})$. Per-moment weights: levels 10, κ 50 (×1.5 for $\kappa_h^{z=1}$), ψ 40 (×1.5 for ψ_{hw}), SD 0.25, Cov 0.1, Wage 0.2. These block-level scale weights are applied on top of inverse-variance normalization so that each moment block contributes approximately equally to the objective function. Resulting block contributions: first moments and slopes 254, SD 329, Cov 408, Wage 432 (total ≈ 1,423).

Let c_i^z for $i = 1, \ldots, 15$ denote the free parameters. The restricted upper-triangular Cholesky factor is

$$U^z_{\text{restricted}} = \begin{pmatrix} \exp(c_1^z) & c_2^z & c_3^z & 0 & c_4^z & 0 \\ 0 & \exp(c_5^z) & -\dfrac{c_2^z c_3^z}{\exp(c_5^z)} & c_6^z & c_7^z & 0 \\ 0 & 0 & \exp(c_8^z) & c_9^z & U^\dagger_{3,5} & c_{11}^z \\ 0 & 0 & 0 & \exp(c_{10}^z) & U^\dagger_{4,5} & c_{12}^z \\ 0 & 0 & 0 & 0 & \exp(c_{13}^z) & c_{14}^z \\ 0 & 0 & 0 & 0 & 0 & \exp(c_{15}^z) \end{pmatrix}, \tag{38}$$

where

$$U^\dagger_{3,5} = -\frac{c_3^z c_4^z - c_2^z c_3^z c_7^z/\exp(c_5^z)}{\exp(c_8^z)}, \qquad U^\dagger_{4,5} = -\frac{c_6^z c_7^z + c_9^z\, U^\dagger_{3,5}}{\exp(c_{10}^z)}.$$

The constraint $\Sigma^z_{1,4} = 0$ is satisfied by $U_{1,4} = 0$; $\Sigma^z_{1,6} = 0$ and $\Sigma^z_{2,6} = 0$ by $U_{1,6} = U_{2,6} = 0$. The constraint $\Sigma^z_{2,3} = 0$ requires $U_{1,2}U_{1,3} + U_{2,2}U_{2,3} = 0$, yielding $U_{2,3} = -c_2^z c_3^z/\exp(c_5^z)$. The constraint $\Sigma^z_{3,5} = 0$ requires $U_{1,3}U_{1,5} + U_{2,3}U_{2,5} + U_{3,3}U_{3,5} = 0$, yielding $U^\dagger_{3,5}$. Finally, $\Sigma^z_{4,5} = 0$ requires $U_{2,4}U_{2,5} + U_{3,4}U_{3,5} + U_{4,4}U_{4,5} = 0$, yielding $U^\dagger_{4,5}$.

H.3 Transformation to Preference Parameters

The vector $(q_1, \ldots, q_6)$ is drawn from $\mathcal{N}((\boldsymbol{\phi}^z, 0, 0), \Sigma^z)$. The preference and wage shock parameters are

$$\mu_h = \frac{\exp(q_1)}{1 + \exp(q_1) + \exp(q_2)}, \qquad \eta_h = \frac{\exp(q_2)}{1 + \exp(q_1) + \exp(q_2)}, \tag{39}$$

$$\mu_w = \frac{\exp(q_3)}{1 + \exp(q_3) + \exp(q_4)}, \qquad \eta_w = \frac{\exp(q_4)}{1 + \exp(q_3) + \exp(q_4)}, \tag{40}$$

$$\zeta_h = q_5, \qquad \zeta_w = q_6. \tag{41}$$

I Descriptive Statistics

This section presents descriptive statistics from the SIPP and PSID-CDS data used for estimation. The key variables include parental time allocation, labor supply, and earnings, stratified by child disability status.

Table I1 reports descriptive statistics from the pooled SIPP 2004 and 2008 samples. The sample includes 12,338 households, of which 989 have a disabled child. Parents of disabled children are slightly older on average, and their children are also older. Both husbands and wives in households with disabled children work fewer hours than their counterparts in households without disabled children, though earnings conditional on employment are comparable across the two groups.

Table I1: Descriptive Statistics: SIPP 2004 + 2008

Variable	Disabled	Non-Disabled	Difference	N
Husband's age	47.54	43.70	+3.84	12,338
Wife's age	44.94	41.35	+3.59	12,338
Child's age	11.61	9.33	+2.29	12,338
Husband's work hours (weekly)	30.99	31.65	-0.65	12,338
Wife's work hours (weekly)	21.13	23.12	-1.98	12,338
Husband's earnings (monthly)	4470.49	4431.37	+39.12	8,294
Wife's earnings (monthly)	3112.88	3071.80	+41.08	6,978
Husband's non-labor income (weekly)	78.27	61.41	+16.86	12,338
Wife's non-labor income (weekly)	56.86	32.07	+24.79	12,338
Father's time with child (weekly)	13.70	15.47	-1.77	859
N (households)	989	11,349		12,338

Table I2 presents parallel statistics from the PSID-CDS 2007 and 2014 waves. The PSID-CDS sample is smaller (2,094 households, with 519 having a disabled child) but provides higher-quality childcare time measures for both parents. A notable pattern emerges: mothers spend more time with disabled children

(+2.62 hours weekly), while fathers spend less time (-3.16 hours weekly), suggesting asymmetric responses to child disability within the household.

Data Requirements. Note that wages must be constructed, which requires conditioning on the exogenous variables. Therefore, households that lack information on each spouse's age, each spouse's education level, non-labor income, or the child's age are dropped from the sample.

All time variables are in weekly hours ($T = 112$, as in Section 3.1). Non-labor income is converted from annual to weekly ($Y^{NL}_{week} = Y^{NL}_{year}/52$), and Child SSI is set at \$625/month (≈\$156/week, using four weeks per month).[4]

Table I2: Descriptive Statistics: PSID-CDS 2007 + 2014

Variable	Disabled	Non-Disabled	Difference	N
Husband's age	45.14	46.73	-1.59	1,805
Wife's age	42.92	44.68	-1.76	1,805
Child's age	12.48	10.72	+1.76	2,074
Husband's education (years)	13.01	13.06	-0.05	1,750
Wife's education (years)	13.36	13.39	-0.03	1,774
Husband's work hours (weekly)	30.95	31.40	-0.44	1,805
Wife's work hours (weekly)	22.23	22.18	+0.05	1,805
Husband's earnings (monthly)	3965.27	4158.97	-193.70	1,389
Wife's earnings (monthly)	2723.18	2737.16	+13.98	1,218
Husband's non-labor income (weekly)	62.25	66.20	-3.95	1,805
Wife's non-labor income (weekly)	54.65	53.73	+0.93	1,805
Mother's time with child (weekly)	20.78	18.16	+2.62	1,224
Father's time with child (weekly)	15.63	18.79	-3.16	1,224
N (households)	519	1,575		2,094

J Missing Data in PSID-CDS Time Diary

I examined the pattern of missing time-diary data in the PSID-CDS following the approach of Keane and Wasi (2013). The analysis indicates that the time-diary observations are missing at random with respect to observable household characteristics, suggesting no systematic selection into non-response. However, selection on unobservable characteristics—such as parents with stronger childcare preferences being more likely to respond—cannot be tested directly. Reassuringly, when the same specification is estimated on

[4]This approximation of the Child SSI amount is used as the benchmark in the estimation. In principle, the amount could be made more flexible by incorporating institutional details such as the household's state of residence and income level. However, the estimation results are virtually unchanged under such variations.

different subsamples, the coefficient estimates remain similar, indicating that the data sources are close in terms of the underlying distributions.

K Disability Variable Correspondence

Table K1 documents the correspondence between disability variables across the SIPP and PSID-CDS datasets.

Table K1: Correspondence of Disability Variables between SIPP and PSID-CDS

Category	SIPP	PSID 2007	PSID 2014	Status
Panel A: Variables Used in Both Datasets				
Developmental Delay	EDDELAY	Q31A4M, Q31A9B	P14A10N	✓
Special Education	ESPECED	Q31B17, Q31B17A	P14B20	✓
Learning Disability	ELERNDIS	Q31A4H	P14A10I	✓
Mental Retardation	EKMR	Q31A4H, Q31A9C	P14A10I	✓
Developmental Disability	EKDEVDIS	Q31A4M	P14A10N	✓
ADHD/Hyperactivity	EADHD	Q31A4O	P14A10P	✓
Speech Disorder	EKSPECHD	Q31A4E	P14A10F	✓
Emotional Disturbance	EKSOCIAL	Q31A4I	P14A10J	✓
Orthopedic Impairment	EARMLEG	Q31A4L	P14A10M	✓
Panel B: Variables Used in SIPP Only				
Run/Play Limitation	ERUNPLAY	—	—	SIPP only
Sports Limitation	ESPORTS	—	—	SIPP only
School Work Limitation	ESKOOLWK	—	—	SIPP only
Panel C: Variables Used in PSID Only				
Autism	—	Q31A4N	P14A10O	PSID only
Panel D: Variables Excluded from Both Datasets				
Vision Problems	—	Q31A4G	P14A10H	Excluded
Hearing Problems	—	Q31A4F	P14A10G	Excluded
Asthma	—	Q31A4B	P14A10B	Excluded
Diabetes	—	Q31A4C	P14A10D	Excluded
Allergies	—	Q31A4P	P14A10Q	Excluded

Notes: Panel A shows disability categories used in both SIPP and PSID-CDS. Panel B shows SIPP-specific physical limitation variables. Panel C shows PSID-specific conditions (autism only, as emotional disturbance now has SIPP correspondence via EKSOCIAL). Panel D shows excluded conditions.