Deliverable Number: D3.2

Deliverable Title: Cooperative siblings model

Work Package: WP3

Deliverable type: DEC

Dissemination status: PU

Submitted by: Axel Börsch-Supan

Authors: Daniel Barczyk Yu Kyung Koh Matthias Kredler

Date Submitted: 29 June 2024

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under project ID 101093849.

www.mpisoc.mpg.de/max-planck-emeritus-gruppe/bb-future

In the mid2030s, the health of the baby boomers will have deteriorated and many in these large cohorts will be in need of formal and/or informal long-term care.

This “**care wave**” will transform two generations: the baby boomers in need of care and their children who may supply care. It will have significant implications for labour supply, especially for women, saving behaviour, and therefore for productivity, economic growth and its inclusiveness.

The overarching objective of BB-Future is to understand the size and the implications of the care wave on economic and social outcomes, to appreciate the quality of this second ageing-related transformation and to develop policy recommendations for advance planning on the EU and Memberstate levels.

This deliverable is a scientific paper on one of two micro models that describe how care choices are determined within a family. In this second model, the siblings cooperatively decide who will take care of the parent, depending on heterogeneous characteristics of, and economic incentives for, each sibling.

Please cite this deliverable as: Deliverable 3.2 of the BB-Future project funded under the European Union’s Horizon 2020 research and innovation programme GA No: 101093849.

Available at: www.mpisoc.mpg.de/max-planck-emeritus-gruppe/bb-future

Preference Heterogeneity versus Economic Incentives: What Determines the Choice to Give Care?

Daniel Barczyk
McGill University and CIREQ

Yu Kyung Koh
McGill University

Matthias Kredler
Universidad Carlos III de Madrid

June 28, 2024

Abstract

Family is a primary source of care, yet significant variations in care arrangements exist both across families and countries. We explore the factors contributing to these variations by estimating a discrete-choice model derived from a parsimonious structural model of the family. Parents and children bargain over care arrangements, choosing between child-provided and formal care. Children, heterogeneous in attributes such as labor income and geographical proximity, collectively decide on the potential caregiver. We find that although economic incentives matter, unobserved preference heterogeneity substantially reduces the elasticity of informal care in response to policy changes compared to a model in which only economic motivations for the care choice are included. This suggests that including preference heterogeneity is essential when it comes to policy analysis.

Contact information: Daniel Barczyk (daniel.barczyk@mcgill.ca): McGill University, Department of Economics, Leacock Building, Room 321b, 855 Sherbrooke Street West, Montreal, QC, H3A 2T7. Yu Kyung Koh (yu.koh@mcgill.ca): McGill University, Department of Economics. Matthias Kredler (matthias.kredler@uc3m.es): Universidad Carlos III de Madrid, Departamento de Economía, C. Madrid, 126, 28903 Getafe.

1 Introduction

Aging populations, together with changes in family structure (e.g., fewer children, increasing female labor-force participation, rising divorce rates), will pose significant challenges to governments. The aging of the baby boomers, combined with longer life expectancy and decreasing fertility rates, will dramatically increase the ratio of the elderly to the working-age population. For example, the share of the oldest-old (80 and over) in the population is projected to double over the next few decades, leading to a rise in long-term care expenditures. As a percentage of GDP, these expenditures are projected to increase by 168% in Germany and 149% in Spain between 2000 and 2050 (Comas-Herrera et al. 2003). Demographic changes will tighten government budgets, necessitating reforms in long-term care policy. Germany and Japan have already undertaken such reforms in 1996 and 2000, respectively.

Long-term care is defined as needing assistance with everyday activities (e.g., getting in and out of bed, showering, eating). Family members are a major source of this care. Barczyk & Kredler (2018) found that in the United States, most care is provided by family members, particularly spouses and adult children. Economic characteristics of the family, such as children's opportunity costs in the labor market and parents' wealth, significantly influence care arrangements. Informal home care is also crucial in many OECD countries (see Sundstroem et al., 2002; Zukewich, 2003; Moise et al., 2004), though there is substantial variation across countries. For instance, Barczyk & Kredler (2019) found that the average monthly hours provided by children to disabled elderly parents in Spain is 12.4, while in Sweden it is only 3.4.

A central question in long-term care policy design is how much families' informal care choices change when policies – and thus effective costs of formal care – change. In other words, it is important to understand the elasticity of informal care provision in response to formal care policies. Most existing models (e.g., Barczyk & Kredler, 2018) only consider economic forces such as opportunity costs in the labor market and care subsidies. However, caregiving preferences are likely to be heterogeneous across families. For some families, caregiving may be seen as a moral duty, while others might view it as burdensome or even shameful. This means that not all families will react the same way to policy changes related to formal care. The diversity in caregiving attitudes raises the question of how our understanding of the impact of policy changes on informal care might change if we account for these preference differences. Furthermore, questions that naturally arises include (i) how much caregiving preferences are explained by the observable factors (such as the caregiver's gender, geographical proximity, or psychological distance, including relationships with in-laws or stepchildren) and (ii) how much is left to unobserved heterogeneity (preference shocks in our setting) to explain caregiving decisions.

Our research aims to integrate this heterogeneity in caregiving preferences into our understanding of how families make caregiving decisions. We explore how our estimates of the elasticity of caregiving arrangements in response to policy changes are affected when we take into account preference heterogeneity seriously. Specifically, we investigate how data on which child provides informal care can inform us about child-specific economic and psychic costs of caregiving, and we examine how policy analysis changes when accounting for preference heterogeneity as well as changing demographics and family structure.

To this end, we develop a simple static model where parents and children bargain on care choices (formal or informal, and which child provides care). This model incorporates both idiosyncratic care preferences and those related to observable characteristics, across families with varying numbers of children. Theoretical results from our model allow us to estimate the informal care elasticity in response to policy and demographic changes using a logit-type estimator. We apply our model to SHARE (The Survey of Health, Ageing, and Retirement in Europe) data, exploiting both variation across families and cross-country policy variations. This provides a relatively simple methodology to predict care arrangements in various counterfactual scenarios. These scenarios include policy changes related to formal care and demographic shifts, such as the decline in the number of children and the narrowing of the gender wage gap.

We proceed with our analysis in the following steps. First, we start by examining formal care usage and informal care provision using several datasets. Our main data is SHARE, which provides rich information on the demographic and socioeconomic characteristics of the senior population in Europe, their adult children, and caregiving arrangements. To understand the economic costs in providing informal care for each SHARE child, we use the Eurostat public data to construct potential income based on the child's gender, education, country of residence, and survey year. Regarding formal care costs, we combine (a) the OECD statistics on the out-of-pocket nursing home costs as a percentage of old-age income in each country and (b) Eurostat statistics on old-age income for each country. It is worth emphasizing that our measures of earning potentials and formal care costs are preliminary in the current version. In our future analyses, we plan to construct more detailed measures of formal care costs by each country and by different care needs, as well as a more granular measure of income potential for each SHARE child.

Our descriptive analyses reveal substantial heterogeneity in care arrangements across countries and families with different characteristics. First, we show that families living in countries with lower formal care costs have higher propensities to use nursing homes than in other countries. Second, descriptive statistics show systematic differences in caregiving probabilities based on the child's characteristics. Caregiving children in our sample are more likely to be female, are living closer to parents, and have lower potential income relative to non-caregiving children. We additionally show that daughters are more likely than sons to provide care even conditional on distance from parents, potential income level, and education. We do not find meaningful differences in informal care probability based on the child's biological status, which is partially due to the small sample size of non-biological children in our current sample. Lastly, we show that there exist cross-country differences in children's distance from parents and in the prevalence of non-biological children. These cross-country differences may contribute to the differing informal care propensities across countries.

Next, we develop a theoretical model of how a family coordinates the selection of a caregiving mode for an elderly parent in need. This model incorporates heterogeneity in caregiving preferences based on children's observed characteristics and unobserved factors. This model extends the one introduced by Barczyk and Kredler (2018) by accounting for households with multiple children and modeling how these children coordinate in selecting the primary caregiver for their parents.[1] We derive a simple theoretical result: a child with the lowest combination of labor market opportunity cost and psychic cost is chosen as a primary potential caregiver. Furthermore, we demonstrate that non-caregiving siblings will compensate the caregiving child with extra consumption.

We then estimate how the characteristics of children affect their utility cost of providing informal care. In our theoretical framework, the psychic cost of caregiving can be shaped by an array of child characteristics such as gender, biological child status, and the distance from their parents. How much each characteristic affects the utility cost of caregiving is an empirical question, which we estimate using the theoretical lemma and SHARE data. We estimate these utility parameters using a discrete choice model framework and maximum likelihood estimation method. The estimates show that (a) there is a significant disutility attached to using formal care instead of informal care, (b) being a daughter lowers the utility cost of informal care, (c) no effect of being a non-biological child, which may be partly driven by the small sample size of non-biological children, and (d) larger distance away from parents significantly increases the utility cost of informal care.

Finally, using the estimated parameters, we conduct counterfactual analyses to examine how families' caregiving arrangements change in response to hypothetical changes in policies and demographics. In our first set of counterfactual analyses, we examine how much care arrangements would change if we made the provision of care by the government more generous. Our model predicts that if all countries have as generous formal care subsidies as the low-cost group countries, the formal care use would increase by 17% in the the middle-cost countries, while it would increase by 18% in high-cost countries. We also show that if we do not account for preference heterogeneity, then we would substantially overestimate the elasticity of formal care usage in response to government policies. This highlights the importance of accounting for preference heterogeneity to obtain realistic predictions.

In the second set of our counterfactual analyses, we examine how demographic and societal changes will affect care arrangements in the long run.

[1]It is worth noting that while Barczyk and Kredler (2018) model how a parent and a (representative) child bargain to determine the optimal provision of informal care, their initial model does not account for scenarios involving multiple children.

Based on the forecasts for 2050, we derive the following predictions. We predict that the declining number of children (driven by declining fertility/marriage rates) would drive the largest increase in formal care usage. Other demographic changes, such as an increase in female wages and an increase in the number of step-families, would play a relatively minor role in driving up the demand for formal care. Lastly, we find that the more generous formal care subsidies would lead to a modest increase in formal care usage, but not as much as the one predicted by the declining number of children. All in all, our current analysis suggests that demographic changes would have a more significant impact on the use of formal care than policy changes. However, it is important to note that our current results are preliminary, and we plan to refine our estimates with better construction of formal care costs and potential wages, as well as with a larger SHARE sample.

The organization of this paper is as follows. Section 2 describes data, sample selection, and construction of variables. Section 3 documents descriptive statistics regarding formal care and informal care using our estimation sample. Section 4 presents our theoretical model of family caregiving arrangement. Section 5 presents the discrete choice model framework to estimate the utility parameters. Section 6 describes the estimation procedures for the discrete choice model and presents the estimates. Section 7 performs counterfactual analyses. Section 8 describes our plan to incorporate the current model to the life-cycle framework. Section 9 concludes.

2 Data

2.1 SHARE

Our analyses rely on SHARE (The Survey of Health, Ageing, and Retirement in Europe), a panel of the European population aged 50 and above. This dataset provides extensive information on the demographic and socioeconomic characteristics of the senior population and their children, household caregiving arrangements, and health-related information. SHARE has been available bi-annually from 2004 to 2020.

Sample selection. We select the analysis sample based on several criteria. First, to focus our attention on the elderly with care needs, we limit the sample to households with at least one person aged 65+ who has at least one mobility limitation. Second, we only keep households with at least one child aged 20 to 60. This is because children outside this age range are either too young or too old to be the potential caregiver. We retain only the households where the demographic variables required for model estimation are complete for all children. Third, we limit to households that either use nursing homes (formal care) or have one intense informal caregiver among children. In the current analyses, we exclude households with multiple child caregivers to facilitate estimating the discrete choice model. Fourth, we drop households that cannot be matched with the constructed potential incomes of children or potential formal care costs. The construction of potential incomes and formal care costs is documented in later subsections.

For the current analyses, we only use the baseline surveys due to several issues with panel dimensions. The first issue is regarding the distance between the child and the parent. Although the distance is reported for baseline surveys, it is updated in later surveys *only if* the child moves. Distance is *not* updated when the parent moves, making it difficult to capture the correct distance information in non-baseline surveys. The second issue is regarding tracking the same child over time. Child's index does not remain the same across different waves, especially when the respondent for the child module changes over time. We plan to add non-baseline samples in the future after addressing these challenges. Appendix Table A1 compares the sample size between the full sample and baseline sample.

Furthermore, we do not use Waves 3, 4, and 7 in the current analyses for the following reasons. Waves 3 and 7 differ from other waves in that they are retrospective: they focus on respondents' life histories, not respondents' current life circumstances. Wave 4 is omitted because we cannot identify *which* child provided informal care. This is different from other waves where it is possible to identify the identities of the child caregivers through explicit questions in the Social Support (SP) module. In contrast, in Wave 4, the SP module only asks whether any child provided care, without specifying which one, thus preventing accurate identification of the caregiving child.[2]

After imposing the above sample selection criteria, we have a final sample of 1,829 households with 4,135 parent-child pairs. Appendix Table A2 reports how sample size changes after imposing each of the sample selection criteria.

Variable definition. We describe how we define care needs, formal care, and intense informal care. In the current analyses, the elderly with "care needs" are defined using a question in the Physical Health (PH) module that asks how many mobility limitations each respondent has (**ph048***). We characterize "need for care" if the respondent reports having at least one mobility limitation.

In the current analyses, formal care is defined as permanently staying in a nursing home (NH). We exclude temporary nursing home care. In future analyses, we plan to include formal home care (FHC) – which is care provided by paid helpers in the elderly's home. Barczyk & Kredler (2019) report that a larger portion of formal care in Europe is provided as NH than as FHC.

Intense informal care (IC) by children is defined using the frequency of informal care. SHARE differentiates between informal care from outside the household (OIC), e.g. from adult children living elsewhere, and informal care from inside the household (IIC), e.g. from the spouse or co-residing children. How OIC and IIC are reported and the associated care frequencies differ across waves, as summarized in Table 1.

[2]One way to infer the identity of the child caregiver in Wave 4 is to use the social network (SN) module. In Wave 4, the SP module asks whether parents received informal care from "social network" person, which is defined in the SN module. This "social network" person can be one of the respondent's children. Specifically, SN module documents (i) whether the social network person is a child, (ii) gender of the social network person, and (iii) distance between the respondent and the social network person. However, the caveat is that even the SN module in Wave 4 does not tell us *which* child is reported as a social network person. We can only infer his/her identity by matching the gender and distance information to children's information. Note that this may lead to imprecise matching if the household has multiple children of same gender and distance.

Table 1: Overview of Informal care (IC) variables in SHARE

	Informal care from outside hh. (OIC)	Informal care from inside hh. (IIC)
Wave 1	*Level:* Couple *Frequency:* 4 categories *Type:* Specified	*Level:* Individual *Frequency:* Defined as daily
Wave 2	*Level:* Couple *Frequency:* 4 categories *Type:* Specified	*Level:* Individual *Frequency:* Defined as daily
Wave 5	*Level:* Couple *Frequency:* 4 categories *Type:* NOT specified	*Level:* Individual *Frequency:* Defined as daily
Wave 6	*Level:* Individual *Frequency:* 4 categories *Type:* Specified	*Level:* Individual *Frequency:* Defined as daily
Wave 8	*Level:* Individual *Frequency:* 4 categories *Type:* Specified	*Level:* Individual *Frequency:* Defined as daily

Note: This table reports which information on informal care is available in SHARE for each wave and type of informal care. *Level:* whether the IC is reported at the couple level or at the individual level. *Frequency:* How the frequency of specified care is reported. 4 categories refer to (i) about daily, (ii) about every week, (iii) about every month, and (iv) less often. *Type:* refers to the types of OIC care provided, which has 3 categories (personal care, practical household help, and help with paperwork). Note that Waves 3, 4, 7 are not reported because Waves 3 and 7 are retrospective surveys and Wave 4 does not report the identity of child caregiver.

There are a few challenges in defining intense IC consistently across waves. First, in the earlier waves, OIC is reported at the couple level, not at the individual level; in other words, we only know if the respondent *and/or* the spouse received OIC, but not *who* received OIC. In the current analyses, the care need and care is defined at the *couple* level, so this does not pose a problem.[3] However, if we want to do future analyses at the individual parent level, then we would need to identify which of the parents received OIC. Second, the type of OIC (personal care, practical household help, and help with paperwork) is not reported in Wave 5. While this information is useful in determining intense IC, we decide not to distinguish among the types of OIC for consistency across waves.[4] Lastly, only about 21% of OIC by child occurs "about daily," as shown

[3]Specifically, our definition of child caregiver is the child who provided IC to any of the parents.

[4]Only about 10% of caregivers only provided help with paperwork, which can be considered

in Appendix Table A4. To increase the sample size, we define both "about daily" and "about every week" OIC as intense informal care. Additionally, we classify all IIC as intense informal care, since by definition in the SHARE survey, IIC occurs on an almost daily basis.

2.2 Potential income

SHARE does not provide income information on respondents' children. However, even if such data were available, it would not reflect the *potential* income of the children since observed income can be influenced by caregiving choices. For instance, a caregiving child might have a low observed income despite having a high potential income based on her education and abilities.

We construct the potential income for each child based on their demographic characteristics and the local labor market conditions. Specifically, we assign the potential annual income to each child based on the child's gender, education, and country of residence for each survey year. Income data is sourced from Eurostat's Structure of Earnings Survey for the years 2006, 2010, 2014, and 2018. Specifically, we use "mean hourly earnings by economic activity, sex, education attainment level" and "number of employees by economic activity, sex, educational attainment level." We exclude 2002 Eurostat data due to its lack of information for many countries in SHARE, primarily because many of the current EU countries joined the EU after 2004. To address differing prices across countries, we use the Purchasing Power Standard (PPS) instead of Euro. PPS is a common currency that adjusts national account aggregates for price level differences using Purchasing Power Parities (PPPs). We convert the hourly earnings to potential annual incomes by multiplying them by 40 hours per week and 52 weeks per year.

We construct two versions of potential wages. The first version does not consider the labor force participation rates of different social groups, while the second version incorporates these participation rates. The rationale behind incorporating participation rates in the second version is to address the overestimation of potential income, particularly for social groups with lower labor

as a light care. Hence, the majority of reported OIC can be considered as substantial care (personal care, household help).

force participation, such as women. If individuals in these groups are unlikely to participate in the labor force even when not providing informal care, it is crucial to account for this in their potential income estimates.

Specifically, we adjust the first version of potential wage to account for the labor force participation as follows. Let $PotentialWage_{gecy}$ denote the first version potential wage for gender g, education e, country c, and year c. The second version of potential wage is constructed as follows:

$$\begin{aligned} PotentialWage^{Adjusted}_{gecy} = & \; LFPR_{gcy} * PotentialWage_{gecy} \\ & + (1 - LFPR_{gcy}) * \frac{1}{2}(MinimumWage_{gcy}) \end{aligned}$$

where $LFPR_{gcy}$ is the labor force participation rate of gender g in country c in year y, and $MinimumWage_{gcy}$ is the minimum wage for gender g in country c in year y. The idea is to weight the potential wage by the labor force participation rate. We assume that for individuals participating in the labor force, the potential wage is the full amount derived in the first version. For those not in the labor force, we assume that their potential wage is the half of the country's minimum wage.

For the current analysis, $LFPR_{gcy}$ is based on the labor force participation rate of people aged 45-65 in each gender and country group for each year. The reason why we chose this age range is because approximately 75% of caregiving children in our SHARE sample are over age 45, as shown in Appendix Figure A2. Due to the limitations of the available data, we currently cannot further refine potential wages by age group or differentiate labor force participation rates by education level.[5] We plan to update our potential wage estimates once we gain access to the Eurostat microdata.

Appendix A.2 documents imputation strategies for potential wage construction. These strategies address several challenges, including (a) missing wage information for some years in Eurostat, (b) changes in educational classifications over time in Eurostat, and (c) differing survey years between SHARE and Eurostat.

[5]The public version of Eurostat data does not provide labor-related statistics categorized by age, education, gender, and country.

2.3 Formal care cost

We construct formal care costs that each SHARE household faces. Out-of-pocket FC costs vary widely depending on country, household income level, and the severity of care needs. Ideally, we aim to incorporate all these factors when assigning the FC costs to each SHARE household.

In the current version, our out-of-pocket costs are based on OECD statistics. Specifically, we use the OECD report on "Out-of-pocket costs of long-term care as a share of old age median disposable income after public support, for care recipients holding no net wealth, by severity of needs and care setting," as shown in Appendix Figure A1.

To construct formal care costs, we proceed with the following steps. First, we group European countries into three groups based on the expensiveness of formal care: (1) low cost (10∼40% of old-age income), (2) medium cost (50∼80% of old-age income), and (3) high cost (80∼120% of old-age income). The grouping of countries is as follows:

- Group 1 (Low FC cost): Sweden, Netherlands, Germany, Latvia, Denmark, Malta
- Group 2 (Medium FC cost): Italy, Ireland, Slovak Republic, Luxembourg, Finland, France, Slovenia, Austria, Belgium, Lithuania, Greece
- Group 3 (High FC cost): Croatia, Spain, Czech, Poland

Note that not all SHARE countries can be matched to the countries in the OECD report, so unmatched countries are dropped in current analyses.

Second, we take the midpoint for the FC cost share for each group[6] and multiply these values of FC cost share by old-age mean annual income for each country. The old-age mean income data is sourced from Eurostat's Structure of Earning Survey 2004-2018, which contains the mean annual earnings of people aged 65+ for each EU country.

In future analyses, we plan to construct a more detailed version of FC costs that better reflect variations across countries, household income levels, and the severity of care needs.

[6] This is 22.5% for Group 1, 65% for Group 2, and 100% for Group 3

3 Empirical Facts

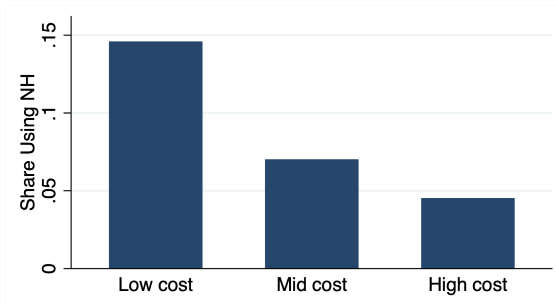
In this section, we document descriptive statistics using our estimation sample. Statistics regarding the full SHARE sample are documented in Appendix A.1.

Nursing home usage: First, out of the final 1,829 households, 150 households (8.2%) use nursing home (NH) care and the remaining households have one intense child caregiver. It is worth noting that the nursing home usage reported in SHARE is generally minimal; Appendix Table A3 shows that only 1% of all SHARE parents aged 65+ who have at least one mobility limitation use nursing home care.[7]

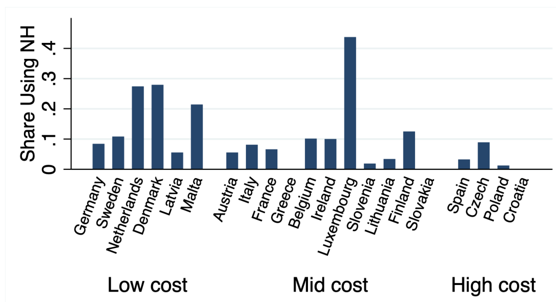
Nursing home usage differs across countries. Figure 1 shows that higher formal care costs are associated with a higher likelihood of nursing home usage in our estimation sample; Panel (a) shows that 14.6% of households from countries with low formal care costs utilize nursing home in our sample, compared to just 4.5% in countries with high formal care costs. Panel (b) also reveals that countries in the low FC cost group generally have a higher fraction of households using nursing homes than other groups. However, country-level statistics should be interpreted with caution, because many countries have small sample sizes in our estimation sample, as indicated in Appendix Table A6. For instance, although Panel (b) indicates that Luxembourg has a high probability of nursing home use, this may not be a reliable estimate because our sample includes only 32 households from Luxembourg. In future analyses, we aim to gain a better understanding of country-level nursing home usage by cleaning the complete SHARE data and collecting more accurate country-specific measures of formal care costs.

[7] Barczyk & Kredler (2019) discuss the under-sampling issue concerning the nursing home population in SHARE.

Figure 1: Nursing Home (NH) Probability by Country



(a) By country group based on FC cost



(b) By country

Note: This figure reports the proportion of households that are permanently using nursing home for sick parents for each country group (Panel (a)) and country (Panel (b)), using our estimation sample. Details on how the estimation sample is selected are provided in Section 2.1. In panel (a), “Low cost" group has FC cost as 10∼40% of old-age income in each country; “Mid cost" group has FC cost as 50∼80% of old-age income ; “High cost" group has FC cost as 80∼120% of old-age income. For more details on the grouping of countries, refer to Section 2.3.

Characteristics of caregiving children: We now turn to characteristics of caregiving children in our sample. Table 2 compares the characteristics of caregiving children versus non-caregiving children. Caregiving children are more likely to be female, are living closer to parents, and have lower potential income relative to non-caregiving children. There is no difference in age between caregiving children and non-caregiving children. Additionally, we do not find meaningful differences regarding biological child status, but this is partially because the sample size for non-biological children is very small in our sample (only 70 cases).

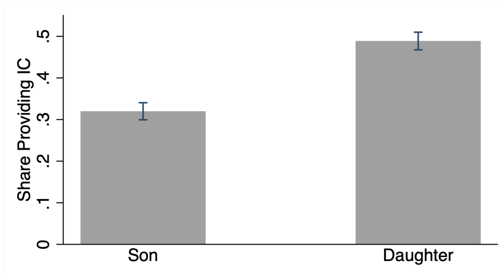
Table 2: Characteristics of caregiving children vs. non-caregiving children

	Caregiving Child	Non-Caregiving Child
Age	48.47	48.31
	(7.456)	(7.206)
Female	0.622	0.448
	(0.485)	(0.497)
Non-Biological	0.015	0.018
	(0.124)	(0.133)
Distance	9.39	86.08
	(35.96)	(159.56)
Potential Income (LFP not adjusted)	23,466.45	24,975.44
	(10,120.38)	(10,854.9)
Potential Income (LFP adjusted)	18,450.33	20,163.06
	(8,331.73)	(9,085.43)
Count	1,679	2,441

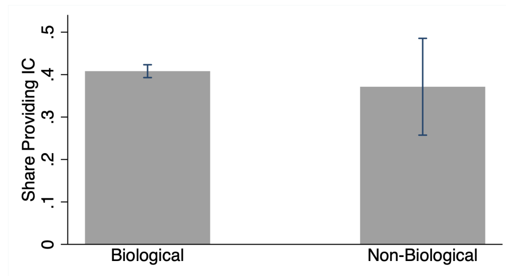
Note: This table reports the mean value of characteristics of caregiving children vs. non-caregiving children in our final SHARE sample. "Non-Biological" is indicator for step-child, adopted child, or foster child. "Distance" refers to the distance between child and parent. Distance for each child is assigned as a mid-value of the reported distance categories: (1) In the same household or building, (2) Less than 1 km, (3) 1-5 km, (4) 5-25 km, (6) 25-100 km, (7) 100-500 km, (8) 500+ km. Potential income for each child is constructed using the procedures in Section 2.2. For details on how the estimation sample was selected, refer to Section 2.1. For details on how the potential wages are constructed, refer to Section 2.2.

To better understand how children's characteristics relate to the probability of providing IC, we report the fraction of children providing IC conditional on the specified characteristic in Figure 2. Panel (a) shows a gender difference in providing IC: 48.9% of daughters in our sample provide IC, whereas it is 32% of sons. Panel (b) shows that there is no clear relationship between biological child status and IC probability, partly driven by the small sample size for non-biological children in our sample. Although biological children are slightly more likely to provide IC than non-biological children, large confidence intervals reveal that this difference is not statistically significant.

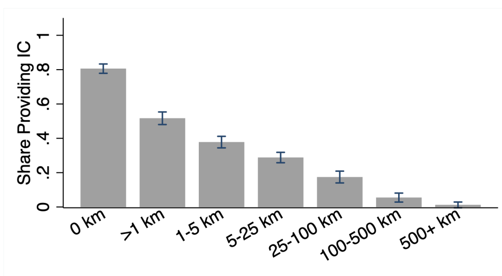
Panel (c) of Figure 2 demonstrates a strong negative relationship between the distance between children and their parents and the likelihood of children providing IC. Approximately 80.6% of children who live with their sick parent



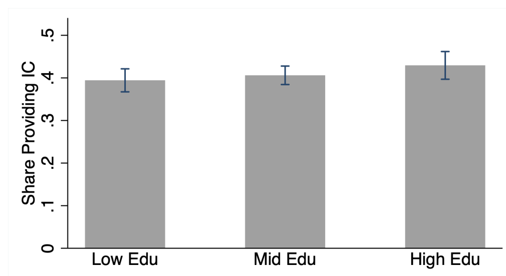
(a) By Gender



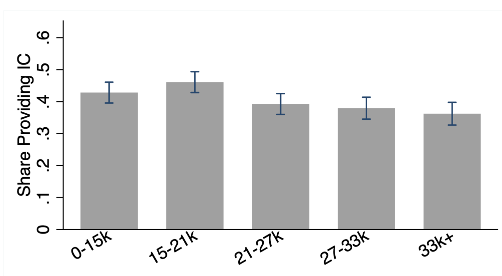
(b) By Biological Child Status



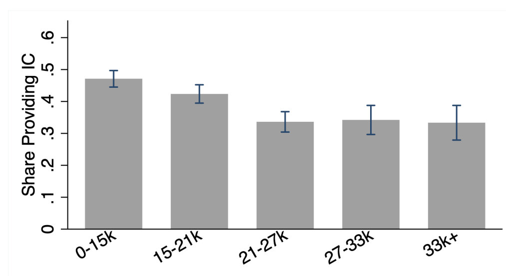
(c) By Distance



(d) By Education



(e) By Potential Income (No LFP adjustment)



(f) By Potential Income (LFP adjusted)

Figure 2: Informal Care Probability by Child's Characteristics

Note: This figure reports the proportion of children providing IC by each characteristics among our estimation sample. How the estimation sample is selected is documented in Section 2.1. In Panel (b), "Non-biological" children include stepchildren, adopted children, and foster children. In Panel (c), "Distance" is reported as km away from the parent. In Panel (d), "Potential income" is reported as Purchasing Power Standard (PPS), a common currency that adjusts national account aggregates for price level differences using Purchasing Power Parities (PPPs). In Panel (e), "Low Edu" is up to lower-secondary education (middle school), which is ISCED 2011 Levels 0-2. "Mid Edu" is up to high school graduation, which is ISCED 2011 Levels 2-4. "High Edu" is some college education or more, which is ISCED 2011 Levels 5-8. 95% confidence interval is reported.

provide IC, and this fraction decreases as the distance increases. It is important to note that these descriptive statistics do not establish a causal relationship between a child's distance and his or her probability of providing IC. For instance, a child might have moved closer to their parent's home, or the parent might have relocated to the child's home, facilitating the child's ability to provide care after the parent became ill. Conversely, it could be that the physical distance itself causally influences the likelihood of a child providing IC. In future analyses, we aim to better understand this relationship by examining the distances before and after the parent's illness.

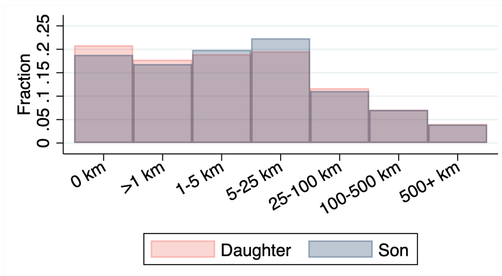
Panel (d) reveals no clear relationship between a child's education and IC probability. If anything, the higher education category is associated with a slightly larger IC probability.

Regarding the relationship between child's potential income and IC probability, we see different patterns depending on whether potential wage accounts for the labor force participation rates of different demographic groups. When potential wage does not account for labor force participation rates, we see no clear relationship between clear relationship between a child's potential income and IC probability, as shown in Panel (e) of Figure 2. In contrast, when potential wage accounts for differing labor force participation rates across groups, we see a negative relationship between a child's potential income and his or her IC probability, as shown in Panel (f).

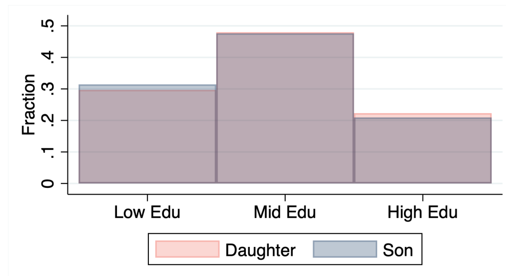
In future analyses, we plan to construct more detailed measures of the potential wage for each child once we get access to the Eurostat microdata.

Exploring gender differences in IC probability: In this section, we further explore the possible reasons behind the observed gender differences in the probabilities of providing IC. One possibility is that daughters tend to live closer to their parents compared to sons. Another consideration is that daughters may have lower levels of education or potential income than sons, which could affect their likelihood of providing IC.

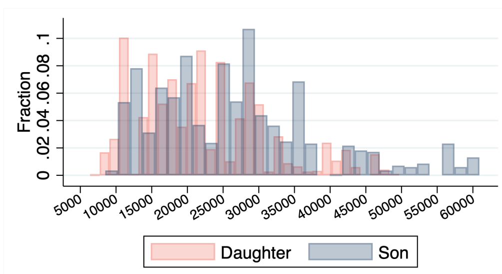
Figure 3: Gender Difference in Distribution of Child Characteristic



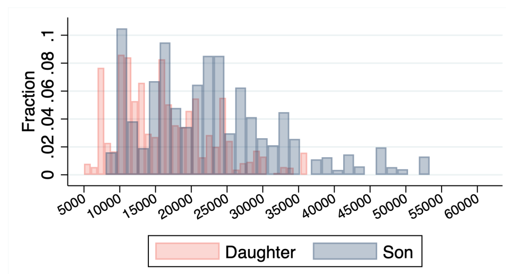
(a) Distance from parents



(b) Education



(c) Potential Income (No LFP adjustment)

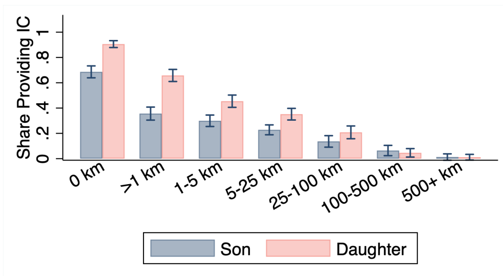


(d) Potential Income (LFP adjusted)

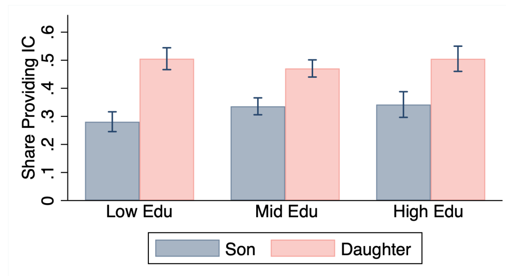
Note: This figure presents histograms for various characteristics of daughters and sons in our estimation sample. The pink bars represent the share of specified characteristics for daughters, while the blue bars represent those for sons. For details on how the estimation sample was selected, please refer to Section 2.1.

In Figure 3, we show the distribution of children's characteristics by gender. In Panel (a), we observe that the distances from parents are similarly distributed for both daughters and sons in our estimation sample, although a slightly higher proportion of daughters live closer to their parents. Panel (b) reveals that the education levels of daughters and sons are quite similar, showing no significant disparity. However, Panel (c) and (d) indicate that daughters generally have lower potential incomes compared to sons, which may reflect the gender wage gap in the labor market. Taken together, these statistics suggest that the gender difference in IC probabilities is not primarily driven by the gender differences in distance or education, as these characteristics do not show substantial gender gaps.

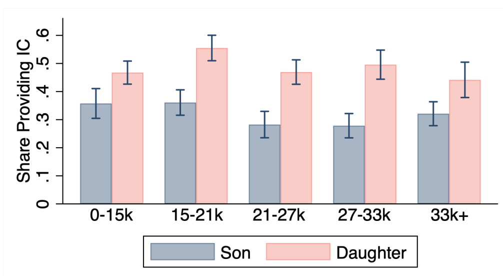
Figure 4: Gender Difference in IC Probabilities by Child's Characteristic



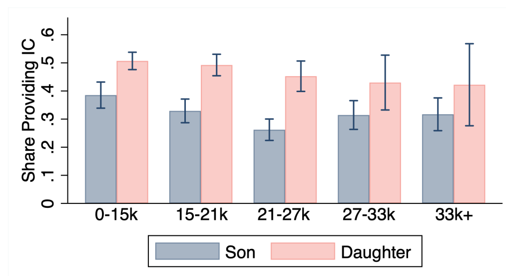
(a) Distance from parents



(b) Education



(c) Potential Income (No LFP adjustment)



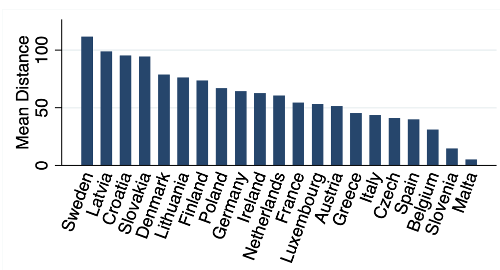
(d) Potential Income (LFP adjusted)

Note: This figure reports the proportion of children providing IC conditional on each specified characteristic, calculated separately for daughters and for sons in our estimation sample. How the estimation sample is selected is documented in Section 2.1. In Panel (b), "Low Edu" is up to lower-secondary education (middle school), which is ISCED 2011 Levels 0-2. "Mid Edu" is up to high school graduation, which is ISCED 2011 Levels 2-4. "High Edu" is some college education or more, which is ISCED 2011 Levels 5-8. In Panel (c), "Potential income" is reported as Purchasing Power Standard (PPS), a common currency that adjusts national account aggregates for price level differences using Purchasing Power Parities (PPPs). 95% confidence interval is reported.

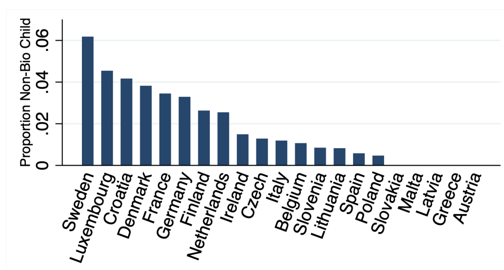
Figure 4 confirms that conditional on each and every characteristic, daughters have a higher probability of providing IC than sons do. For example, Panel (a) shows that although the larger distance from parents is associated with lower IC probabilities for both sons and daughters, daughters have higher IC probabilities than sons across all distance categories. We see similar patterns in terms of education and income categories. All in all, these statistics suggest that daughters may have systematically lower emotional or psychological costs associated with providing IC compared to sons. Alternatively, it could indicate the presence of a strong gender norm favoring daughters in IC provision that is not explained by other observable characteristics.

Cross-country differences in child characteristics: Finally, we explore cross-country differences in children's characteristics in our estimation sample. Panel (a) of Figure 5 presents the average distance between children and their parents in each country. Countries such as Sweden, Latvia, and Croatia exhibit higher average distances compared to others, which could contribute to varying levels of informal care provision by children across countries. Panel (b) shows the proportion of non-biological children in each country. It is important to note that this should be interpreted cautiously due to the small sample size of non-biological children in our current dataset (only 70 observations). Nevertheless, countries like Sweden and Luxembourg show a higher proportion of non-biological children compared to others. Given that the emotional and psychological costs of caregiving may differ based on biological child status, these cross-country differences in step-families could potentially influence variations in IC probabilities.

Figure 5: Cross-country Differences in Child's Characteristic



(a) Distance from parents



(b) Non-biological children

Note: This figure reports cross-country differences in children's characteristics in our estimation sample. For details on how the estimation sample was selected, please refer to Section 2.1. Panel (a) shows the mean distance between a child and her parents in each country. Panel (b) shows the fraction of non-biological children (i.e. stepchild, adopted child, foster child) in each country.

4 The Model of Family Caregiving Choice

We model how each family coordinates to choose a caregiving mode for an elderly parent, building on a unitary model, which is a workhorse model of intra-household decision-making (Chiappori & Mazzocco 2017). Our model consists of two stages. The first stage is the sibling cooperation stage, where adult children collectively decide who among them will become the (potential) primary caregiver for their parents. The second stage is the bargaining stage, where the selected caregiving child bargains with the parent to determine whether to actually provide informal care and negotiate the exchange-motivated transfer. Below, we outline the specifics of our model.

4.1 Sibling cooperation stage

Preferences. For simplicity, let's consider a family with two adult children. Each child i derives utility from their own consumption (c_i) and from the utility value of caregiving, which is expressed in terms of consumption, in case he or she provides informal care. We specify child i's utility function as the following:

$$u_i(c_i, ic_i) = \frac{1}{1-\gamma}(c_i - \theta_i ic_i)^{1-\gamma}, \quad ic_i = \begin{cases} 1 & \text{if informal care} \\ 0 & \text{otherwise} \end{cases}$$

where θ_i is the cost or benefit of providing informal care for child i. Note that if $\theta_i > 0$, caregiving reduces effective consumption that yields utility, acting as a 'cost' to the child. Conversely, if $\theta_i < 0$, caregiving increases effective consumption, acting as a 'benefit' to the child. The utility value of IC depends on an array of child characteristics, including the child's gender, biological child status, and distance from parents. Hence, we specify θ_{ij} as:

$$\theta_{ij} = \beta_0 + \beta_1 female_{ij} + \beta_2 biological_{ij} + \beta_3 distance_{ij} + \varepsilon_{ij}$$

where ε_{ij} captures unobservable characteristics affecting the utility value of caregiving.

Despite being parsimonious, our utility specification can include a rich set of children's characteristics, which allows us to quantitatively examine how one's and siblings' characteristics play a role in determining the optimal caregiver. Additionally, from a computational standpoint, our specification allows us to maintain a manageable state space by reducing the dimensionality of child attributes into a single utility value of caregiving for each child.

Child generation's problem. Children choose consumption allocation and potential caregiver by solving the maximization problem below, which follows a unitary model structure:

$$\begin{aligned} \max_{c_1,c_2,ic_1,ic_2} \quad & \mu_1 \frac{1}{1-\gamma}(c_1-\theta_1 ic_1)^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}(c_2-\theta_2 ic_2)^{1-\gamma} \\ s.t. \quad & c_1 + c_2 = Q + (1-ic_1)y_1 + (1-ic_2)y_2 \end{aligned}$$

where μ_1 and μ_2 are the weights for each child that together make up the unit. Q is the exchange-motivated transfer from the parent that will be determined in the second stage. y_i denotes the potential income of child i. Note that if child 1 becomes a caregiver ($ic_1 = 1$), the budget constraint reflects that the child generation forgoes the potential labor income of child 1 (y_1). Similarly, if child 2 takes on the caregiving role, then they forgo child 2's potential earning.

We assume that either of the two children must be the potential caregiver. Then, we can set $ic_2 = 1 - ic_1$ and rewrite the child generation's problem as:

$$\begin{aligned} \max_{c_1,c_2,ic_1} \quad & \mu_1 \frac{1}{1-\gamma}(c_1-\theta_1 ic_1)^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}(c_2-\theta_2 (1-ic_1))^{1-\gamma} \\ s.t. \quad & c_1 + c_2 = (1-ic_1)y_1 + ic_1 y_2 + Q \end{aligned}$$

Solution. We solve the maximization problem in two steps. First, conditioning on each caregiver choice, we obtain the indirect utility function for each choice. Second, the caregiver choice that leads the highest indirect utility for the child generation is selected.

In the first step, we get the following *conditional* indirect utility functions:

1. Conditional on child 1 being the caregiver ($ic_1 = 1$):

$$W_{ic_1=1} = \frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}}\right)^{\gamma}(y_2 + Q - \theta_1)^{1-\gamma}$$

2. Conditional on child 2 being the caregiver ($ic_1 = 0$):

$$W_{ic_1=0} = \frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}}\right)^{\gamma}(y_1 + Q - \theta_2)^{1-\gamma}$$

We provide the full derivation for these indirect utility functions in the Appendix Section A.5.1.

The second step is to compare the conditional indirect utilities above. Specifically, child 1 is selected as the caregiver if and only if

$$\begin{aligned} W_{ic=1} &> W_{ic=0} \\ \Rightarrow \quad \frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}}\right)^{\gamma}(y_2 + Q - \theta_1)^{1-\gamma} &> \frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}}\right)^{\gamma}(y_1 + Q - \theta_2)^{1-\gamma} \\ \Rightarrow \quad y_2 + \theta_2 &> y_1 + \theta_1 \end{aligned}$$

In simple terms, if the combined opportunity cost from the labor market and the utility cost of caregiving is higher for child 2 than for child 1, then child 1 will be the potential caregiver. Conversely, if the combined cost is higher for child 1, child 2 will be selected. This principle extends to any number of n adult children in the family. Thus, we derive the following lemma from solving the sibling cooperation stage.

Lemma 1. *Let C_i denotes the set of children in the child generation i. Child j becomes the caregiver in the child generation i if and only if*

$$y_k + \theta_k \;>\; y_j + \theta_j \quad \forall\, k \in C_i \text{ and } k \neq j$$

Regarding the optimal consumption allocation, when child 1 is chosen as a caregiver ($ic_1 = 1$), the resulting consumption allocation is:

$$c^*_{1|ic_1=1} = \left(1 + \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1} (y_2 + Q + \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}} \theta_1)$$
$$c^*_{2|ic_1=1} = \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}} \left(1 + \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1} (y_2 + Q - \theta_1)$$

See Appendix Section A.5.1 for proof. Noteworthy aspects are the following:

- First, when child 1 and child 2 have the same Pareto weights ($\mu_1 = \mu_2$), the above allocation boils down to $c^*_{1|ic_1=1} = \frac{1}{2}(y_2+Q+\theta_1)$ and $c^*_{2|ic_1=1} = \frac{1}{2}(y_2+Q-\theta_1)$. This shows that the caregiving child (child 1) is compensated with additional consumption for providing informal care.

- Second, individual consumption depends on the Pareto weights. Generally, the child with higher Pareto weight has higher consumption than the sibling. See Appendix Section A.5.1 for proof.

4.2 Bargaining with parents

In this stage, the chosen caregiving child from the first stage bargains with the parent to decide (1) whether the chosen child will actually provide the informal care and (2) the amount of exchange-motivated transfer Q from the parent.

Preferences. Parent derives utility from own consumption (c^p), informal care by the selected child (ic^k), and utility value of not receiving informal care from child (C_f). The utility of the parent is given by:

$$u^p(c^p) = \frac{1}{1-\gamma}(c^p - (1 - ic^k)C_f)^{1-\gamma}$$

Child i's utility is the same as in Section 4.1, except that there is k superscript attached to distinguish child's utility from parent's.

$$u^k_i(c^k_i, ic^k_i) = \frac{1}{1-\gamma}(c^k_i - \theta_i ic^k_i)^{1-\gamma}$$

Bargaining problem. For convenience, let's assume that child 1 is selected as a potential caregiver from stage 1 (Section 4.1). In the second stage, the key idea is child 1 provides informal care if there exists a non-negative transfer Q such that both parent and child generations' surpluses are positive. Surplus $S^i(Q)$, where i is either parent or child generation, is defined as the difference between i's utility when IC is provided versus i's utility when IC is not provided (FC provided).

The equilibrium Q^* is determined by the following Nash-bargaining problem:

$$Q^* = \begin{cases} \underset{Q \leq 0}{argmax} \left(S^k(Q)\right)^{\omega} \left(S^p(Q)\right)^{1-\omega} & \text{if } ic^k = 1 \\ 0 & \text{otherwise} \end{cases}$$

where ω is the bargaining weight for the child generation and $1 - \omega$ is the bargaining weight for the parent.

Parent's surplus. To solve the bargaining stage, we need to first derive the surplus functions for the parent and child generations, respectively. The parent's surplus function is:

$$S^p(Q) = \underbrace{u^p(c^{p*}_{ic^k=1})}_{\text{IC takes place}} - \underbrace{u^p(c^{p*}_{ic^k=0})}_{\text{FC takes place}}$$

We get the solution for the parent's surplus function by solving the parent's utility maximization problem (i) when IC takes place and (ii) when FC takes place, respectively. Specifically, the parent's utility maximization problem when IC takes place is:

$$\underset{c^p}{max} \frac{1}{1-\gamma} c^{p\,1-\gamma} \quad s.t. \quad c^p = y^p - Q$$

where y^p is the parent's income or asset. Hence, when IC takes place, the parent has to pay the exchange-motivated transfer Q to the child, but there is no utility cost incurred.

The parent's problem when IC does not take place, that is, when FC takes place is:

$$\underset{c^p}{max}\,\frac{1}{1-\gamma}(c^p - C_f)^{1-\gamma} \quad s.t. \quad c^p = y^p - p_{bc}$$

where p_{bc} is the cost of formal care. This shows that when IC does not happen, the utility cost C_f is incurred to the parent. Given our assumption that the elderly parent opts for formal care when informal caregiving does not occur, they must pay the formal care cost p_{bc} in this case.

When we solve the above, we get the following surplus function for the parent:

$$S^p(Q) = \frac{1}{1-\gamma}(y^p - Q)^{1-\gamma} - \frac{1}{1-\gamma}(y^p - (p_{bc} + C_f))^{1-\gamma} \tag{1}$$

The full derivation is provided in the Appendix Section A.5.2.

Child generation's surplus. For the child generation, the surplus function is defined for the whole generation, not for the individual child. The surplus function is specified as follows:

$$S^k(Q) = \underbrace{U^k(c^{k*}_{1|ic^k=1}, c^{k*}_{2|ic^k=1})}_{\text{IC takes place}} - \underbrace{U^k(c^{k*}_{1|ic^k=0}, c^{k*}_{2|ic^k=0})}_{\text{formal care takes place}}$$

The solution for the child generation's surplus function is obtained by solving the child generation's problem (i) when IC takes place and (ii) when FC takes place, respectively. Specifically, the problem when IC takes place is:

$$\underset{c^k_1, c^k_2}{max} \quad \mu_1 \frac{1}{1-\gamma}(c^k_1 - \theta_1)^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}(c^k_2)^{1-\gamma} \quad s.t. \quad c^k_1 + c^k_2 = y^k_2 + Q$$

In this scenario, since child 1 is chosen as the caregiver, she experiences the utility cost or benefit of caregiving (θ_1). Additionally, the child generation forgoes child 1's labor income (y^k_1) but receives the transfer Q from the parent.

The child generation's problem when FC takes place is:

$$\underset{c^k_1, c^k_2}{max} \quad \mu_1 \frac{1}{1-\gamma}(c^k_1)^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}(c^k_2)^{1-\gamma} \quad s.t. \quad c^k_1 + c^k_2 = y^k_1 + y^k_2$$

In this case, there is no utility value of caregiving incurred to any child, and children retain their full potential labor income.

When we solve the above, we obtain the following surplus function for the child generation:

$$S^k(Q) = \frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}}\right)^{\gamma}(y_2^k + Q - \theta_1)^{1-\gamma} - \frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}}\right)^{\gamma}(y_1^k + y_2^k)^{1-\gamma} \tag{2}$$

The full derivation is provided in the Appendix Section A.5.2.

Solution for optimal Q^*. Although there is no closed-form solution for the optimal transfer Q^* from the Nash bargaining problem, the surplus functions give us the bounds for Q that make IC happen.

First, the parent's surplus function (Eq. 1) reveals that $S^p(Q)$ is positive as long as $p_{bc} + C_f > Q$. In words, the parent gets a positive surplus from IC if the combination of the cost of not receiving IC (C_f) and the cost of FC (p_{bc}) is greater than the transfer that they give to the child generation. Therefore, $p_{bc}+C_f$ is the **maximum transfer** $\bar{Q}$ that makes the parent indifferent between IC and FC.

Second, the child generation's surplus function (Eq. 2) reveals that $S^k(Q)$ is positive as long as $Q > y_1^k + \theta_1$. This means that the child generation received positive surplus from IC if the combination of the caregiving child's labor market opportunity cost and psychic cost is less than the transfer they receive from parents for providing IC. Therefore, $y_1^k + \theta_1$ is the **minimum transfer** $\underline{Q}$ that makes the child generation indifferent between IC and FC.

In summary, the optimal choice of the caregiver only depends on which choice maximizes total child generation resources. Informal care takes place if and only if $Q \in (p_{bc} + CF, y_1^k + \theta_1)$, which leads to positive surpluses for both the parent and the child generation.

Main result. We now state our main result on the caregiving decision. Define the *monetary cost* of care choices as

$$MC_j = \begin{cases} p_{bc} & \text{if } j = 0, \\ y_j & \text{if } j > 0. \end{cases}$$

Theorem 1 (Efficient care choice) *The bargaining solution implements the care choice with the minimal effective (=monetary+psychic) cost, i.e.*

$$j^*_{eff} = \arg \min_{j \in \{0,1,\ldots,K\}} \{MC_j + \theta_j\},$$

which is efficient.

Remark: Any other bargaining protocol that satisfies **efficiency** yields **same care outcome** (e.g.: parent makes take-it-or-leave-it offer Q, children bid auction-style, collective model of entire family...)

5 Empirical Discrete Choice Model

In this section, we outline the discrete choice model of caregiving mode to estimate the parameters of the utility cost function θ_{ij} shown in the previous section. The main goal is to quantify how economic costs and various child characteristics – such as gender, biological relationship, and proximity to the parent – affect the provision of informal care for elderly parents.

A family i chooses its caregiving mode j among the following options: formal care (FC) or informal care (IC) provided by one of its children. Hence, the **choice set** $\mathcal{C}_i$ for each family i is the following:

$$\mathcal{C}_i = \{0, 1, 2, \ldots, J_i\}$$

where J_i is the total number of children in family i. $j = 0$ refers to FC, and $j > 0$ refers to IC by child j.

Each caregiving option j has its associated cost, C_{ij}, which includes both monetary and psychic components:

$$C_{ij} = p_{ij} + \theta^*_{ij}$$

Monetary costs are given by

$$p_{ij} = \begin{cases} p_i^{bc} & \text{if } j = 0 \\ y_{ij} & \text{if } j > 0 \end{cases}$$

where p_i^{bc} is the price of basic formal care faced by parent i and y_{ij} is the potential income of child j from full-time employment. Psychic costs are given by

$$\theta^*_{ij} = \begin{cases} \theta^*_{FC} + \varepsilon^*_{i0} & \text{if } j = 0 \\ \theta^*_{IC} + \theta_{ij} + \varepsilon^*_{ij} & \text{if } j > 0 \end{cases}$$

where we decompose psychic costs into a systematic part, common across all individuals, and individual-specific parts. The individual-specific part of the FC choice is an unobservable preference shock ε^*_{i0}. For the IC choice, there are also child-specific characteristics contained in θ_{ij}, such as, gender, step-child status, and distance from parents, in addition to an unobservable child-specific preference shock ε^*_{ij}.

To frame the choice problem in the language of standard discrete-choice models, we define the utility benefit stemming from the deterministic part of care arrangement j as the negative of the costs

$$V_{ij} = -p_{ij} - \begin{cases} \theta^*_{FC} & \text{if } j = 0 \\ (\theta^*_{IC} + \theta_{ij}) & \text{if } j > 0 \end{cases}$$

Family i implements care arrangement $j = j^*$ if and only if:

$$U_{ij^*} \equiv V_{ij^*} + \epsilon^*_{ij^*} \geq V_{ij} + \epsilon^*_{ij} \equiv U_{ij}, \quad \forall j \in \mathcal{C}_i,\ j \neq j^*$$

Following a large literature on estimating discrete-choice models, we assume that the unobservable choice-specific taste shock ϵ^*_{ij} is i.i.d. and follows an extreme value distribution with scale parameter σ and location parameter zero (EV-1/Gumbel with scale sigma). The probability of observing option j in family i is then given by

$$P_{ij} = \frac{\exp\left(\frac{V_{ij}}{\sigma}\right)}{\sum_{j=0}^{J_i} \exp\left(\frac{V_{ij}}{\sigma}\right)}$$

Here we can see that if the scale parameter is $\sigma > 1$, it acts to attenuate the impact of the "true" deterministic part of the monetary and the psychic costs. In the limit, as σ becomes large, the choice probability converges to the unconditional probability $1/(1 + J_i)$ as observable attributes become uninformative due to the vast heterogeneity in unobserved preferences.

We can also write $\epsilon^*_{ij} = \sigma\eta^*_{ij}$, where η^*_{ij} follows a standard Gumbel distribution with location parameter $\mu = 0$ and scale parameter $\sigma = 1$. The variance of η^*_{ij} is $\pi^2/6$ and that of ϵ^*_{ij} is $\sigma^2(\pi^2/6)$.

6 Estimation

To estimate the discrete choice model, we specify the utility benefit of informal care provided by child j for family i as the following.

$$U_{ij} = V_{ij} + \epsilon^*_{ij}$$

$$\text{where} \quad V_{ij} = \begin{cases} -p_i^{bc} - \theta^*_{FC} & \text{if } j = 0 \\ -y_{ij} - (\theta^*_{IC} + \beta^*_1 \text{female}_{ij} + \beta^*_2 \text{non-bio}_{ij} + \beta^*_3 \text{dist}_{ij}) & \text{if } j > 0 \end{cases}$$

and ϵ^*_{ij} follows a Gumbel distribution with location parameter 0 and scale parameter σ. We note that the utility is denominated in the same unit as the consumption good which follows from our specification of the utility functions where psychic costs act to reduce utility-generating consumption. Also, the labelling of children, $1, \ldots, K$, in our setting has no meaning so that the constant θ^*_{IC} is the same for all children.

To identify the parameters of interest, we need to impose two normalizations.

First, we need to normalize one of θ^*_{FC} and θ^*_{IC} . We cannot identify both θ^*_{FC} and θ^*_{IC} because only differences in utility matter in the caregiving choice. There are infinite combinations of θ^*_{FC} and θ^*_{IC} that can rationalize the same choice. We choose to normalize θ^*_{IC} to 0, and consequently need to interpret θ^*_{FC} relative to the reference point θ^*_{IC}.

Second, we need to normalize the scale of utility benefit, as the overall scale of utility is irrelevant in discrete-choice models. Following the standard approach, we normalize the scale by normalizing the variance of ε^*_{ij} to $\pi^2/6$, corresponding to a scale parameter $\sigma = 1$. The specification of utility is then given by

$$\frac{U_{ij}}{\sigma} = \frac{V_{ij}}{\sigma} + \eta_{ij}$$

where η_{ij} follows a Gumbel distribution with location parameter 0 and scale parameter 1, and

$$\tilde{V}_{ij} = \begin{cases} -\frac{1}{\sigma}p_i^{bc} - \frac{\theta^*_{FC}}{\sigma} & \text{if } j = 0 \\ -\frac{1}{\sigma}y_{ij} - (\frac{\beta^*_1}{\sigma}\text{female}_{ij} + \frac{\beta^*_2}{\sigma}\text{non-bio}_{ij} + \frac{\beta^*_3}{\sigma}\text{dist}_{ij}) & \text{if } j > 0 \end{cases} \tag{3}$$

$$= (-1) * \begin{cases} \theta^{FC} + \beta^0 p_i^{bc} & \text{if } j = 0 \\ \beta^0 y_{ij} + \beta^1\text{female}_{ij} + \beta^2\text{non-bio}_{ij} + \beta^3\text{dist}_{ij} & \text{if } j > 0 \end{cases} \tag{4}$$

Here, we can see that the estimated coefficients capture the true effect of a variable relative to the size of the unobserved factors.

To estimate the unknown coefficients, we maximize the likelihood (or log-likelihood) of observing the actual choices made in the data, i.e. coefficient estimates that best explain the observed choices given the assumptions of the model. Our log-likelihood function is the following:

$$
\begin{aligned}
LL(\boldsymbol{\beta}) &= \sum_{i=1}^{N} \sum_{j \in C_i} 1\{d_{ij} = 1\} \ln(P_{ij}) \\
&= \sum_{i=1}^{N} \sum_{j \in C_i} 1\{d_{ij} = 1\} \ln\bigg(\frac{e^{\tilde{V}_{ij}}}{\sum_{j \in C_i} e^{\tilde{V}_{ij}}} \bigg) \\
&= \sum_{i=1}^{N} \sum_{j \in C_i} 1\{d_{ij} = 1\} \bigg(\tilde{V}_{ij} - \ln\Big(\sum_{j \in C_i} e^{\tilde{V}_{ij}} \Big) \bigg)
\end{aligned}
$$

Using a standard maximum likelihood estimation (MLE) approach, we estimate the parameters in Equation 4: $\{\theta^{FC}, \beta^0, \beta^0, \beta^2, \beta^3\}$. In practice, we multiply all constants and explanatory variables by -1 so that the interpretation of each coefficient becomes the effect of each characteristic on the utility "cost" of caregiving.

Table 3: Estimated parameters

	LFP unadjusted		LFP adjusted		
Coefficient	Estimate	S.E.	Estimate	S.E.	Underlying Parameter
θ^{FC}	1.990	(0.112)	1.962	(0.108)	$\frac{\theta^{*FC}}{\sigma\sqrt{\pi^2/6}}$
β^0 (for monetary cost)	0.110	(0.049)	0.154	(0.057)	$\frac{1}{\sigma\sqrt{\pi^2/6}}$
β^1 (for female)	-0.771	(0.080)	-0.726	(0.085)	$\frac{\beta^{*1}}{\sigma\sqrt{\pi^2/6}}$
β^2 (for non-biological)	0.181	(0.430)	0.161	(0.431)	$\frac{\beta^{*2}}{\sigma\sqrt{\pi^2/6}}$
β^3 (for distance)	1.677	(0.146)	1.668	(0.146)	$\frac{\beta^{*3}}{\sigma\sqrt{\pi^2/6}}$

Notes: This table reports estimated parameters from Equation 4. The last column shows the relations to underlying true parameter (Equation 3) for each estimated parameter. "LFP unadjusted" shows the parameter estimates when using the potential wage that does not adjust for differing labor force participation across groups. "LFP adjusted" shows the parameter estimates when using the potential wage that adjusts for labor force participation, as described in Section 2.2. Estimation is done using the SHARE data.

Table 3 shows the estimation results. We show the results for (a) the case when we use potential wage that does *not* account for labor force participation rates ("LFP unadjusted") and (b) the case when we use potential wage that accounts for labor force participation rates ("LFP adjusted"), as described in

Section 2.2. The interpretations of each coefficient in terms of "true" parameters are as follows:

- θ^{FC}: This is the effect of choosing FC instead of IC on utility cost "*relative to*" the standard error of unobserved factors. The estimated value shows statistically significant positive utility cost of choosing FC instead of IC.

- β^0: With our quasilinear form of utility cost of caregiving, this parameter is interpreted as the inverse of the standard error of unobserved factors (multiplied by $\sqrt{\pi^2/6}$). Note that while using LFP-adjusted potential wages increases β^0 compared to the LPF-unadjusted versions, the estimates remain similar across the two versions.

- β^1: This is the effect of being a female on the utility cost of caregiving "*relative to*" the standard error of unobserved factors. The estimate shows that being a female lowers the utility cost of caregiving, and this effect is statistically significant.

- β^2: This is the effect of being a step-child on the utility cost of caregiving "*relative to*" the standard error of unobserved factors. The estimate shows a very noisy effect of being a step-child, as shown by a large standard error relative to the estimated value.

- β^3: This is effect of increasing a distance by 100 km on the utility cost of caregiving "*relative to*" the standard error of unobserved factors. The estimate shows that a larger distance is associated with higher utility cost of caregiving, and this effect is statistically significant.

We acknowledge that our potential wage measures might contain measurement errors, as they are constructed only based on one's gender, education, country, and survey year. In reality, there can be many other unobserved factors that shape one's earning potentials. In our future analyses, we plan to address this issue and examine how our estimates change when we account for the potential measurement errors. Specifically, we plan to perform MLE estimation with wage residuals, which is described in Appendix Section A.6.

7 Counterfactuals

In this section, we describe how we carry out counterfactuals and present their results.

Methodology We follow closely the methodology from the literature on discrete-choice models, making adjustments only where needed. In general, we treat the estimated coefficients β as deep, policy-invariant preference parameters and leave them unchanged in the counterfactuals. However, we make changes to the distribution of right-hand-side variables (specifically of the variables p_{bc}, y_{ij}, $dist_{ij}$, $step_{ij}$ and to the numer of children, K, thereby]creating matrices $\tilde{\mathbf{X}} \neq \mathbf{X}$).[8] The counterfactuals thus answer the question: How would care choices change if economic and demographic conditions ($\mathbf{X}$) changed, but the preferences underlying choices (β) stay unchanged? Specifically, denote by $X_{ij} = (pb_i, y_{ij}, step_{ij}, dist_{ij})$ observation in family i, child j and by $\tilde{X}_{ij}$ the same vector in the counterfactual. We then compute the probability of care outcome $j \in \{0, 1, \ldots, K_i\}$ in family i in the counterfactual as

$$\tilde{P}_{ij} = \frac{\exp(\tilde{X}'_{ij}\beta)}{\sum_{j=0}^{K} \exp(\tilde{X}'_{ij}\beta)}. \tag{5}$$

We then aggregate these probabilities over all families. Concretely, to compute the prevalence of option $j = 0$ (formal care) in the popluation, for example, we compute the sum

$$FC = \sum_{i=1}^{N_{fam}} \tilde{P}_{i0}, \tag{6}$$

where N_{fam} is the number of families in our data.[9] For the counterfactuals that involve only changes to p_{bc}, y_{ij} and $dist_{ij}$, this completely describes our

[8]Essentially, in this step we use the observed joint distribution over $\mathbf{X}$ in the baseline as a non-parametric estimate of the joint cdf of regressors. An alternative approach would be to estimate the joint distribution of the regressors parametrically and then to make changes to this distribution, however, this would require to make parametric assumptions on the distribution, which we avoid here.

[9]Note here that this prediction is preferable to drawing preference shocks for all children, since the logit formula (11) removes sampling noise. Indeed, $\tilde{P}_{i0}$ is the frequency with which family i would choose formal care if we drew an infinite amount of shocks.

algorithm.

For the remaining counterfactuals (in which K or *step* or both of them change), however, we do resort to simulation. For these counterfactuals, we run $N_{sim} = 1,000$ simulations in which we randomly change K and/or $dist$, but maintaining all *other* variables unchanged. Specifically, we proceed as follows:

K: We delete each child in the sample with probability p_K, where p_K is set such that we obtain the projected number of children in the future (see below).

step: We change each biological child to a step child with probability p_{step} in order to match the prevalence of non-biological children in the future.

For each simulation $n \in \{1, \ldots, N_{sim}\}$, we then calculate the population change in the FC probability as

$$\Delta FC(\Delta, n) = \frac{1}{N_{fam}} \Big[\sum_{i=1}^{N_{fam}} \hat{P}_{i0}(\Delta, n) - P_{i0} \Big], \tag{7}$$

where P_{i0} is the predicted FC probability of family i in the baseline estimation and where $\hat{P}_{i0}(\Delta, n)$ is the FC prob. of i in counterfactual Δ in simulation n, which we calculate using (11). Finally, to obtain the predicted change in FC from counterfactual Δ, we average over the N_{sim} simulations to obtain

$$\Delta FC(\Delta) = \frac{1}{N_{sim}} \sum_{n=1}^{N_{sim}} \Delta FC(\Delta, n). \tag{8}$$

7.1 Transplanting LTC systems

We first run a set of policy counterfactuals, In particular, we ask by how much care arrangements would change if we make the provision of care by the government more or less generous. We first classify countries by the ratio ρ of median out-of-pocket nursing-home expenses to old-age income (OECD data):

North 10-40% ($\rho_{North} = 0.225$): Sweden, Netherlands, Germany, Latvia, Denmark, Malta

Middle 50-80% ($\rho_{Middle} = 0.65$): Italy, Ireland, Slovak Republic, Luxembourg, Finland, France, Slovenia, Austria, Belgium, Lithuania, Greece

East 80-120% of median old-age income ($\rho_{East} = 1$): Croatia, Spain, Czech Rep., Poland.

In the counterfactuals, we explore the impact of applying the ratio $\{\rho_i\}$ of each region $i \in \{1, 2, 3\}$ to all other regions, essentially "transplanting" LTC systems across regions. Table 4 summarizes the results.

Region	$\rho_{North} = 0.225$	$\rho_{Middle} = 0.65$	$\rho_{East} = 1$
North	12.3%	11.1%	10.1%
Middle	8.4%	7.2%	6.4%
East	7.3%	6.7%	6.2%

Table 4: Formal-care usage when transplanting LTC systems

Generosity of LTC system (ρ varies by column. Rows show formal-care usage for each region under the three policies.

We highlight here the left column, i.e. the formal-care usage that our model would predict if countries adopted the policies in the most generous countries (North). We then compare this counterfactual to the status quo (the numbers on the diagonal). The model predicts that formal-care use would increase by 17% (=(8.4-7.2)/7.2) in the Middle countries, while it would increase by 18% (=(7.3-6.2)/6.2) in the Eastern countries. This would require a substantial, but not overwhelming, increase in the supply of nursing homes and/or formal home care.

It is interesting to ask here how much the inclusion of preference heterogeneity matters for the response of formal-care use (i.e. the price elasticity of FC) that we identify. To do so, we also estimate a *no-heterogeneity* (or "pure-economics") version of our model in which we shut down preference heterogeneity (both idiosyncratic shocks and the systematic variation in psychic informal-care costs in observables); in this alternative model, families only decide based on the effective cost of care (p_{bc}) and the opportunity costs that children face (y_{ij}), facing a *uniform* (non-idiosyncratic) disutility from formal care that we estimate to match observed FC rates. Table 5 reports the elasticity of FC use in the two models. As the point of departure, we choose $\rho_{low} = 0.35$ such that both models feature realistic levels of FC usage; we then increment the effective nursing-home costs by 10% to $\rho_{high} = 1.1\rho_{low}$. Interestingly, the elasticity in the no-heterogeneity model is easily an order of magnitude higher (about 30x) than in the baseline model, which suggests that is essential to include preference heterogeneity for care to obtain realistic estimates for care elasticities.[10]

Model	$\rho_{low} = 0.35$	$\rho_{high} = (1 + 0.1)\rho_{low}$	**elasticity**
baseline	8.75%	8.66%	**-0.11**
no heterogeneity	8.60%	6.24%	**-3.36**

Table 5: Elasticity of formal-care usage: baseline vs. no-heterogeneity model

7.2 Forecasting counterfactual: Europe in 2050

In a second set of counterfactuals, we ask how demographic and societal changes will affect care choices in the long run. Based on available forecasts for 2050, we model four broad trends:

- **Population aging** (K) will lead to a lower number of potential child caregivers. To simulate this change, we delete each child in the sample with prob. $p_K = (K - \hat{K})/K$, where $K = 2.24$ is the avg. number of potential caregivers per elderly in 2010 and $\hat{K} = 1.25$ is the number of potential caregivers per elderly forecasted for 2050. We obtain these numbers as the ratio of 45-to-65-year-olds over 70-to-90-year-olds in the EU27 from Eurostat.

- **Changes in female labor-force attachment** (y_{ij}) will increase children's opportunity cost of caregiving and thus increase formal-care use.

[10] We remark here, though, that the FC elasticity in the baseline model is likely underestimated due to two kinds of measurement error: i) we use noisy measures of opportunity costs (wages predicted by observables that omit idiosyncratic variation) and ii) we assume that the effective care cost, p_{bc}, is uniform within a country, while in reality this cost is lower for poor than for rich households in many countries (means-testing). i) and ii) downward-bias the coefficient β_{MC} on monetary costs. In future versions of the paper, we will address i) by explicitly modeling an unobserved idiosyncratic wage component and ii) by identifying country-specific specifications for p_{bc} that take into account income-specific reimbursement rates. We expect these changes to increase the coefficient on β_{MC} and thereby to diminish (but certainly not close) the gap between the elasticities in Table 5.

For the gender-wage gap, Eurostat data for the year 2010(?) tell us that the gender-wage gap in the EU27(?) was 14.3%. In our SHARE data, the gap is 17.4%. Averaging the two, we obtain a gender-wage gap of 16%. For the counterfactual, we make the optimistic assumption that this gap closes by 2050 and increment all female wages to $\tilde{y}_{ij} = 1.16 y_{ij}$, which we see as an upper bound on effects through this channel.

- **Changing family structures** (*step*) will lead to a lower supply of informal caregivers. To reflect this change, we convert each biological child in the baseline to a step child with probability $p_{step} = 1.1$.[11]

- **Higher mobility** (*dist*) of children will decrease informal-care supply. Again, it is hard to find forecasts for mobility of children. To identify a quantitatively reasonable factor by which the distance of children to parents may increase, we use SHARE data to estimate how the distance child-parent currently varies between low- and high-mobility countries. We obtain a factor 1.8 between Northern and Southern countries and thus set $\tilde{dist}_{ij} = 1.8 dist_{ij}$ in the counterfactual.[12]

Table 6 shows how formal-care usage increases when introducing the four changes sequentially. Comparing first the column "total" with "baseline", we see that the model predicts a very large increase in FC: about three-fold in the North, four-fold in the Middle, and five-fold in the East. When looking at the separate contributions of the distinct changes in the environment , we see that the decrease in the number of children is essentially the sole driver of this result (this result is independent of the order in which introduce the four changes). The change in female wages and in family structures play relatively minor roles, and the change in family structures has almost nil effects. Finally,

[11] We could not obtain direct measures on the number of non-biological children in European families. Eurostat reports that the number of single-parent families increased by 3.6% in the period 2005-2030 (Eurostat). Extrapolating to 2050 we use the factor 10%, which we see as an upper bound.

[12] Specifically, in SHARE's Wave 2, we observe that 80% (60%) of children lived $\leq$25km ($\leq$50km) away from parents in Southern countries, while these numbers were 50% (30%) in Northern countries. Averaging over the two (and assuming that Northerners' distances are a multiple of Southerner's across the distribution), we obtain a ratio North-South of $1.8 = (80/50 + 60/30)/2$.

the last column shows how much higher FC use would be if, additionally, all countries implemented Northern policies, showing a modest increase compared. This exercise thus suggests that the raw demographic forces at play will have stronger effects on care choices than policy changes. These results are still preliminary at this stage and should be interpreted with caution.

Region	baseline	#kids↓	$y_{fem}\uparrow$	step↑	distance↑ **(total)**	$+\rho_{North}$
North	12.3%	37.3%	37.6%	37.6%	**39.7%**	39.7%
Middle	7.2%	31.4%	31.5%	31.5%	**33.0%**	34.2%
East	6.2%	30.5%	30.5%	30.6%	**32.0%**	33.2%

Table 6: Formal-care use in counterfactual *Europe in 2050*

Table presents % of families opting for formal care, switching on changes to environment sequentially (i.e.#kids changes only number of children with respect to baseline, column $y_{fem}\uparrow$ shows counterfactual with a change in number of children and change in female wage etc.) Column in bold (with "total") shows counterfactual with our forecast for 2050, where all four changes are swtiched on. Last column additionally sets $\rho = \rho_{North}$ for all countries.

8 Embedding the 'Cooperative Siblings Model' into the multi-generational life-cycle model

The purpose of the cooperative siblings model as part of the project is to have a micro-foundation of who among multiple children becomes the designated caregiver in a dynamic context. Particularly, children face idiosyncratic income risk so that changes in income (or other time-varying child characteristics) can lead to changes in who provides care. In addition to the role heterogeneous opportunity costs in the labor market play, the model also takes into account various other sources of heterogeneity which we have found in our analysis above to matter for the caregiving decision and are likely subject to change over the foreseeable future. These include factors such as the number of children, the geographical distance between children and parents, and the complexity of the family structure, such as, patchwork families, where step children may differ from biological children in their propensity to provide care. Modelling these factors explicitly allows us to better predict the evolution of the supply of informal caregiving in an aging population, and to more comprehensively quantify the response of informal caregiving to changes in government policy.

One major challenge for our full dynamic structural model is how to accommodate the various sources of heterogeneity while ensuring that the dimensionality of the state space remains tractable. Our strategy for parsimony is to incorporate a "utility cost" (θ) in the child generation's state, which serves as a proxy for the psychological factors that are behind the pre-disposition to provide care. By using this approach, we introduce merely one more state variable, namely, the utility cost, in lieu of several state variables (gender, distance, step-child, etc.). It is precisely this approach that has given rise to our discrete-choice specification above which we have exploited so as to get our hands on the empirical estimates informing us about the importance of the various factors using SHARE data. These estimates are the central novel element required for embedding the Cooperative Siblings Model into our project framework.

The second part of our strategy for parsimony is that we will stick to a unitary framework for the child generation, i.e., the weights on the siblings are constant over time. Preferences are specified as in our static model but otherwise the environment should be as in Barczyk & Kredler (2018). We need to draw the utility costs from some distribution that we still need to think about how to parameterize. We aim for estimating a process for the cost θ that the marginal child incurs from the micro data that we have used in this paper.

9 Conclusion

In this paper, we study how families' caregiving arrangements for elderly parents would change in response to policy changes and demographic shifts. Importantly, we account for preference heterogeneity in providing care that differs across families. To this end, we build a model where parents and children bargain on care choices, incorporating both preference heterogeneity in caregiving mode and differences in formal care policies across countries. Our model enables us to estimate how children's characteristics affect their psychic cost of providing informal care and to predict how families' caregiving arrangements respond to various policy counterfactuals and predicted demographic shifts.

Our findings highlight that accounting for preference heterogeneity is crucial for better understanding families' caregiving decisions and for making realistic prediction about future caregiving arrangements. We show that there exists substantial heterogeneity in caregiving preferences, based on child's gender, distance from parents, and unobserved factors that cannot be explained by economic incentives. Our counterfactual exercises show that if we do not account for this preference heterogeneity, we would substantially overestimate the elasticity of formal care usage in response to policy changes. Furthermore, our model predicts that demographic changes, particularly the declining number of children due to decreasing fertility rates and marriage rates, would play a bigger role in increasing the demand for formal care compared to the impact of more generous subsidies for formal care. However, we emphasize that our current results are preliminary, as we currently do not have detailed measures of formal care costs due to a lack of available data.

We conclude by discussing the future steps for our paper. First, we plan to construct better measures for formal care costs by country and by care needs. Second, we aim to construct more granular measures of the potential wage for each SHARE child by using Eurostat microdata. We also plan to examine how measurement errors in potential wages affect our estimates. Third, we will include the non-baseline sample of SHARE survey in our estimation, after addressing data issues documented in Section 2. Finally, we plan to embed the cooperative siblings model into the multi-generational life-cycle model that would allow us to better predict the evolution of caregiving arrangements in an aging population.

References

Barczyk, D. & Kredler, M. (2018), 'Evaluating long-term care policies, taking the family seriously', *Review of Economic Studies* **85**(2), 766–809.

Barczyk, D. & Kredler, M. (2019), 'Long-term care across europe and the u.s.: The role of formal and informal care', *Fiscal Studies* **40**(2), 329–373.

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

Comas-Herrera, A., Costa-Font, J., Gori, C., di Maio, A., Patxot, C., Pickard, L., Pozzi, A., Rothgang, H. & Wittenberg, R. (2003), 'European study of long-term care expenditure: investigating the sensitivity of projections of future long-term care expenditure in germany, spain, italy and the united kingdom to changes in assumptions about demography, dependency, informal care, formal care and unit costs.', *Report to the European Commission, Employment and Social Affairs DG.PSSRU Discussion Paper 1840* .

Moise, P., Schwarzinger, M. & Um, M. (2004), 'Dementia care in 9 oecd countries: A comparative analysis', *OECD Health Working Papers* **13**.

Sundstroem, G., Johansson, L. & Hassing, L. B. (2002), 'The shifting balance of long-term care in sweden', *The Gerontologist* **42**, 350–355.

Zukewich, N. (2003), 'Unpaid informal caregiving', *Canadian Social Trends* **70**, 14–18.

A Appendix

A.1 Additional information on SHARE

A.1.1 SHARE sample selection

As documented in Section 2.1, we only use baseline samples in the current analyses due to the data issues with using panel dimensions. Baseline sample includes households that participated in SHARE for the first time in each wave. Table A1 compares sample size between full sample and baseline samples. Note that these counts are before applying any of our sample selection criteria. Further note that the reported sample sizes are not at the household level; it includes both respondents and their spouses.

Table A1: # of Respondents and Spouses, Full sample vs. baseline sample, SHARE

Wave	**Full**	**Baseline**	**Note**
1	30,419	30,419	
2	37,143	14,405	
3	28,463		Retrospective survey
4	58,000	36,717	Not used in current analyses because child caregivers cannot be identified
5	66,065	21,356	
6	68,085	10,769	
7	77,202		Retrospective survey
8	46,733	9,349	Baseline sample was added in Wave 7 (retrospective survey), but these respondents participated in the regular survey for the first time in Wave 8
Total	383,647	123,015	

Note: This table reports sample size for respondents and spouses for each wave in SHARE. “Full" column shows the sample size for *all* respondents and their spouses. “Baseline" column shows the sample size for respondents and spouses who participated in SHARE for the first time in the corresponding wave. These are raw counts before applying any sample selection criterion.

Table A2: Number of parent-child pairs after applying selection criteria, Baseline SHARE only

		After applying each sample selection criterion, subsequently					
	None	1. Sick elderly aged 65+	2. Has child(ren) aged 20-60	3. Either NH or one IC child	4. Matched with FC and wage	5. No missing X vars	6. Convert to (hh, child)-level
Count	194,860	55,255	49,013	6,465	5,262	4,684	4,120

Table A2 shows how the sample size changes after applying each of the sample selection criteria. Note that these counts are at the parent-child level, not at the household level. Column “1. Sick elderly aged 65+" shows that the sample size substantially decreases after limiting to respondents and their spouses who report having at least one mobility limitation and are aged 65+. Column “3. Either NH or one IC child" reports sample size after limiting to elderly who either (i) are in nursing home care or (ii) have one caregiving child. This selection criteria further reduces the sample size by a large margin.

In fact, Table A3 shows that most parents with at least one mobility limitation does not get formal care or is cared by any of their children. Specifically, 84% of parents aged 65+ with at least one mobility limitation are not cared for by any of their children, and 99% of such parents are not in nursing home care. In future analyses, we will examine whether these individuals are cared for by their spouses or if altering the definition of "care need" affects the frequencies of informal and formal care.

After imposing additional sample criteria as shown in Table A2, we have a final sample size of 4,120 household-child pairs and 1,829 households.

Table A3: Distribution of IC and FC for parent-child pairs, Sick elderly aged 65+ with child aged 20-60, Baseline SHARE

Informal Care		Formal Care	
Number of IC children	Frequency	NH status	Frequency
0	41,236	No	48,533
1	5,640	Yes	480
2	1,638		
3	478		
4	21		
Total	49,013	Total	49,013

Note: This table reports the distribution of number of caregiving children and formal care. The sample includes parent-child observations where parent has at least one mobility limitation and is aged 65+ and child is aged 20-60. Informal care by child is defined to be either within-household IC or outside-household IC that happens at least weekly.

Table A4: Distribution of OIC and IIC intensity, Sick elderly aged 65+ with child aged 20-60, Baseline SHARE

Outside-HH IC		Inside-HH IC	
Intensity	Frequency	Status	Frequency
None	43,120	No	48,461
Daily	1,206	Yes	465
Weekly	2,043		
Monthly	1,232		
Less Often	1,124		
Total	48,725	Total	48,926

Note: This table reports the distribution of the intensity for outside-household informal care (OIC) by children, and inside-household informal care (IIC) status. The sample includes paren-child observations where parent has at least one mobility limitation and is aged 65+ and child is aged 20-60. Note that IIC is defined to happen almost daily by definition.

Table A4 reports the distribution of OIC frequency as well as IIC status for children aged 20-60 who have parent aged 65+ with at least one mobility limitation. There are two main observations. First, most caregiving children provide OIC, as shown by much lower frequencies of IIC by children. Second, weekly OIC is the most common intensity among OIC caregiving children.

A.1.2 Notes on Children (CH) module

In this section, we outline the details of the Children (CH) module of SHARE that complicate the data cleaning process.

1. Only one spouse answers questions in the CH module

As a result, children's information is missing for non-responding spouses in each wave. We need to import children's information for non-responding spouses from the responses of the responding spouses. The respondent for the CH module can change over the panel.

2. Many questions are not asked again from one wave to another if the responses are the same

Information including the child's distance from parent and education are not asked again in the subsequent waves if the responses have not changed. Child's distance is recorded again if child moves, but not when parent moves. This complicates measuring the current distance between parents and children in non-baseline surveys.

3. Children may not have same index across different waves.
For instance, Child 1 in wave 1 may be listed as Child 3 in wave 4. This complicates the data cleaning process, especially since many questions are not repeated in subsequent waves. To track the same child across waves, we need to rely on the child's gender and year of birth. However, in cases involving twins, accurately tracking the same child over time may not be possible.

4. In waves 1 and 2, some information are only recorded up to 4 children
Characteristics like child's education, stepchild status, and employment are recorded only up to 4 children in waves 1 and 2. For subsequent waves, these characteristics are recorded for all children. Hence, for waves 1 and 2, we have missing information for children for households with more than 4 children. Furthermore, these 4 children are not necessarily child indexed 1, 2, 3, 4. Hence, it is crucial to carefully check which child's information is being recorded in waves 1 and 2.

The above four points are the main challenges regarding the CH module. In addition to these points, there are minor challenges including the reported number of children being different from the number of children's characteristics, etc. It is crucial to check each variable carefully in the data cleaning process.

A.1.3 Notes on Social Support (SP) module

In this section, we outline the details of the Social Support (SP) module of SHARE that complicate the data cleaning process.

1. The questions about informal care differ across waves
Waves 1, 2, and 5 share a similar format of questions regarding informal care, while waves 6 and 8 also follow a similar format. Unlike other waves, wave 4 does not have any questions that identify *which* child provided informal care.

2. There are different sets of questions for caregiver within the household and outside the household

See Table 1 to check which questions are available for each wave.

3. Some families do not correctly report OIC and IIC caregiving children.
For example, some families report the same child for different OIC caregivers (which can be reported up to 3 caregivers). Furthermore, some families report same child as being both OIC and IIC caregiver.

A.2 Additional details on potential wage construction

Our goal is to construct the potential income for each SHARE child based on country, gender, education, and year. To this end, we need imputation strategies to address several challenges. Below, we describe the challenges and the strategies to address them.

1. Dealing with inconsistent education categories: First, education categories differ across survey years in Eurostat, as shown in Table A5. For consistency, we need to construct synchronized educational categories that are consistent across years.

Table A5: Education Categories, Eurostat's Structure of Earnings Survey

Survey Year	Classification	Education Categories
2006	ISCED 1997	Levels 0-1, Level 2, Level 3-4, Level 5A, Level 5B, Level 6
2010	ISCED 1997	Levels 0-1, Level 2, Level 3-4, Level 5A, Level 5B, Level 6
2014	ISCED 2011	Levels 0-2, Levels 3-4, Levels 5-6, Levels 7-8
2018	ISCED 2011	Levels 0-2, Levels 3-4, Levels 5-8

Note: This table reports educational categories in Eurostat's structure of earnings survey for each year. For more information about what each category means and how to map between ISCED 1997 and ISCED 2011, click [ILO link].

We construct the potential income for synchronized education categories based on the broadest education categorization – which is in survey year 2018. Specifically, the synchronized education categories have 3 levels: (1) ISCED 2011 Levels 0-2: Less than lower secondary education, (2) ISCED 2011 Levels 3-4: Upper secondary and post-secondary non-tertiary education, (3) ISCED 2011 Levels 5-8: College education or more. The mapping between ISCED 1997 and 2011 is done using the ILO classification [ILO link].

To construct wages based on the synchronized education categories, we calculate weighted averages of multiple sub-categories as needed. As a demonstration, consider the survey year 2014. We need to combine gender wages for Levels 5-6 and Levels 7-8 to create the gender wages for the synchronized category Levels 5-8. How we combine is by taking the weighted average, where the weights are the share of workers in each education category relative to the total number of workers for the combined categories. Specifically, for each gender g and country c, the weighted average for education levels 5-8 in year 2014 is calculated as follows:

$$\begin{aligned} Wage_{g,\, c,\, year=2014,\, edu=5-8} = & \underbrace{\left(\frac{NumEmployees_{g,\, c,\, year=2014,\, edu=5-6}}{NumEmployees_{g,\, c,\, year=2014,\, edu=5-8}}\right)}_{\text{Weight for level 5-6}} Wage_{g,\, c,\, year=2014,\, edu=5-6} \\ & + \underbrace{\left(\frac{NumEmployees_{g,\, c,\, year=2014,\, edu=7-8}}{NumEmployees_{g,\, c,\, year=2014,\, edu=5-8}}\right)}_{\text{Weight for level 7-8}} Wage_{g,\, c,\, year=2014,\, edu=7-8} \end{aligned}$$

The synchronization procedure is similarly applied to other education categories and survey years.

2. Dealing with missing wages: To apply the synchronization procedure above, ideally, the data should have full information about wages for each gender, education category, country, and year. However, Eurostat data lacks wage information for some cells in year 2006 and 2010. For years 2014 and 2018, we have full information on wages. We document our imputation strategies for the missing wages for several cases:

- **Case 1:** Only one of female or male wages is missing for country c, education e, and year y

  To demonstrate, consider a scenario where only the female wage is missing. In this case, we impute the female wage using the male wage and the total wage. We assume that the total wage is the weighted average of male wage and female wage:

$$TotalWage_{c,y,e} = \left(\frac{MaleEmployees_{c,y,e}}{TotalEmployees_{c,y,e}}\right) MaleWage_{c,y,e} + \left(\frac{FemaleEmployees_{c,y,e}}{TotalEmployees_{c,y,e}}\right) FemaleWage_{c,y,e}$$

When $FemaleEmployees_{c,y,e}$ is missing, we impute this using the following assumption:

$MaleEmployees_{c,y,e} + FemaleEmployees_{c,y,e} = TotalEmployees_{c,y,e}$.

Once we impute $FemaleEmployees_{c,y,e}$, we can impute $FemaleWage_{c,y,e}$ using the above formula. Imputation for cases where only the male wage is missing is performed similarly.

- **Case 2:** Both female and male wages are missing for country c, education e, and year y

  In these cases, we impute missing wages using information on other years. For example, let's consider that country c has missing gender wages for education e for the year 2010, but not for the year 2006. We impute the missing wages in 2010 using the following formula:

$$\underbrace{GenderWage_{c,y=2010,e}}_{Imputed} = GrowthGenderWage^{2006-2010}_{c=EU,e} \underbrace{GenderWage_{c,y=2006,e}}_{Observed} \tag{9}$$

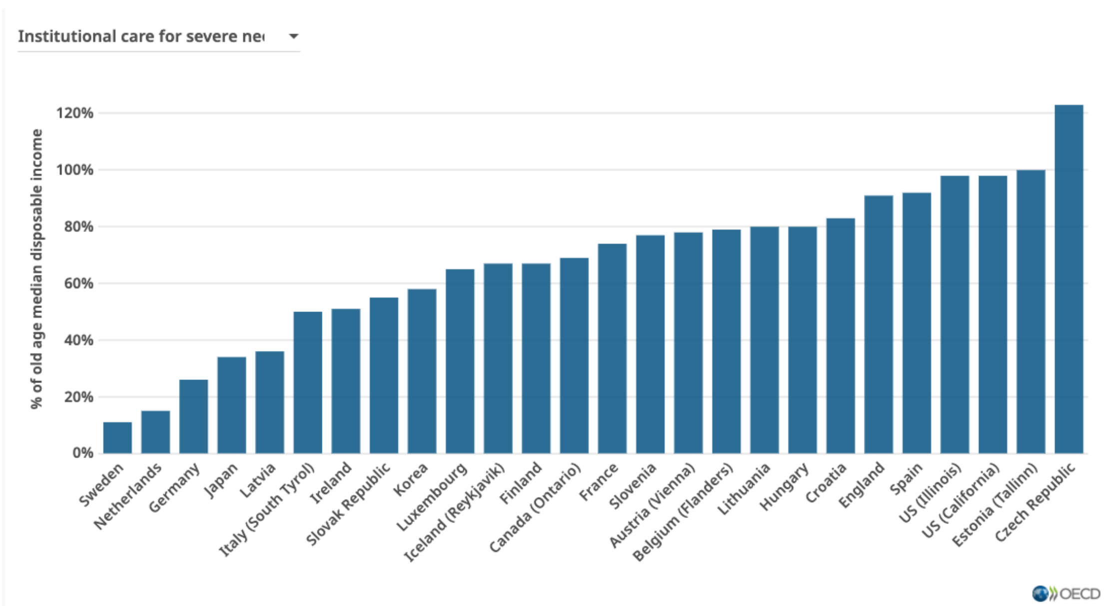
  where $GrowthGenderWage^{2006-2010}_{c=EU,e}$ is the gender wage growth rate between 2006 and 2010 for education e at the EU-level. Note that there is no wage information at the EU-level.

  The cases where only wages for 2006 are missing, but not for year 2010, imputation is done similarly. For the cases where both wages for 2006 and 2010 are missing, we address the issue in the next step.

3. Dealing with differing survey years between SHARE and Eurostat: Even after addressing missing values and synchronizing education categories, we still cannot match Eurostat wages to SHARE children due to differing survey years. To resolve this, we linearly interpolate and extrapolate potential wages for each gender g, education e, and country c to fill wage information for all years between 2004 and 2018. Note that for cases where gender wages are missing for both 2006 and 2010, the interpolation/extrapolation procedures also fill these gaps using wage information from 2014 and 2018, which are available for all cases.

A.3 Additional details on formal care cost construction

Figure A1: Out-of-pocket costs of long-term care as a share of old age median disposable income after public support, for care recipients holding no net wealth, by severity of needs and care setting



Source: OECD "Social protection for older people with long-term care needs" [Link to the web source]

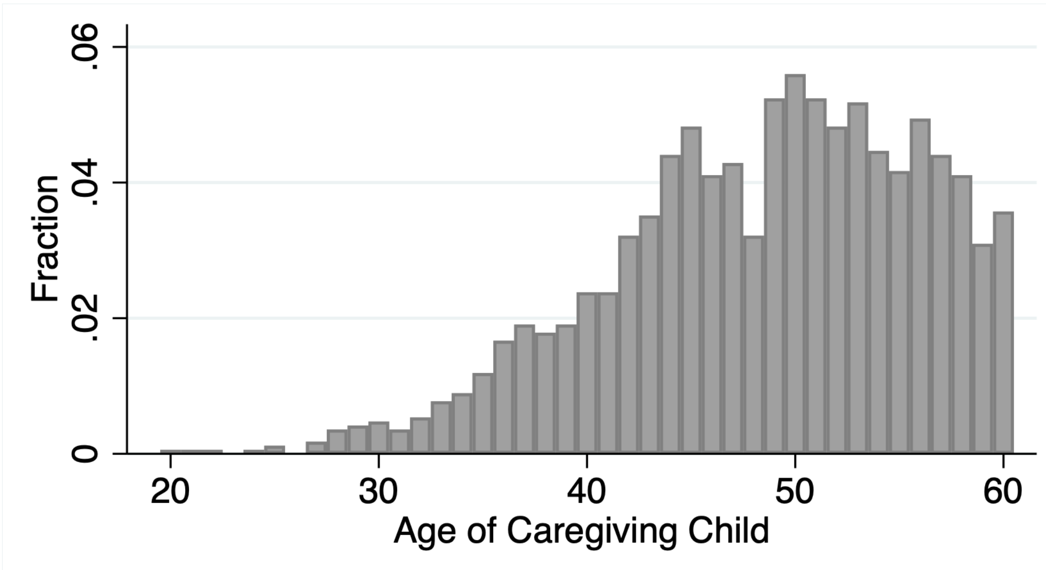
Table A6: Counts by Nursing Home Status for Each Country

Country	Not using NH	Using NH	Total
Austria	51	3	54
Germany	196	18	214
Sweden	74	9	83
Netherlands	45	17	62
Spain	149	5	154
Italy	136	12	148
France	99	7	106
Denmark	49	19	68
Greece	171	0	171
Belgium	142	16	158
Czech	174	17	191
Poland	160	2	162
Ireland	18	2	20
Luxembourg	18	14	32
Slovenia	52	1	53
Croatia	22	0	22
Lithuania	57	2	59
Finland	14	2	16
Latvia	17	1	18
Malta	11	3	14
Slovakia	24	0	24

Note: This table shows the number of households utilizing nursing homes (NH) for sick elderly parents versus those not using them in each country, based on our estimation sample. Details on how the estimation sample is selected are provided in Section 2.1.

A.4 Additional Statistics

Figure A2: Age distribution of caregiving children in the SHARE sample



Note: This figure presents the age distribution of caregiving children in the SHARE sample. For details on how the estimation sample was selected, please refer to Section 2.1.

A.5 Proofs for the Model Solutions

A.5.1 Sibling cooperation stage

We provide detailed proofs for deriving solutions from the child generation's problem:

$$\begin{aligned}\max_{c_1,c_2,ic_1} \quad & \mu_1 \frac{1}{1-\gamma}(c_1-\theta_1 ic_1)^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}(c_2-\theta_2(1-ic_1))^{1-\gamma} \\ s.t. \quad & c_1+c_2=(1-ic_1)y_1+ic_1y_2+Q\end{aligned}$$

Case 1: Child 1 is the caregiver ($ic_1 = 1$). Then the household problem becomes

$$\begin{aligned}\max_{c_1,c_2} \quad & \mu_1 \frac{1}{1-\gamma}(c_1-\theta_1)^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}(c_2)^{1-\gamma} \\ s.t. \quad & c_1+c_2=y_2+Q\end{aligned}$$

When solving the above by taking the first order conditions with respect to c_1 and c_2 respectively, we get the following solutions for optimal c_1^* and c_2^*, conditional on $ic_1 = 1$:

$$\begin{aligned}c^*_{1|ic_1=1} &= \left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_2+Q+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\theta_1) \\ c^*_{2|ic_1=1} &= \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_2+Q-\theta_1)\end{aligned}$$

We then plug the optimal consumption into the utility function to obtain the conditional indirect utility function for Case 1.

$$
\begin{aligned}
W_{ic_1=1} = \quad & \mu_1 \frac{1}{1-\gamma}(c^*_{1|ic_1=1} - \theta_1)^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}(c^*_{2|ic_1=1})^{1-\gamma} \\
= \quad & \mu_1 \frac{1}{1-\gamma}\left(\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_2 + Q + \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\theta_1) - \theta_1\right)^{1-\gamma} \quad + \quad \mu_2 \frac{1}{1-\gamma}\left(\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_2+Q-\theta_1)\right)^{1-\gamma} \\
= \quad & \mu_1 \frac{1}{1-\gamma}\left(\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_2 + Q + \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\theta_1 - \left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)\theta_1)\right)^{1-\gamma} \\
& + \mu_2 \frac{1}{1-\gamma}\left(\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_2+Q-\theta_1)\right)^{1-\gamma} \\
= \quad & \mu_1 \frac{1}{1-\gamma}\left(\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_2+Q-\theta_1)\right)^{1-\gamma} \quad + \quad \mu_2 \frac{1}{1-\gamma}\left(\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_2+Q-\theta_1)\right)^{1-\gamma} \\
= \quad & \mu_1 \frac{1}{1-\gamma}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}(y_2+Q-\theta_1)^{1-\gamma} \quad + \quad \mu_2 \frac{1}{1-\gamma}\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1-\gamma}{\gamma}}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}(y_2+Q-\theta_1)^{1-\gamma} \\
= \quad & \left(\mu_1 \frac{1}{1-\gamma}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1} + \quad \mu_2 \frac{1}{1-\gamma}\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1-\gamma}{\gamma}}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}\right)(y_2+Q-\theta_1)^{1-\gamma} \\
= \quad & \left(\frac{1}{1-\gamma}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}\left(\mu_1 + \mu_2\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1-\gamma}{\gamma}}\right)\right)(y_2+Q-\theta_1)^{1-\gamma} \\
= \quad & \left(\frac{1}{1-\gamma}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}\left(\mu_1 + \frac{\mu_2^{\frac{1}{\gamma}}}{\mu_1^{\frac{1-\gamma}{\gamma}}}\right)\right)(y_2+Q-\theta_1)^{1-\gamma} \\
= \quad & \left(\frac{1}{1-\gamma}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}\left(\mu_1(1+\mu_2^{\frac{1}{\gamma}}\mu_1^{\frac{\gamma-1}{\gamma}-1})\right)\right)(y_2+Q-\theta_1)^{1-\gamma} \\
= \quad & \left(\frac{1}{1-\gamma}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}\mu_1\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)\right)(y_2+Q-\theta_1)^{1-\gamma} \\
= \quad & \left(\frac{1}{1-\gamma}\mu_1\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma}\right)(y_2+Q-\theta_1)^{1-\gamma} \\
= \quad & \left(\frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)\right)^{\gamma}\right)(y_2+Q-\theta_1)^{1-\gamma} \\
= \quad & \frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}}\right)^{\gamma}(y_2+Q-\theta_1)^{1-\gamma}
\end{aligned}
$$

Case 2: Child 2 is the caregiver ($ic_1 = 0$). Then the household problem becomes

$$
\begin{aligned}
\max_{c_1, c_2} \quad & \mu_1 \frac{1}{1-\gamma}(c_1)^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}(c_2 - \theta_2)^{1-\gamma} \\
s.t. \quad & c_1 + c_2 = y_1 + Q
\end{aligned}
$$

When solving the above by taking the first order conditions with respect to

c_1 and c_2 respectively, we get the following solutions for optimal c_1^* and c_2^*, conditional on $ic_1 = 0$:

$$c^*_{1|ic_1=0} = \left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}} \left(1 + \left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{-1} (y_1 + Q - \theta_2)$$
$$c^*_{2|ic_1=0} = \left(1 + \left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{-1} (y_1 + Q + \left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}} \theta_2)$$

We then plug the optimal consumption into the utility function to obtain the conditional indirect utility function for Case 2.

$$\begin{aligned}
W_{ic_1=0} =& \quad \mu_1 \frac{1}{1-\gamma}(c^*_{1|ic_1=0})^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}(c^*_{2|ic_1=0} - \theta_2)^{1-\gamma} \\
=& \quad \mu_1 \frac{1}{1-\gamma}\left(\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\left(1+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_1+Q-\theta_2)\right)^{1-\gamma} + \quad \mu_2 \frac{1}{1-\gamma}\left(\left(1+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_1+Q+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\theta_2) - \theta_2\right)^{1-\gamma} \\
=& \quad \mu_1 \frac{1}{1-\gamma}\left(\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\left(1+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_1+Q-\theta_2)\right)^{1-\gamma} + \quad \mu_2 \frac{1}{1-\gamma}\left(\left(1+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_1+Q-\theta_2)\right)^{1-\gamma} \\
=& \quad \mu_1 \frac{1}{1-\gamma}\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1-\gamma}{\gamma}}\left(1+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}(y_1+Q-\theta_2)^{1-\gamma} + \quad \mu_2 \frac{1}{1-\gamma}\left(1+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}(y_1+Q-\theta_2)^{1-\gamma} \\
=& \quad \left(\frac{1}{1-\gamma}\left(1+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}\right)\left(\mu_1\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1-\gamma}{\gamma}} + \mu_2\right)(y_1+Q-\theta_2)^{1-\gamma} \\
=& \quad \left(\frac{1}{1-\gamma}\left(1+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}\right)\left(\frac{\mu_1^{\frac{1}{\gamma}}}{\mu_2^{\frac{1-\gamma}{\gamma}}} + \mu_2\right)(y_1+Q-\theta_2)^{1-\gamma} \\
=& \quad \left(\frac{1}{1-\gamma}\left(1+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}\right)\left(\mu_2\left(1+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)\right)(y_1+Q-\theta_2)^{1-\gamma} \\
=& \quad \frac{1}{1-\gamma}\mu_2\left(1+\left(\frac{\mu_1}{\mu_2}\right)^{\frac{1}{\gamma}}\right)^{\gamma}(y_1+Q-\theta_2)^{1-\gamma} \\
=& \quad \frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}}\right)^{\gamma}(y_1+Q-\theta_2)^{1-\gamma}
\end{aligned}$$

More on optimal consumption allocation: Here, we show how optimal consumption allocation depends on Pareto weights. When child 1 is chosen as a caregiver, then as shown above, the resulting consumption allocation is:

$$c^*_{1|ic_1=1} = \left(1 + \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1} (y_2 + Q + \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}} \theta_1)$$
$$c^*_{2|ic_1=1} = \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}} \left(1 + \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1} (y_2 + Q - \theta_1)$$

Let $\mu = \frac{\mu_1}{\mu_2}$ denote the relative Pareto weight for child 1. Then, the consumption allocation is rewritten as:

$$
\begin{aligned}
c^*_{1|ic_1=1} &= \left(1 + \left(\frac{1}{\mu}\right)^{\frac{1}{\gamma}}\right)^{-1} (y_2 + Q + \left(\frac{1}{\mu}\right)^{\frac{1}{\gamma}} \theta_1) \\
c^*_{2|ic_1=1} &= \left(\frac{1}{\mu}\right)^{\frac{1}{\gamma}} \left(1 + \left(\frac{1}{\mu}\right)^{\frac{1}{\gamma}}\right)^{-1} (y_2 + Q - \theta_1)
\end{aligned}
$$

How do these consumptions change when μ changes? If we take the derivative of child 1's optimal consumption with respect to μ,

$$
\begin{aligned}
\frac{\partial c^*_{1|ic_1=1}}{\partial \mu} &= -(1+\mu^{-\frac{1}{\gamma}})^{-2}\Big(-\frac{1}{\gamma}\Big)\mu^{\frac{-1-\gamma}{\gamma}}(y_2+Q+\mu^{-\frac{1}{\gamma}}\theta) + (1+\mu^{-\frac{1}{\gamma}})^{-1}\theta_1\Big(-\frac{1}{\gamma}\Big)\mu^{\frac{-1-\gamma}{\gamma}} \\
&= \frac{1}{\gamma}(1+\mu^{-\frac{1}{\gamma}})^{-2}\mu^{\frac{-1-\gamma}{\gamma}}(y_2+Q+\mu^{-\frac{1}{\gamma}}\theta) - \frac{1}{\gamma}(1+\mu^{-\frac{1}{\gamma}})^{-1}\mu^{\frac{-1-\gamma}{\gamma}}\theta_1 \\
&= \frac{1}{\gamma}\mu^{\frac{-1-\gamma}{\gamma}}(1+\mu^{-\frac{1}{\gamma}})^{-2}\Big(y_2+Q+\mu^{-\frac{1}{\gamma}}\theta_1 - (1+\mu^{-\frac{1}{\gamma}})\theta_1\Big) \\
&= \frac{1}{\gamma}\mu^{\frac{-1-\gamma}{\gamma}}(1+\mu^{-\frac{1}{\gamma}})^{-2}(y_2+Q-\theta_1)
\end{aligned}
$$

Note that as long as $y_2 + Q - \theta_1 > 0$, then $\frac{\partial c^*_{1|ic_1=1}}{\partial \mu}$ is positive. This means that as long as $y_2 + Q - \theta_1 > 0$, the higher Pareto weight for child 1 leads to higher consumption for child 1. This outcome is intuitive, as we expect the child with greater decision-making power to have higher consumption compared to their siblings. Regarding the condition $y_2 + Q - \theta_1 > 0$, note that this implies that the utility cost of caregiving (θ_1) must be lower than the total income available to the child generation ($y_2 + Q$) when child 1 provides care. This is likely to hold in most cases, because it is unreasonable for child 1 to provide care if the utility cost is so huge that the cost is greater than the total income available.

Similarly, if we take the derivative of child 2's optimal consumption with respect to μ,

$$
\frac{\partial c^*_{2|ic_1=1}}{\partial \mu} = -\frac{1}{\gamma}\mu^{\frac{-1-\gamma}{\gamma}}(1+\mu^{-\frac{1}{\gamma}})^{-2}(y_2+Q-\theta_1)
$$

As long as $y_2 + Q - \theta_1 > 0$, then $\frac{\partial c^*_{2|ic_1=1}}{\partial \mu}$ is negative. This means that as long as

$y_2+Q-\theta_1>0$, the higher Pareto weight for child 1 leads to lower consumption for child 2.

A.5.2 Bargaining stage

Parent's surplus. The parent's surplus function is:

$$S^p(Q) = \underbrace{u^p(c^{p*}_{ic^k=1})}_{\text{IC takes place}} - \underbrace{u^p(c^{p*}_{ic^k=0})}_{\text{FC takes place}}$$

We get the solution for the parent's surplus function by solving the parent's utility maximization problem (i) when IC takes place and (ii) when FC takes place, respectively.

The parent's utility maximization problem (i) when IC takes place is:

$$\underset{c^p}{max}\ \frac{1}{1-\gamma}c^{p1-\gamma} \quad s.t. \quad c^p = y^p - Q$$

In this case, $c^{p*}_{ic^k=1} = y^p - Q$, because the parent's utility (u^p) goes up as c^p goes up, leading to a corner solution. The indirect utility in this case is therefore:

$$u^p(c^{p*}_{ic^k=1}) = \frac{1}{1-\gamma}(y^p - Q)^{1-\gamma}$$

The parent's problem (ii) when IC does not take place is:

$$\underset{c^p}{max}\ \frac{1}{1-\gamma}(c^p - C_f)^{1-\gamma} \quad s.t. \quad c^p = y^p - p_{bc}$$

Similarly, $c^{p*}_{ic^k=0} = y^p - p_{bc}$ as the parent's utility (u^p) goes up as c^p goes up, leading to a corner solution. The indirect utility in this case is therefore:

$$u^p(c^{p*}_{ic^k=1}) = \frac{1}{1-\gamma}(y^p - (p_{bc} + C_f))^{1-\gamma}$$

Combining the two solutions above, we get the following surplus function for the parent:

$$S^p(Q) = \frac{1}{1-\gamma}(y^p - Q)^{1-\gamma} - \frac{1}{1-\gamma}(y^p - (p_{bc} + C_f))^{1-\gamma}$$

Child generation's surplus. The surplus function is specified as follows:

$$S^k(Q) = \underbrace{U^k(c^{k*}_{1|ic^k=1}, c^{k*}_{2|ic^k=1})}_{\text{IC takes place}} - \underbrace{U^k(c^{k*}_{1|ic^k=0}, c^{k*}_{2|ic^k=0})}_{\text{formal care takes place}}$$

The solution for the child generation's surplus function is obtained by solving the child generation's problem (i) when IC takes place and (ii) when FC takes place, respectively.

The problem for the child generation (i) when IC takes place is:

$$\max_{c_1^k, c_2^k} \quad \mu_1 \frac{1}{1-\gamma}(c_1^k - \theta_1)^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}(c_2^k)^{1-\gamma} \quad s.t. \quad c_1^k + c_2^k = y_2^k + Q$$

Here, child generation's indirect utility is exactly the same as in the Case 1 of Appendix Section A.5.1, because we assumed that child 1 is the chosen caregiver from the sibling cooperation stage. Hence, the indirect utility function is:

$$U^k({c^k}^*_{1|ic^k=1}, {c^k}^*_{2|ic^k=1}) = \frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}}\right)^{\gamma}(y_2 + Q - \theta_1)^{1-\gamma}$$

The child generation's problem (ii) when IC does not take place is:

$$\max_{c_1^k, c_2^k} \quad \mu_1 \frac{1}{1-\gamma}(c_1^k)^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}(c_2^k)^{1-\gamma} \quad s.t. \quad c_1^k + c_2^k = y_1^k + y_2^k$$

When solving the above by taking the first order conditions with respect to c_1 and c_2 respectively, we get the following solutions for optimal c_1^* and c_2^*, conditional on $ic^k = 0$

$$
\begin{aligned}
{c^k}^*_{1|ic^k=0} &= \left(1 + \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_1 + y_2) \\
{c^k}^*_{2|ic^k=0} &= \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\left(1 + \left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_1 + y_2)
\end{aligned}
$$

We then plug the optimal consumption into the utility function to obtain the conditional indirect utility function:

$$\begin{aligned}
U^k({c^k}^*_{1|ic^k=0}, {c^k}^*_{2|ic^k=0}) = & \quad \mu_1 \frac{1}{1-\gamma}({c^k}^*_{1|ic^k=0})^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}({c^k}^*_{2|ic^k=0})^{1-\gamma} \\
= & \quad \mu_1 \frac{1}{1-\gamma}\left(\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_1+y_2)\right)^{1-\gamma} + \mu_2 \frac{1}{1-\gamma}\left(\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_1+y_2)\right)^{1-\gamma} \\
= & \quad \left(\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_1+y_2)\right)^{1-\gamma}\left(\frac{\mu_1}{1-\gamma} - \frac{\mu_2}{1-\gamma}\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1-\gamma}{\gamma}}\right) \\
= & \quad \left(\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_1+y_2)\right)^{1-\gamma}\left(\frac{1}{1-\gamma}\left(\mu_1 + \mu_2\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1-\gamma}{\gamma}}\right)\right) \\
= & \quad \left(\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{-1}(y_1+y_2)\right)^{1-\gamma}\left(\frac{1}{1-\gamma}\mu_1\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)\right) \\
= & \quad \left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma-1}(y_1+y_2)^{1-\gamma}\left(\frac{1}{1-\gamma}\mu_1\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)\right) \\
= & \quad \frac{1}{\gamma}\mu_1\left(1+\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma}(y_1+y_2)^{1-\gamma} \\
= & \quad \frac{1}{\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_1^{\frac{1}{\gamma}}\left(\frac{\mu_2}{\mu_1}\right)^{\frac{1}{\gamma}}\right)^{\gamma}(y_1+y_2)^{1-\gamma} \\
= & \quad \frac{1}{\gamma}(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}})^{\gamma}(y_1+y_2)^{1-\gamma}
\end{aligned}$$

Combining the two solutions above, we get the following surplus function for the child generation:

$$S^k(Q) = \frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}}\right)^{\gamma}(y_2^k + Q - \theta_1)^{1-\gamma} - \frac{1}{1-\gamma}\left(\mu_1^{\frac{1}{\gamma}} + \mu_2^{\frac{1}{\gamma}}\right)^{\gamma}(y_1^k + y_2^k)^{1-\gamma}$$

A.6 MLE with wage residuals

MLE with unobserved wage residuals We denote by y_{ij} the (true) opportunity cost of child j in family i, which we cannot observe in the SHARE data. We decompose the opportunity cost into an observed and an unobserved component by writing

$$\ln y_{ij} = \bar{y}_{ij} + \eta_{ij} \tag{10}$$

We impute $\bar{y}_{ij} \equiv \mathbb{E}[\ln y_{ij}|X_{ij}]$ using our estimated coefficients from the Eurostat data based on child j's observables in the SHARE data (gender, country, distance, step status) contained in X_{ij}. η_{ij} is an unobserved wage residual with mean zero.

We assume that the wage residuals among siblings are jointly normally distributed and denote by $\phi_K(\eta_{i1}, \eta_{i2}, \ldots, \eta_{iK})$ the joint normal PDF. The variance of the unobserved wage residual is the same for all children, $\mathbb{E}[\eta_{ij}^2] = \sigma_\eta^2$ for all j; note that we can obtain σ_η^2 from a Mincer regression in our Eurostat data. We allow that the unobservable wage residuals are correlated among siblings, $\mathbb{E}[\eta_{ij}\eta_{ik}] = \rho_\eta \sigma_\eta^2$ for $j \neq k$, where $\rho_\eta \in [-1, 1]$ is the correlation of siblings' earnings. Finally, we assume that η_{ik} is independent of observables $\{X_{ij}\}_{j=1}^{N_i}$[13], where N_i is the number of children in family i.

We will now derive the likelihood function for the care choice in family i given observables $X_i \equiv \{X_{ij}\}_{j=1}^{K_i}$ in family i. The probability that care choice j is chosen given a residual wage vector η is

$$\tilde{P}_j(X_i, \eta; \beta) = \frac{e^{V_{ij}}}{\sum_{j=0}^{K_i} e^{V_{ij}}}, \tag{11}$$

Now V_{ij} for the children $j = 1, \ldots, N_i$ is given by

$$V_{ij} = -\exp\left(\bar{y}_{ij} + \eta_{ij}\right) - \left(\beta_1^* \text{gender}_{ij} + \beta_2^* \text{dist}_{ij} + \beta_3^* \text{bio}_{ij}\right) \tag{12}$$

The crucial difference is that here the shock η enters into the opportunity-cost term. For formal care, V_{i0} is as before.

[13] In Eq. 10, we see that $\mathbb{E}[\eta_{ij}|X_{ij}] = 0$, but independence, also across children, is a slightly stronger (but not unreasonable) assumption.

In our MLE estimation above we were done at this stage and in a position to construct the likelihood function. Here we still have to deal with the unknown wage residuals, but for which we have imposed distributional assumptions. Since we have assumed η_{ij} to be independent of observables X_i, the likelihood that care option $j \in \{0, 1, \ldots, K_i\}$ is chosen in family i conditional on observables X_i and given parameter vector β is

$$P_{ij} = \int \tilde{P}_j(X_i, \eta; \beta)\phi_{K_i}(\eta)\mathrm{d}\eta. \tag{13}$$

where the integration is over all possible combinations of child wage shocks $\eta = (\eta_{i1}, \ldots, \eta_{K_i})$ using the multi-variate normal density function ϕ_{K_i}.

We approximate the integral in Eq. 13 using multivariate Gaussian quadrature.[14] The integral is approximated by a weighted sum of function values evaluated at quadrature nodes. For example, in the bivariate case we have that

$$P_{ij} \approx \sum_{k_1=1}^{N} \sum_{k_2=1}^{N} w_{k_1} w_{k_2} \tilde{P}_j(X_i, x_{k_1}, x_{k_2}; \beta)$$

where w_{k_i} are quadrature weights and x_{k_i} are quadrature nodes. The quadrature nodes take the form:

$$(x_{k_1}, x_{k_2}) = (\sigma_\eta z_{k_1}, \rho\sigma_\eta z_{k_1} + \sigma_\eta\sqrt{1-\rho^2} z_{k_2})$$

where (z_{k_1}, z_{k_2}) are the two-dimensional Gaussian quadrature nodes for a standard normal variable, and (w_{k_1}, w_{k_2}) are the corresponding quadrature weights.

Why? In order to use Gaussian quadrature we need to express the random variable $\eta = (\eta_1, \eta_2)$ in terms of the standard normal variable $z = (z_1, z_2)$ in the following way:

$$\eta = \mu + Lz = \begin{bmatrix} \sigma_\eta & 0 \\ \rho\sigma_\eta & \sigma_\eta\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

The lower-triangular matrix L is the Cholesky square root of the covariance

[14] See https://www.r-bloggers.com/2015/09/notes-on-multivariate-gaussian-quadrature-with-r-code/ for an implementation in R.

matrix, $\Sigma = LL^T$. The Cholesky decomposition of the covariance matrix is given by

$$\Sigma = \begin{bmatrix} \sigma_\eta^2 & \rho\sigma_\eta^2 \\ \rho\sigma_\eta^2 & \sigma_\eta^2 \end{bmatrix} = \begin{bmatrix} \sigma_\eta & 0 \\ \rho\sigma_\eta & \sigma_\eta\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} \sigma_\eta & \rho\sigma_\eta \\ 0 & \sigma_\eta\sqrt{1-\rho^2} \end{bmatrix}$$

where the lower triangular matrix L is given by

$$L = \begin{bmatrix} \sigma_\eta & 0 \\ \rho\sigma_\eta & \sigma_\eta\sqrt{1-\rho^2} \end{bmatrix}$$

We can obtain the quadrature nodes and weights from pre-existing computational routines. Likely, we just have to supply the covariance matrix and the number of nodes per dimension. But we have to be mindful about the number of nodes that we stipulate. For example, if we choose $n = 10$ quadrature nodes per dimension, then we have $n^4 = 10,000$ nodes in families with 4 children.

We sketch here the modification of our baseline MLE algorithm:

1. Before the MLE routine:
   (a) Divide families into groups K by the number of children $K \in \{1, \ldots, K_{max}\}$.
   (b) For each group K, obtain a set of quadrature nodes $\{\eta_{K,s}\}_{s=1}^{S_K}$ and weights $\{\omega_{K,s}\}_{s=1}^{S_K}$ from the multivariate Gauss-Hermite quadrature rules for the K-dimensional normal distribution. We aim for S_k on the order of 10 to 100 for a good trade-off between speed and precision. We can use pruning, i.e. set very small weights to zero.
   (c) Compute wages $y_{i,j,s} = \exp(\bar{y}_{i,j} + \eta_{K_i,s})$ for all quadrature nodes s for each child j.

2. Inside the MLE routine:
   (a) Compute the odds ratios $E_{i,j,s}$ for each quadrature node (for each child) from (12), using the wage $y_{i,j,s}$ computed in Step 1(c).
   (b) Use the logit formula (11) to compute the likelihood $\tilde{P}_{i,j,s}$ of observing outcome $j^*(i) = j$ for the quadrature node s.

(c) To obtain the likelihood of outcome j in familiy i, approximate the integral in (13) by the quadrature formula

$$\mathbb{P}(j^*(i) = j | X_i, \beta) \simeq \sum_{s=1}^{S_{K_i}} \omega_{K_i,s} \tilde{P}_{i,j,s}. \tag{14}$$

We expect this procedure to yield higher estimates for β_1, the coefficient on monetary costs, than for the baseline case without wage residuals ($\sigma_\eta^2 = 0$). Why? Consider the care-choice probabilities that the model predicts as a function of $\bar{y}_{ij}$. The logistic function is i) convex on the lower part, increasing predicted probabilities when taking into account shocks η_{ij} (and thus averaging over neighboring wages), and ii) concave on the upper part, decreasing predicted probabilities when averaging over neighboring wages. Thus, we expect the function $\mathbb{P}$ to be flatter in $\bar{y}_{ij}$ as we increase σ_η^2 to above zero, for a fixed β. Thus, to match the observed slope in IC probabilites in the approximate opportunity cost, $\bar{y}_{ij}$, the procedure will identify a higher β.