Australian School of Business

Working Paper

Never Stand Still Australian School of Business

Australian School of Business Research Paper No. 2012 ECON 15

Externality and Strategic Interaction in the Location Choice of Siblings under Altruism toward Parents

Meliyanni Johar
Shiko Maruyama

Externality and Strategic Interaction in the Location Choice of Siblings under Altruism toward Parents

Meliyanni Johar and Shiko Maruyama*,†

February 10, 2012

Abstract

When siblings wish for the well-being of their elderly parents, the cost of caregiving and long-term commitment creates a free-rider problem among siblings. We estimate a sequential game to investigate externality and strategic interaction among adult siblings regarding their location choice relative to their elderly parents. Using the US Health and Retirement Survey, we find a positive externality and strategic interaction. The first-mover advantage of eldest children and the prisoner's dilemma are likely to exist but their magnitudes are negligible compared with inefficiency in joint utility. Inefficiency is large in a family with an educated, widowed mother and with educated siblings who are younger (relative to parents), married, and similar to each other. Had siblings fully internalized externality and jointly maximized utility sum in 2010, 17% more parents with multiple children would have had a child nearby. Public policies that reduce children's private costs may enhance social welfare.

*Corresponding author. E-mail: s.maruyama@unsw.edu.au

†The authors gratefully acknowledge financial support from the Australian Research Council's *Discovery Projects* funding scheme (project number DP110100773).

1 Introduction

While adult children wish for the well-being of their elderly parents, the burden of caring for elderly parents has been well-documented. Potentially more important but much less documented is the opportunity cost of living near or with the parent and forgoing opportunities somewhere else. In a family with multiple adult children, altruism toward the elderly parent and the cost of caregiving and proximate living create a textbook public good problem. In addition, in the course of location decisions by siblings, there exists a potential strategic commitment device, the decision order. The eldest children may choose whether to move away from their parents once they finish their schooling, which is earlier than their younger siblings. Non-negligible relocation costs allow them to make a strategic commitment. Consistent with this sequential nature of the location decision of siblings, Konrad et al (2002) find adult children with younger siblings in Germany more likely to move farther away from their parents.

This paper studies externality and strategic interaction among adult siblings regarding their location decision relative to their elderly parents. We use the Health and Retirement Study (HRS) to study American families with a non-institutionalized elderly parent. We build on Konrad et al (2002) by estimating a sequential game played by adult siblings. The game-theoretic structural framework sheds light on (1) the degree of externality, such as altruism toward parents and cooperation among siblings, (2) the value of strategic commitment, or the first-mover advantage, (3) associated efficiency loss due to externality, and (4) how externality and inefficiency vary across families.

These empirical questions have significant policy implications. In the recent trend of population aging, elderly parents, particularly widowed mothers, are more likely to live alone for longer, often with disabilities. Despite the trend toward formal care, informal care still plays an important role. In the case of the elderly with a disability or severe medical condition, for example, around 80% of the hours of care are provided informally (OECD, 2005). Despite declining intergenerational coresidence and the increasing mobility of young generations, the majority of adult Americans still live within 25 miles of their mothers (Compton and Pollak, 2009). Though much of the informal care can be replaced by formal care, family assistance, such as companionship, attention, mental and emotional support, and frequent visits, contributes to the well-being of elderly parents and enables them to remain in the community (Matthews and Rosner, 1988). A good understanding of adult children's location decisions serves as an important step in designing public policies to promote the well-being of families in aging societies.

We estimate a sequential discrete game with perfect information, a simple yet robust framework to examine strategic interactions among siblings. Estimation relies on the maximum simulated likelihood in which the game is fully solved for an equilibrium outcome. After the preference parameters of children are recovered, counterfactual simulations reveal the structure of the location choice game and its implications for externality, strategic behavior, and inefficiency. Though our empirical framework is cross-sectional, we examine different waves of the HRS to ensure the robustness of results. We also estimate a joint utility maximization model as a test against the non-cooperative assumption.

There are myriad studies on informal care and living arrangements for families with elderly parents. We advance the literature in two ways. First, we are the first in the literature to develop an econometric model that captures the sequential aspect of decision making among siblings and to quantify its empirical importance. Existing studies that test the first-mover advantage take the nonstructural approach (Konrad et al, 2002; Holmlund et al, 2009), while all structural game-theoretic studies in this literature assume the simultaneous move of siblings. Second, we are the first to apply an empirical game to the living arrangement decision among siblings. Existing structural studies that concern strategic interaction among siblings focus on long-term care arrangements, taking siblings' locations as given.

Our focus on location choice rather than informal care per se is motivated as follows. First, living arrangements and location patterns are critical determinants of formal and informal care arrangements. Checkovich and Stern (2002) and Engers and Stern (2002) show that distance to parents is one of the strongest predictors of care provision by children. Most existing studies, however, rely upon cross-sectional care arrangement data without addressing the endogeneity between living arrangements and informal care intention.[1] Furthermore, since care arrangements may change over time, living arrangements may better capture the long-term commitment to looking after elderly parents. Second, strategic interactions and externality are most relevant in location choice, because of its discrete, irreversible, and long-term nature. For this reason, the application of the game-theoretic framework is fruitful and the associated efficiency analysis has a significant implication. Informal care arrangements and monetary transfers, on the other hand, are more negotiable and adjustable among siblings over time, which justifies the use of the bargaining approach in many existing structural studies (e.g. Pezzin and Schone, 1999; Engers and Stern, 2001).

The findings are summarized as follows. First, we find a positive externality and strategic interaction. The non-cooperative model fits better with data than the joint utility maximization model. Second, the value of sequential strategic moves, or the first-mover advantage, is found to be almost negligible compared with inefficiency in joint utility (Kaldor-Hicks inefficiency). Third, some families exhibit a Pareto improving but non-equilibrium location configuration (the prisoner's

[1] Hiedemann and Stern (1999) use instrumental variables to address this endogeneity.

dilemma), but the size of Pareto inefficiency is much smaller than the efficiency loss in joint utility. Fourth, the extent of externality and associated inefficiency varies across families. Externality is larger in a family with a widowed mother and with educated children who are younger (conditional on the parent's age). For these families, the efficiency loss is especially large when the widowed mother is educated and does not own a home and when children are married and similar to each other. Fifth, the impact of this public good problem is striking. Of the 2010 HRS families with multiple children, had families fully internalized externality and jointly maximized utility sum, 17% more parents would have had at least one child living nearby. The development of future research and public policies needs to take this non-negligible externality into account.

The paper is organized as follows. Section 2 explains the data and provides descriptive analysis to identify the key empirical patterns to be explained and to motivate our structural model. We review related literature in Section 3, and proceed to the econometric model in Section 4. A simulation analysis follows to illustrate the working of the model. After discussion of the estimation strategy in Section 5, the results are presented in Section 6. Section 7 concludes.

2 Data and Descriptive Results

2.1 Health and Retirement Study

Data are drawn from the Health and Retirement Study (HRS), a nationally representative longitudinal survey of Americans over 50. The HRS tracks the health, wealth, and well-being of these elderly individuals and their spouses. The HRS also asks the respondents about the demographics and location of all their children. Our main conclusions are based on the latest wave of the HRS conducted in 2010. Because our empirical framework is cross-sectional, we compare the results from the previous biannual waves, in particular the 1998 wave, to ensure the robustness of our results. The 1998 and 2010 waves are fairly distant from each other, and provide a good robustness check. See Appendix A for further discussion of the HRS.

2.2 Population

We study the cross-sectional living arrangement patterns of elderly parents and their adult children. Specifically, our sample consists of elderly individuals: (1) who do not live in a nursing home or institution; (2) who do not have a spouse younger than 50; (3) who have at least one surviving biological child; (4) who do not have more than 4 children; (5) who have no step or foster children; (6) whose youngest child is aged 30 or older and whose eldest child is younger than 65; (7) whose eldest child is at least 16 years younger than the parent (or the spouse, if the spouse is younger); and (8) who have no same age children.

In HRS 2010, about 3% of the elderly population live in nursing homes and fewer than 7% have no child. Hence our study covers the vast majority of elderly Americans. We limit the number of children to 4 to limit computational burdens. This group represents about 75% of parents who have a child. For our research question, we expect to learn little from adding very large families. Furthermore, families with many children may have family preferences considerably different from those of the majority.

We focus on relatively older children because the moves of younger children are often temporary; for example they may relate to schooling. The location decision of those above 30 is more likely to involve serious long-term commitment. We also limit the age of children to 65, because individuals around retirement age tend to be highly mobile. We focus our study on biological children, to avoid complications from potentially different family preferences in relation to non-biological children.[2] Finally, we exclude same age children[3] because sequential decision making is one of our main interests and because our estimation method utilizes the decision order.

For this population, we create a child-level data set. The spousal information is retained as explanatory variables. From the 15,372 respondents in HRS 2010, the sample selection creates our final data of 11,150 child observations in 4,619 families.

2.3 Location Patterns of Siblings

The location of the children relative to the parent defines our dependent variable. Conceptually, there are three distinctive categories of a child's location: (1) living with the parent; (2) living close to the parent; and (3) living far from the parent. In this study, we group the first two together and refer to this as living *near* the parent, to make our empirical framework tractable while maintaining our primary interest—a long-term commitment in providing (or not providing) care and attention to the parent. Though most existing studies focus on the primary caregiver (Hiedemann and Stern 1999; Engers and Stern, 2001) or coresidence (Pezzin and Schone, 1999), shared caregiving is observed in a non-negligible proportion of families (Matthews and Rosner 1988; Checkovich and Stern, 2002). Siblings living nearby can also contribute to the family by other means—by frequent visits and as a backup in case of caregiving burnout of the primary caregiver. Moreover, coresidence is becoming less common. These reasons motivate us to focus on living proximity, rather than coresidence. Distinguishing (1), (2), and (3) is an interesting but challenging task that we leave to future research.

[2] Because we do not know whether the relationship between the child and the respondent's spouse is biological, our restriction on non-biological children is approximate.

[3] In the HRS, we know the ages of children but not their birth dates. Hence, the "same age" children may or may not have the same birth dates.

Proximity is defined as a distance of less than 10 miles. This definition is used in HRS reports and previous studies (e.g. McGarry and Schoeni, 1995; Byrne et al, 2009).[4] Table 1 presents the location patterns of siblings based on our sample from HRS 2010. The top panel shows that 54.1% of only children live far from their parents. The second panel concerns two-child families, the most common case in the US. Elderly parents with two children are most likely to have no child nearby (40.4%) and least likely to have both of them nearby (19.9%). This implies that the probability that each child lives near the parent is lower than that of only children. At the same time, the four panels show that the chance of having at least one child nearby increases with the number of children; while the majority of parents with one child live without a child nearby (54.1%), parents of four children are much less likely to live without any child nearby (22.4%). Similar patterns are reported in Checkovich and Stern (2002).

[Insert Table 1: Sibling Location Patterns by Birth Order: 2010]

Conditional on one child living near the parent, two-child families have two possible location patterns: (near, far) and (far, near). The fact that the former is less likely than the latter is consistent with the first-mover advantage of the eldest child. Although this difference is quite small, the three- and four-child family panels prove the robustness of this asymmetry among siblings, across the number of siblings and across the number of siblings living near the parent. For example, 35.9% of three-child families have one child near the parent. Of these families, 14.8% have the last child nearby, while 10.5% and 10.6% have the first and second child nearby, respectively.

Figure 1 visualizes these patterns and also shows the location patterns in 1998. Compared with 1998, parents in 2010 are more likely to have no child living nearby, regardless of the number of

[4] While many children never leave their parents, elderly parents sometimes relocate closer to their children. We do not distinguish how the proximate living is formed.

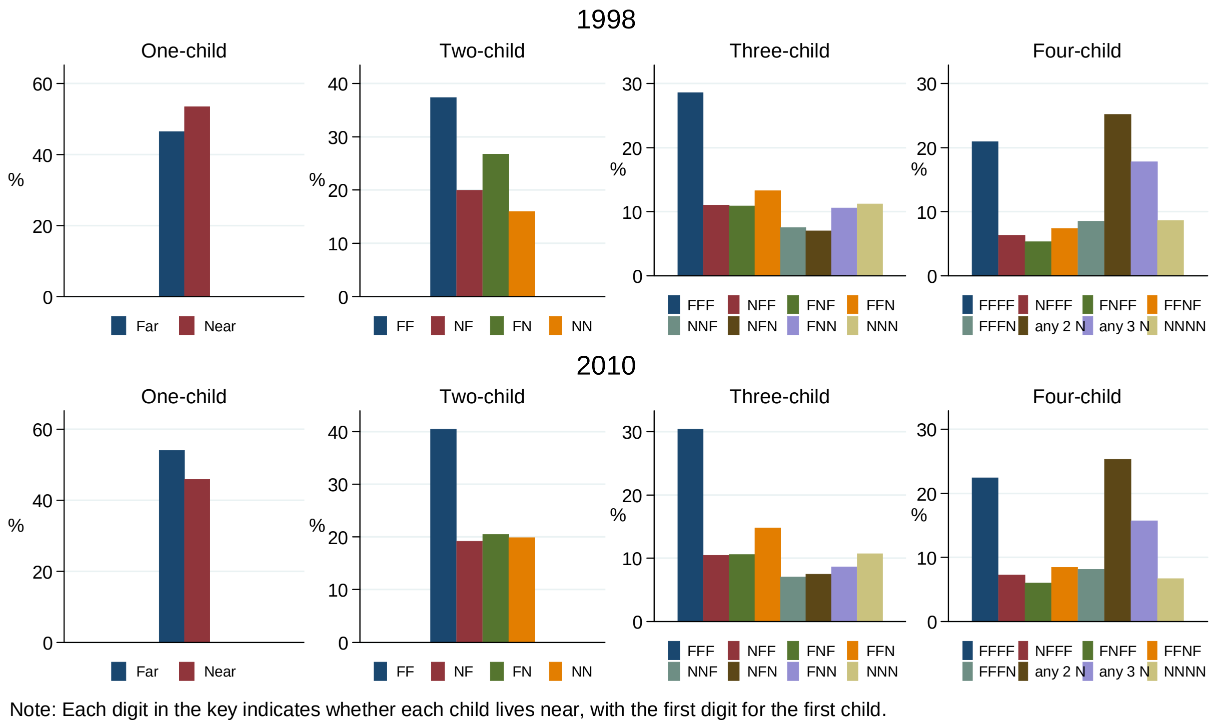

Figure 1: Observed Location Patterns: 1998 and 2010

children, which probably reflects increasing mobility and declining intergenerational coresidence. Despite this time trend, Figure 1 consistently shows that elder children are more likely to live far from their parents in both 1998 and 2010. The youngest child, particularly, is likely to live near the parent.

The robust birth order asymmetry is in line with Konrad et al's (2002) argument of the first-mover advantage. This may simply reflect the systematic difference between elder and younger siblings, however. Elder children are in the later stage of their life, and they are more likely to have better outside options and greater commitment to their own family. It is also a well-known fact that elder children tend to have more education.[5] Hence, how much of the observed birth

[5] In our sample, the share of those who have a university degree is 39.1% and 36.8% for the first and second children in two-child families, and 36.8%, 32.0%, and 31.7% for the first, second, and third children in three-child families.

order asymmetry is attributed to the first-mover advantage is an empirical question. To answer this, we include observable family and child characteristics in our analysis and then proceed to the game-theoretic model.

2.4 Explanatory Variables

We use both parents' and children's characteristics. Parental characteristics include demographics (age, sex, marital status, and ethnicity), education, health conditions, residential location, and housing. For children, we use demographics (age, sex, education, and marital status) and information on grandchildren.

Table 2 provides the definitions of the explanatory variables and their summary statistics. The majority of the parents in the sample are single, with single mothers being the most common. Parental health conditions are poor with a high number reporting poor eyesight and limitations in performing daily tasks, known as Activities of Daily Living (ADL), which consist of dressing, walking, bathing, eating, getting in and out of bed, and using the toilet. Most children are in their 40s and have a spouse and children.

[Insert Table 2: Definition and Summary Statistics of Explanatory Variables]

2.5 Probit Results

Before presenting the game-theoretic model, it is useful to summarize robust key facts that emerge from a simple probit analysis as the statistical associations between the location pattern and observable family characteristics. Table 3 reports the result for all waves from 1998 to 2010. The result is fairly robust across waves, indicating the similarity of data across waves. The result is overall consistent with existing studies. Parents who have a child living nearby tend to be less educated single mothers over 80, live in urban areas, and have their own home. Parents' disability and existing conditions also affect intergenerational proximity. While parents with a stroke and poor eyesight are more likely to have a child nearby, proximate living with children is less likely when the parent has ADL limitations. These results are not straightforward to interpret, particularly because parents' condition variables capture both parental needs and caregiving costs to children, and hence in this paper, we do not attempt to draw conclusions from the results of health related variables. Previous US studies have also found negative associations of ADL with informal care provision (for further discussion, see Byrne et al, 2009).

[Insert Table 3: Preliminary Probit Regressions]

Child variables are also relevant. Unmarried children, particularly single daughters, are likely to live near their parents. Highly educated children are less likely to live near parents. These findings are consistent with previous studies (e.g. Checkovich and Stern, 2002; Compton and Pollak, 2009; Byrne et al, 2009). Intergenerational proximity is less likely for children who are older and have fewer grandchildren.

The probit analysis also serves as a benchmark for more complex models; the model offers a simple random utility model interpretation under the assumptions that each child makes their location decision independently, their decision has no implications for other children, and their unobserved taste is distributed i.i.d. normal. These assumptions are restrictive: in multi-child families, family specific heterogeneity and within-sibling interactions may well play significant roles. Most existing structural studies have shown that the decision making of family members is not independent (e.g. Checkovich and Stern, 2002). To address the within-family interdependence, we add more structures to the above simple probit model.

3 Related Literature

There are a few structural studies on the family's decision on living arrangements, but none of them address interactions among siblings. Pezzin and Schone (1999) study American families with one daughter using a bargaining model of coresidence, care arrangements, and the child's labor force participation. Sakudo (2008) studies Japanese families with one daughter by a bargaining model of coresidence, monetary transfers, and marriage. Hoerger et al (1996) study living arrangements, allowing multiple children to contribute to caregiving, based on a single family utility function.

A small but tangible body of literature applies the non-cooperative game-theoretic framework to study interactions among siblings in family decisions regarding informal care arrangements (Hiedemann and Stern, 1999; Checkovich and Stern, 2002; Engers and Stern, 2002; Byrne et al, 2009; Knoef and Kooreman, 2011). In these models, each family member acts to maximize his own utility. Informal care contributes to the well-being of parents, from which siblings derive utility. At the same time, informal care is costly to provide, and care provided by another sibling is a substitute. This creates a free-rider problem. Though the game structure and solution concepts in these studies vary, the equilibrium arrangement is fully solved for in estimation. Hiedemann and Stern (1999) and Engers and Stern (2002) study the family decision about the primary caregiver; Checkovich and Stern (2002) study the amount of care, allowing for multiple caregivers. Byrne et al (2009) enrich these studies by also modeling consumption, transfers for formal home care, and labor supply. While these studies use US data, Knoef and Kooreman (2011) estimate a model that is as rich as Byrne et al's (2009) with European multi-country data, the Survey of Health, Ageing, and Retirement in Europe (SHARE). Except for Byrne et al (2009), these studies find support for interdependence in caregiving decisions among siblings. Knoef and Kooreman (2011) find that if siblings engage in joint utility maximization, 50% more informal care will be provided to parents, while the costs to children will increase to a much smaller extent. All these structural studies apply the game-theoretic framework to explain across-family variations in care arrangements, taking families' location choice as given. We contribute to this empirical literature by applying the game-theoretic framework to the living arrangement decision among siblings for the first time.[6]

Given the complexity of care and living arrangements, one model cannot capture all possible aspects of family decision making. The existing structural studies on informal care utilize rich and reliable measures of informal care and other transfers, endogenize labor force participation and formal care decisions, and/or incorporate important policy variables, such as eligibility for Medicaid. We abstract these relevant features to concentrate on modeling strategic interaction and externality. Therefore, our study should be regarded as a complement to existing structural studies.

Our study builds on the nonstructural study by Konrad et al (2002). They estimate an ordered logit model of children's distance from the parent with child-level data of two-child families in the mid-1990s drawn from the German Aging Survey, and find first-born children more likely to live farther from their parents than their younger siblings. They argue that this finding supports their first-mover advantage hypothesis: by locating sufficiently far from the parent, the first-born child can force his younger sibling to locate closer to the parent as the primary caregiver. Holmlund et al (2009) re-examine the conclusion of Konrad et al (2002), using 9,000 individuals in 9 countries (including Germany) drawn from SHARE.[7] Holmlund et al (2009) find no significant asymmetry in the location decisions and time transfers to elderly parents between firstborn and second-born children. These nonstructural studies, however, do not account for interdependence in the deci-

[6] For more details of the family decision literature of informal care, see the literature section in Byrnes et al (2009).
[7] They also argue that the theoretical results of Konrad et al (2002) are not robust to certain generalizations.

sions and unobserved heterogeneity in the preferences of siblings, and thus may reach misleading conclusions.

Though we do not discuss them here due to space limitations, there are myriad other economics and non-economics studies on the living arrangements (Börsch-Supan et al, 1988; Dostie and Léger, 2005; Hank, 2006; Compton and Pollak, 2009; Fontaine et al, 2009; Hotz et al, 2010; Johar and Maruyama, 2011) and on the role of gender and family composition in filial informal care (Matthews and Rosner, 1988; Holroyd, 2001; Lin et al, 2003; Hequembourg and Brailler, 2005; Silverstein et al, 2006).

4 The Model

4.1 Setup

We consider a game played by children. Our goal is to describe the observed cross-sectional snapshot of living arrangement patterns of families by explicitly modeling strategic interactions among siblings. Each child chooses whether to live close to their elderly parent(s). We do not distinguish coresidence and living nearby; for the rest of the paper, living "near" includes living together. Let $a_{i,h} \in \{0,1\}$ denote the *action* of child $i = 1, ..., I_h$ in family $h = 1, ..., H$. If child i lives near the parent, $a_{i,h} = 1$. Child $i = 1$ denotes the eldest child.

We model the location choice of children as a one-shot perfect information sequential game. This approach has several implications. First, we formulate the location problem of families solely as the children's problem, not modeling the role of parents. This simplification helps us to focus on interaction among siblings. In reality, parents may be involved in the family's location decision. In our view, however, this decision is essentially made by children: parents are unlikely to be able to force their child to live nearby. In addition, our model can be regarded as a reduced-form with implicit family bargaining and intergenerational transfers.[8]

Second, the one-shot sequential game ignores the dynamic or repeated aspects of location choice over the family's life-cycle. As discussed in Konrad et al (2002), this simplification is somewhat justified by large relocation costs. Nevertheless, after children enter the labor market and establish own new family, numerous events, such as changes in the family structure, parents' retirement, and parents' health deterioration, may occasionally influence the location decision. This type of dynamics is beyond the scope of this study, as it has been for most previous studies.[9][10]

Third, we rely on the non-cooperative game theoretic framework. Bargaining, negotiations, and side-payment transfers are abstracted and implicitly captured in the payoff function. To address this restriction, we also estimate a model of joint utility maximization and test it against our non-cooperative framework. Fourth, we assume a game with perfect information. Although the majority of empirical games in the industrial organization and labor literature assume incomplete information, in the family setting, the perfect information framework is reasonable, because family members know each other well.[11][12]

[8] Checkovich and Stern (2002) and Knoef and Kooreman (2011) employ the same approach.

[9] There is a line of research on the dynamics of living arrangements (e.g. Börsch-Supan et al, 1988; Dostie and Léger, 2005; Johar and Maruyama, 2011). Interaction among siblings, however, is beyond the scope of this literature.

[10] The structural econometric modeling of entry in the industrial organization literature also started with the static game-theoretic framework to analyze a cross-sectional snapshot of market structures (e.g. Berry, 1992), even though the entry decision of firms is inherently dynamic.

[11] The informal care literature uses both approaches; e.g. Byrne et al (2009) assume a complete information game, while Engers and Stern (2002) assume a game with private information.

[12] Related to this, Engers and Stern (2002) consider both pure and mixed strategies. In our view, mixed strategies are not relevant to the location decision.

4.2 Preferences

4.2.1 Children's Problem

First suppose that child i has no sibling. Denote the utility level of child i by $u_{i,h}(a_{i,h})$. In the rest of the paper, the family subscript, h, is omitted when no ambiguity arises. Child i's problem is, after dropping subscript h, written as

$$\max_{a_i \in \{0,1\}} u_i(a_i).$$

This model can be analyzed as a standard random utility binary choice model. To extend this model to multi-child families, we need to incorporate interdependence among the decisions of siblings. Child i's utility depends not only on his choice, a_i, but also the choices of his siblings, $a_{-i} \in \{0,1\}^{I-1}$. Denote the utility level of child i by $u_i(a_i, a_{-i})$. Given the decisions of his siblings, the child's problem is written as

$$\max_{a_i \in \{0,1\}} u_i(a_i, a_{-i}).$$

4.2.2 Specification of Preferences

To make our functional specification reasonably parsimonious, we assume that child i's utility depends only on his own location choice, a_i, and the number of his siblings who choose to live near the parent. Let $N = \sum_k a_k$ denote the number of siblings who choose to live near the parent, and $N_{-i} = \sum_{k \neq i} a_k$ the number of siblings who choose to live near their parent except for child i. The utility levels when child i lives far from the parent and near the parent are specified as follows:

$$\begin{cases} u_i(a_i = 0, a_{-i}) = u_i^{\alpha}(N), \\ u_i(a_i = 1, a_{-i}) = u_i^{\alpha}(N) + u_i^{\beta} + u_i^{\gamma}(N_{-i}). \end{cases} \tag{1}$$

Utility flow consists of three components, $u_i^\alpha(N), u_i^\beta$, and $u_i^\gamma(N_{-i})$. The first component, $u_i^\alpha(N)$, captures the child's *altruism* toward the parent. It is a utility gain of child i from the parent's well-being (such as happiness, good health, and long-term security) that arises if the parent is taken care of by *any* child. We assume $u_i^\alpha(0) = 0$, that is, we normalize the system without loss of generality so that when every sibling lives far from the parent, everyone receives zero utility. If $u_i^\alpha(N > 0)$ is positive, the parent's well-being (or more precisely, intergenerational proximate living) is a public good with a positive externality. This term may be an increasing function of N if the number of children living nearby means a greater amount of care and attention given to the parent, and the child is concerned about the amount of care and attention.

The next component, u_i^β, captures child i's personal utility costs (or benefits) from living near the parent that are independent of a_{-i}, the decisions of his siblings. For example, this term includes caregiving burdens, opportunity costs (outside options somewhere else), moving costs, frequent contact and close companionship with the parent, monetary transfers to/from parents, housing benefits in case of coresidence, and attachment to parents and their location.

The third component, $u_i^\gamma(N_{-i})$, is child i's personal utility costs or benefits from living near the parent that depend on a_{-i}. This *cooperation* term is likely to be a positive function of other siblings' proximity. Siblings can share the costs of looking after parents. Siblings may also enjoy living close to each other. This term can also be a decreasing function of N_{-i}. It might be costly to get along with siblings' families and coordinate the care and attention given to the parent. Another example is the bequest motive hypothesis discussed in Bernheim et al (1985)—the presence of another sibling taking care of the parent reduces transfers from the parent.

In our general model, these three terms are specified as follows:

$$u_i^{\alpha}(N) = X_i^{\alpha}\alpha^0 \cdot \left(I[N \geq 1] + \alpha^1 \cdot I[N \geq 2] + \alpha^2 \cdot I[N \geq 3]\right), \tag{2}$$

$$u_i^{\beta} = X_i^{\beta}\beta, \quad \text{and}$$

$$u_i^{\gamma}(N_{-i}) = X_i^{\gamma}\gamma^0 \cdot \left(I[N_{-i} \geq 1] + \gamma^1 \cdot I[N_{-1} \geq 2]\right).$$

The relative importance of altruism and cooperation terms $(u_i^{\alpha}, u_i^{\gamma})$ are preference parameters and are likely to vary across families. We allow preference heterogeneity based on observables. In this specification, (α, β, γ) is a set of parameters common to every child, and $\mathbf{X} \equiv \left(X_i^{\alpha}, X_i^{\beta}, X_i^{\gamma}\right)$ is a vector of covariates observable to the econometrician that includes a constant term.

4.3 Equilibrium and Efficiency Benchmarks

The location decision is made by siblings in their birth order. All siblings' preferences and the game structure are known to every sibling. In this sequential game, child i's *strategy*, $s_i \in S_i$, specifies the child's decision at *every decision node* (thus note the difference between a_i and s_i). A subgame perfect pure strategy Nash equilibrium (SPNE) is obtained when no child expects to gain from individually deviating from their equilibrium strategy in *every subgame.* Every finite game with perfect information has a pure strategy SPNE (Zermelo's theorem).

The sequential nature of the game is illustrated in the extensive form representation in Figure 2. The figure shows four possible SPNE when the first child chooses to live nearby. Because the younger child has two decision nodes, his choice set comprises four strategies, which we refer to as "always far", "imitate", "preempted", and "always near". Note that if the elder child lives nearby, the first and third strategies, "always far" and "preempted", lead to the same living arrangement

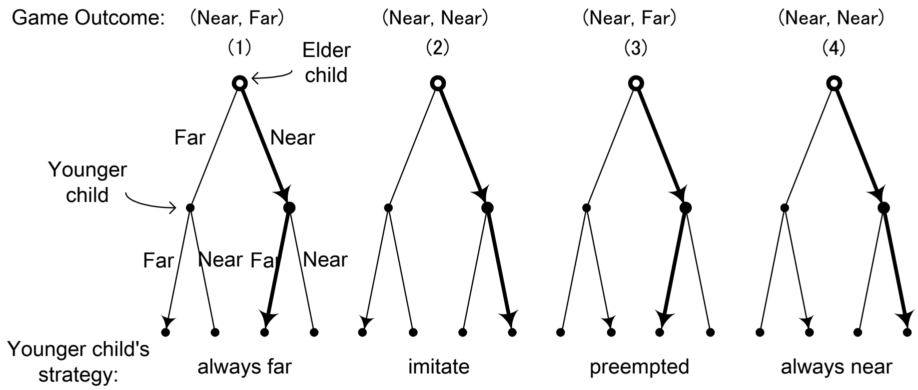

Figure 2: Strategies and Outcomes in Extensive Form Presentation

outcome, (Near, Far). In estimation, we exploit this one-to-many mapping structure of SPNE (see Section 5.2).

To examine how desirable an equilibrium outcome is, we use two efficiency measures: Pareto efficiency and efficiency in joint utility. Though the game has a unique SPNE, it may have a Pareto-improving (non-equilibrium) outcome, which constitutes a well-known prisoner's dilemma. Efficiency in joint utility, or Kaldor-Hicks efficiency, concerns the utility sum of siblings. This assumes the additivity of utility among siblings, which is implicitly imposed in our econometric model. Though it does not guarantee Pareto improvement, this efficiency measure is sensible to examine when families or policy makers consider implementable compensation schemes. The following examples in the normal form illustrate the relationship between these concepts:

Example 1:

	$a_2 = 1$	$a_2 = 0$
$a_1 = 1$	$(2, 2)$	$(-1, 1)$
$a_1 = 0$	$(1, -1)$	$(0, 0)$

Example 2:

	$a_2 = 1$	$a_2 = 0$
$a_1 = 1$	$(1, 1)$	$(-1, 2)$
$a_1 = 0$	$(2, -1)$	$(0, 0)$

Example 3:

	$a_2 = 1$	$a_2 = 0$
$a_1 = 1$	$(-1, -1)$	$(-2, 4)$
$a_1 = 0$	$(4, -2)$	$(0, 0)$

Without the sequential structure, Example 1 has two (simultaneous move) Nash equilibria, (Near, Near) and (Far, Far). The former is Pareto dominating and the latter is so-called coordination failure. Once we introduce the decision order, (Near, Near) becomes the only SPNE outcome. Example 2 exhibits the prisoner's dilemma. A Pareto improving (Near, Near) is not an equilibrium and thus causes Pareto inefficiency.[13] The unique equilibrium in Example 3, (Far, Far), is Pareto efficient but not joint-utility efficient. The family can achieve a larger joint utility at (Near, Far) or (Far, Near)— at the expense of either sibling's compromise. If appropriate compensation is possible, these efficient outcomes are preferred. Note that assuming a constant altruism $\left(\alpha^1 = \alpha^2 = 0\right)$, the payoff matrix in Example 1 implies: $\left(u_i^a, u_i^\beta, u_i^\gamma\right) = (1, -2, 3)$. Similarly, $\left(u_i^a, u_i^\beta, u_i^\gamma\right) = (2, -3, 2)$ in Example 2 and $(4, -6, 1)$ in Example 3. Thus, the degrees of altruism, private costs, and cooperation govern the game structure in each family.

4.4 Unobserved Error Term

To match the model with data, we need an unobserved error term. We assume that the error additively affects the private utility of living near the parent. Formally,

$$\begin{cases} u_i\left(a_i = 0, a_{-i}\right) = u_i^\alpha\left(N\right), \\ u_i\left(a_i = 1, a_{-i}\right) \ = u_i^\alpha\left(N\right) + u_i^\beta + u_i^\gamma\left(N_{-i}\right) + \varepsilon_i. \end{cases} \tag{3}$$

The unobserved error term is assumed to be independent of $\mathbf{X}_h$, and distributed in normal distribution. Under the assumption of perfect information, $\boldsymbol{\varepsilon}_h$ is unobservable to researchers but is observed by the family. The normality assumption implies that the game almost surely has a unique equilibrium, because ties occur with probability measure zero. We can solve for the unique equilibrium

[13] Pezzin et al (2007) discuss the prisoner's dilemma in a closely related context. They develop a two-stage bargaining model to determine the equilibrium resource allocation to disabled elderly parents. The first-stage decision is on coresidence and the second-stage determines how much assistance to provide. They argue that if no family member can make binding arrangements in the first-stage of the game, coresidence may fail to emerge, even if it is Pareto efficient, because coresidence reduces the bargaining power of the coresidence child relative to his siblings.

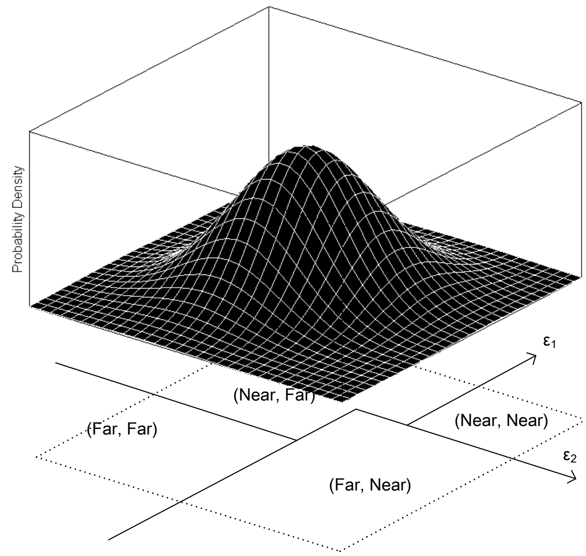

Figure 3: Relationship between Error Terms and Location Outcome

solution by backward induction for any given parameters, (α, β, γ), observed characteristics of the family and siblings, $\mathbf{X}$, and unobservable heterogeneity, $\boldsymbol{\varepsilon}$.

In other words, conditional on $\mathbf{X}$ and (α, β, γ), the location configuration is determined by $\boldsymbol{\varepsilon}_h$. Figure 3 depicts this relationship in a simple, two-child family example. A higher value of ε_i increases the propensity that child i lives near the parent. The correlation of ε_1 and ε_2 is captured by the probability density function. The asymmetry around the center part in Figure 3 is due to sequential strategic interaction. When the gain from living near the parent is modest for both children, the first child can take advantage of the decision order. The first child's commitment to the irreversible decision to move away allows him to free-ride on the second child.

As with the standard random utility models, the level of utility is not identified. We normalize the variance of $\varepsilon_{i,h}$ to one. Formally,

$$\boldsymbol{\varepsilon}_h \equiv \{\varepsilon_{i,h}\}_{i=1,\ldots,I_h} \sim \Phi\left(\Omega^h\right), \tag{4}$$

where Ω^h is the $I_h \times I_h$ covariance matrix whose diagonal elements are unity and whose (i, j) off-diagonal element is $\rho_{i,j} \in (-1, 1)$, a correlation coefficient of $\varepsilon_{i,h}$ and $\varepsilon_{j,h}$. The off-diagonal element can be parameterized as

$$\rho_{i,j} = X^{\rho}_{i,j}\theta^{\rho}, \tag{5}$$

where θ^{ρ} are vectors of parameters and $X^{\rho}_{i,j}$ is a set of relational variables between child i and child j, such as their age and gender differences.

4.5 Simulation: Understanding the Mechanics

Before we proceed to estimation, it is worthwhile to further investigate the quantitative nature of externality and strategic interaction under the given model. To highlight the effect of the values of u^{α}, u^{β}, and u^{γ} on the game outcome, we simulate the game of "symmetric" two-child families, in which the two children have the same values of u^{α}, u^{β}, and u^{γ}. The two children are effectively identical except for birth order and the error term, $\boldsymbol{\varepsilon}$. We also assume that u^{α}, u^{β}, and u^{γ} are constant in the number of siblings living near the parent, i.e. $\alpha^1 = \alpha^2 = \gamma^1 = 0$.

Each panel in Figure 4 shows how the utility levels of the two children change along u^{β}, under various values of u^{α} and u^{γ}, when there is no error term ($\boldsymbol{\varepsilon}_h = 0$). The three panels in the first column, for example, show game outcomes when $u^{\alpha} = 0$. The middle panel in the first column shows that when $u^{\alpha} = u^{\gamma} = 0$, the two utility curves completely coincide. In this case, each child's decision has no externality, and thus the decision order creates no difference. As long as $u^{\beta} < 0$, neither child lives near the parent and $u_i = 0$, but once u^{β} exceeds zero, each child chooses to live near the parent and $u_i = u^{\beta}$. A positive u^{α} creates a first-mover advantage. Although a positive u^{α} increases both children's utility, when the first child has modestly small u^{β}, he can free-ride

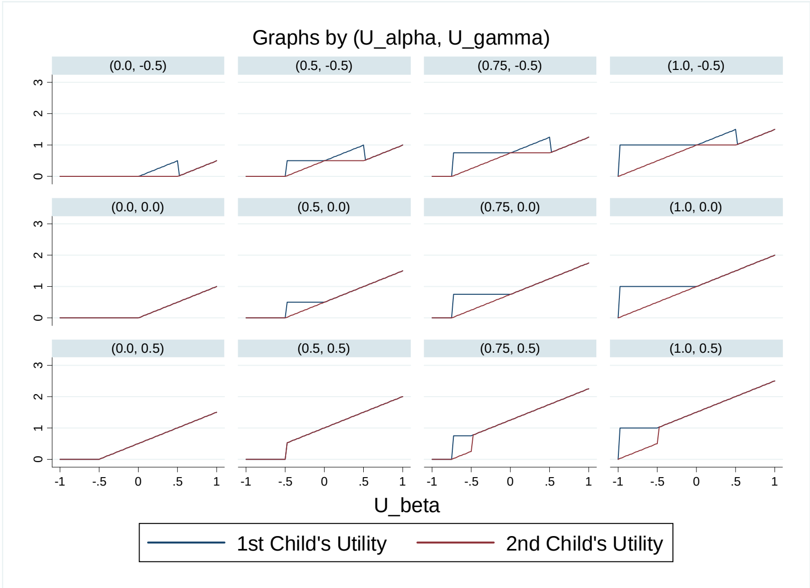

Figure 4: First-Mover Advantage When No Error Term: 2-Child Families

on the second child by committing to "away". Since u^{α} reflects the magnitude of externality, the larger u^{α}, the larger the first-mover advantage.

The other externality term, u^{γ}, also affects the first-mover advantage. In general, the size of a first-mover advantage depends on strategic substitutability. Gal-Or (1985) studies a two-player Stackelberg game and proves that when the reaction functions of the players are downwards (upwards) sloping, the first mover earns higher (lower) profits. The same principle applies here. Observe that

$$u_1\left(a_1=1,a_2=1\right)=u_1^{\alpha}+u_1^{\beta}+u_1^{\gamma},\quad u_1\left(a_1=1,a_2=0\right)=u_1^{\alpha}+u_1^{\beta},$$
$$u_1\left(a_1=0,a_2=1\right)=u_1^{\alpha},\quad u_1\left(a_1=0,a_2=0\right)=0.$$

To see the strategic substitutability in our discrete setup, define the following

$$\left[u_1\left(1,1\right)-u_1\left(0,1\right)\right]-\left[u_1\left(1,0\right)-u_1\left(0,0\right)\right]=-u_1^{\alpha}+u_1^{\gamma}.$$

Analogous to Gal-Or's (1985) argument, when this is negative, i.e. when the objective function exhibits *decreasing difference*, we expect a larger first-mover advantage. Figure 4 clearly shows this effect. A smaller u^γ and larger u^α (in the northeast panel with (1.0,−0.5)) most widen the gap between the two utility curves.

When u^α is small and/or u^γ is large, on the other hand, siblings' decisions may become strategic complements (a supermoduler game). In our simple binary symmetric setup, however, the second-mover advantage never appears, as seen in the bottom panel in the first column. This is because strategic complementarity degenerates the game into the choice between $(1,1)$ and $(0,0)$ and at the same time, the first mover is never worse off.

Now we incorporate the error component to understand the quantitative working of the econometric model. The variance of $\boldsymbol{\varepsilon}$ is one. We set the covariance between ε_1 and ε_2 to 0.5, which is close to the estimate we obtain below. Due to this bivariate normal distribution and the discrete nature of the game, it is infeasible to derive analytical solutions. Hence, we use the Monte Carlo simulation.

Figure 5 illustrates the expected utility of siblings for different values of $\left(u^\alpha, u^\beta, u^\gamma\right)$. Expected utility increases in u^α, u^β, and u^γ. The first-mover advantage, as the difference between the two children's expected utility curves, exists when u^α is large. Figure 6 magnifies this utility gap, and shows that the first-mover advantage becomes larger when u^α is larger and u^γ is smaller (the largest in the northeast panel). The first-mover advantage decreases with u^β, because with a large u^β, the elder sibling would rather live near the parent than free-ride on the younger sibling. As discussed earlier, we observe no second-mover advantage. Though not reported, we have also simulated larger families. In three-child families, the utility gap between the first and third children is almost the same as the gap in two-child families. The middle child experiences a tiny disadvantage and can

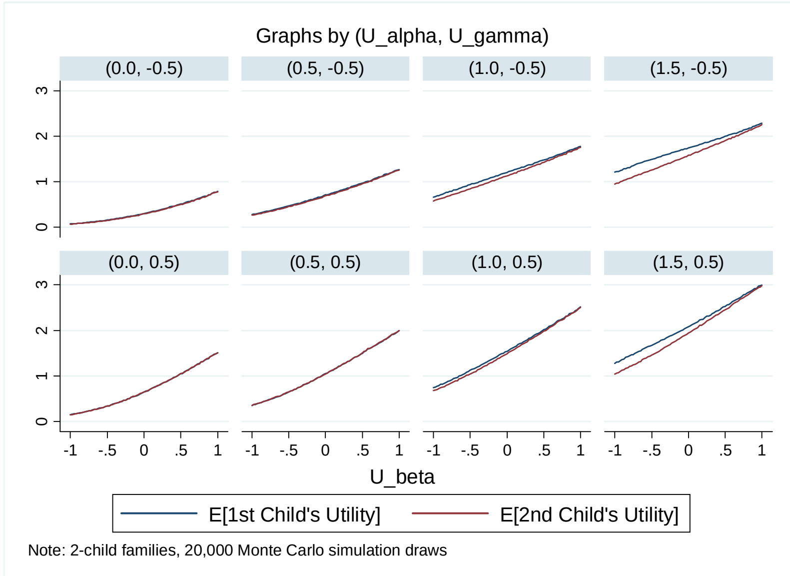

Figure 5: Expected Utility Levels of Siblings

free-ride on the third child almost as much as the eldest child does.

Externality leads to inefficiency. We can confirm this by simulating two kinds of optimal outcomes and comparing the SPNE outcome with them. First, consider a location outcome that maximizes the sum of two children's utility. Figure 7 shows the difference in the joint family utility between the joint-utility-maximizing outcome and SPNE outcome. When there is no externality ($u^\alpha = u^\gamma = 0$), the SPNE outcome maximizes the joint profit (the second panel in the first column). In contrast with the previous figures, the size of the joint-utility inefficiency increases with the *absolute* size of u^α and u^γ. Both positive and negative values of u^γ enlarge inefficiency. When u^α is positive, efficiency loss decreases in u^β, because with larger values of u^β, children are more likely to choose to live near the parent, which alleviates the free-rider problem. Second, Figure 7 also shows the Pareto efficient outcome, which maximizes the utility sum of two children without reducing either child's utility. This "prisoner's dilemma" case only appears when $u^\alpha > 0$ and $u^\gamma > 0$, i.e. when cooperation increases payoffs but the incentive to free-ride exists. The magnitude

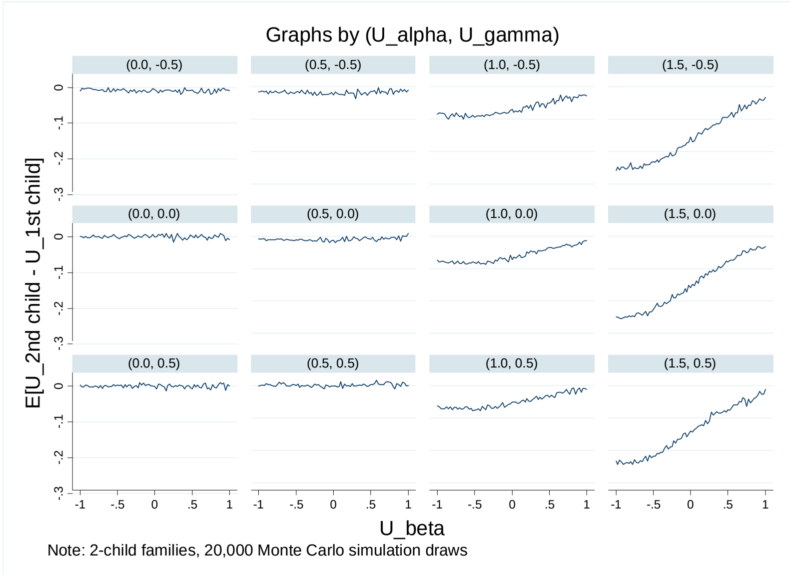

Figure 6: Second-Mover Disadvantage

of joint-utility inefficiency is much larger than that of Pareto inefficiency, if it exists.

Lastly, we also investigate the effect of ρ, the correlation parameter. While the above findings are fairly robust with respect to the value of ρ, a higher value of ρ slightly increases the first-mover advantage when it exists. This is because a high correlation of the error term more frequently leads to a conflict of interests and generates greater room for strategic behavior.

5 Estimation

5.1 Method of Simulated Likelihood

The estimation relies on the maximum likelihood estimation in which the game is fully solved for an equilibrium outcome, a_h^*. Denote the observed family location configuration as $a_h^o \in \{0,1\}^{I_h}$.

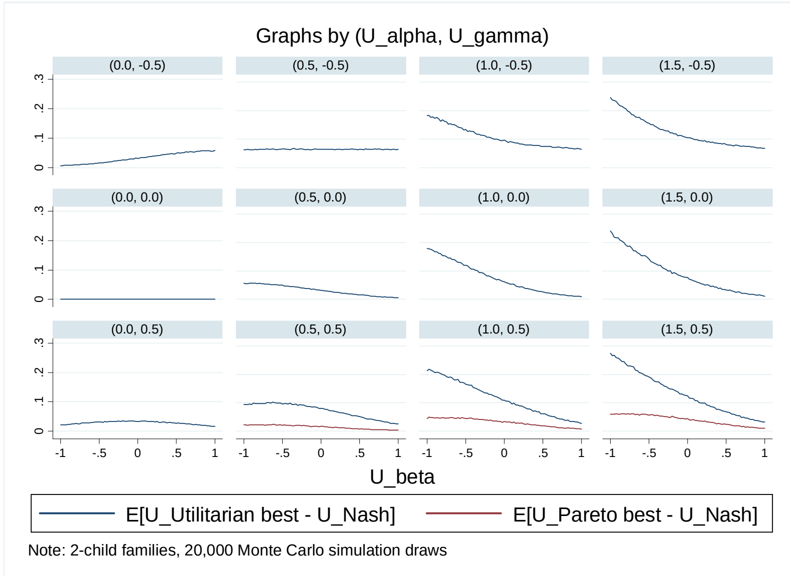

Figure 7: Two Measures of Family Inefficiency

The log-likelihood function is written as

$$\widehat{\theta}_{ML} = \arg\max_{\theta} \left\{ \frac{1}{H} \sum_{h}^{H} \ln \Pr_{\rho} \left[a_h^o = a_h^*(\mathbf{X}_h, \boldsymbol{\varepsilon}_h;\ \alpha, \beta, \gamma) \right] \right\}, \tag{6}$$

where θ is the vector of model parameters, $(\alpha, \beta, \gamma, \rho)$.

The probability in the likelihood does not have an analytical solution due to multidimensional integrals over the $\boldsymbol{\varepsilon}_h$ space. This motivates the use of the maximum simulated likelihood (MSL) method. The multidimensionality becomes a non-trivial problem when error components are correlated among siblings and the decisions of siblings are interdependent. When the dimension of $\boldsymbol{\varepsilon}_h$ becomes large (i.e. more than two or three), computationally demanding numerical approximation, such as a quadrature method, is impractical.

To overcome this computation problem, we use the Monte Carlo integration method developed by Maruyama (2010). This simulation-based integration method utilizes the structure of the perfect information sequential game, and characterizes an observed equilibrium outcome into every subgame perfect equilibrium that rationalizes the observed outcome and then applies the Geweke-Hajivassiliou-Keane (GHK) simulator, a widely used probit simulator, to each equilibrium.

We turn now to a brief overview of the MSL estimation, followed by identification issues. Readers not interested in these details can turn directly to the empirical results.

5.2 Monte Carlo Integration

Maruyama (2010) develops the estimation method applicable to discrete-choice sequential games with perfect information, in which each player makes a decision in publicly known exogenous decision order. The proposed method relies on two ideas. Its first building block is the GHK simulator, the most popular solution to approximate high-dimensional truncated integrals in standard probit models. This powerful importance-sampling simulator recursively truncates the multivariate normal probability density function, by decomposing the multivariate normal distribution into a set of univariate normal distribution using Cholesky triangularization.

Strategic interaction, however, complicates high-dimensional truncated integration, causing interdependence of the truncation thresholds, which undermines the ground of the GHK's recursive conditioning approach. The second building block of the proposed method is the use of the GHK simulator, not for the observed equilibrium outcome per se, but separately for each of the subgame perfect strategy profiles that rationalize the observed equilibrium outcome. In the sequential game framework, the econometrician does not observe the underlying subgame perfect equilibrium. This is because an equilibrium strategy consists of a complete contingent plan, which includes off-the-equilibrium-path strategies as unobserved counterfactuals. From the econometrician's viewpoint, there may exist different realizations of unobservables that lead to different subgame perfect equilibria but generate an observationally equivalent game outcome. Figure 8 visualizes this point. The integration domain of $(\varepsilon_1, \varepsilon_2)$ that leads to the location outcome, (Near, Far), is not rectangular due to strategic interaction between the two children, and hence the standard GHK simulator breaks down for this domain.

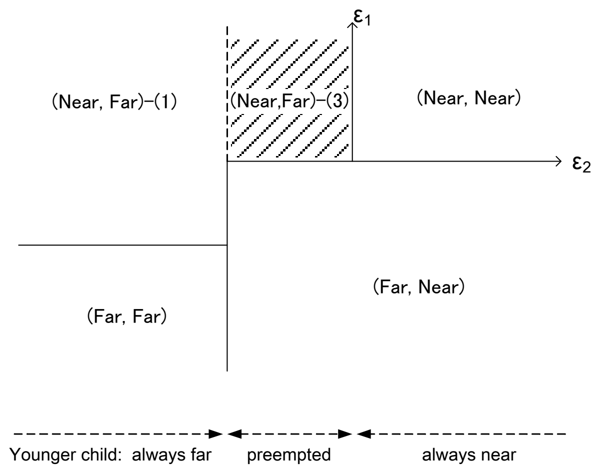

Figure 8: Dividing Observed Location Outcome into Strategy Profiles

The use of subgame perfection resolves this non-rectangular domain problem. In the example in Figure 8, the non-rectangular integration domain for (Near, Far) consists of two rectangular regions that correspond to two sets of SPNE, as labeled (1) and (3), which also correspond to (1) and (3) in the extensive form in Figure 2. Maruyama (2010) proves that the separate evaluation of likelihood contribution for each subgame perfect strategy profile allows us to control for the unobserved off-the-equilibrium-path strategies so that the recursive conditioning of the GHK simulator works by making the domain of Monte Carlo integration (hyper-)rectangular. The econometrician then obtains the probability of the observed outcome by summing the probabilities of each SPNE that rationalizes the observed outcome, and the use of maximum likelihood follows. For more details about the theoretical foundation and performance of this method, see Maruyama (2010).

5.3 Identification

An informal discussion suffices to show the identification of $\left\{u^{\alpha}(N), u^{\beta}, u^{\gamma}(N_{-i})\right\}$. First consider the choice problem of the only child. The utility values the only child receives when he chooses "far" and "near" are 0 and $u_1^{\alpha}(1) + u_1^{\beta}$, respectively. Analyzing this choice problem with only-child data identifies $u_1^{\alpha}(1) + u_1^{\beta}$. Next, consider the choice problem of the second child in a two-child family. He makes a decision after observing the first child's location decision. When the first child lives near the parent, comparing $u_2^{\alpha}(1)$ and $u_2^{\alpha}(2) + u_2^{\beta} + u_2^{\gamma}(1)$ identifies $u_2^{\alpha}(2) - u_2^{\alpha}(1) + u_2^{\beta} + u_2^{\gamma}(1)$. When the first child moves away, what we identify from the second child's problem is the same as the case of the only child.

Our further identification argument relies on subgame perfection. Consider two-child families. The choice problem the first child faces is contingent on the second child's *strategy*, i.e. $s_2 \in \{(\text{Far, Far}), (\text{Far, Near}), (\text{Near, Far}), (\text{Near, Near})\}$, where the first (second) argument of a strategy indicates the decision of the second child when the first child is "Far" ("Near"). Suppose $s_2 =$ (Near, Far), which we label "preempted". The first child's choice problem compares $u_1^{\alpha}(1)$ and $u_1^{\alpha}(1) + u_1^{\beta}$ and hence identifies u_1^{β}. In a simple model with no $u^{\gamma}(N_{-i})$, the identified u_1^{β} in turn leads to the identification of $u_i^{\alpha}(1)$ and $u_i^{\alpha}(2)$, controlling for X_i and under the assumption that the preferences do not vary by birth order. In a similar vein, the identification of $u_i^{\alpha}(3)$ and $u_i^{\gamma}(N_{-i})$ can be shown by examining the choice problems in the three- and four-child families.

It is worthwhile to note several points. First, this identification argument does not rely on the functional form assumption of $u^{\alpha}(N), u^{\beta}$, and $u^{\gamma}(N_{-i})$. Second, for the identification argument in the above two-child family case to hold, it is crucial that the second child chooses (Near, Far) with a positive probability. Since this paper is on the public good problem in which decisions are strategic substitutes, the situation in which (Near, Far) never occurs is not of interest.[14] Third, the standard "reduced-form" simultaneity bias due to the interdependence among players' decisions does not exist in this framework, because the dependence is explicitly solved for an equilibrium. Although strategies, s_i, include off-the-equilibrium-path plans we do not observe in data, our model explicitly distinguishes different strategies.

6 Results

6.1 Models with Within-Family Interactions

Table 4 compares the results of models with various forms of family interactions for our 2010 sample. Model [1] in Column 1 is a replication of the baseline probit result in Table 3 for ease of comparison. In Column 2, we add three dummy variables based on birth order information: the only-child dummy and the dummies for sons and daughters who have no younger sibling. Thus the reference group is children in multi-child families who have a younger sibling. This approach takes within-family relationship into account in a nonstructural way. We find only children more likely to live near parents than children with siblings, as is well-documented in the literature (e.g. Holmlund et al, 2009), and that youngest daughters are somewhat more likely to live near parents than elder siblings are, though this effect is not significant for youngest sons. The sibling structure variables slightly increase the fit of the model in terms of log likelihood. These results indicate the presence of within-family interaction. Younger daughters tend to live with their parents not only because of their age but also because of their birth order. Though the finding is consistent with

[14] In general, the above identification argument hinges on the state contingency of players' strategies. If every player's decision is independent of the decisions his upstream players have made, the identification strategy breaks down, but such a case lacks strategic interaction and does not require a model with externality.

the hypothesis that elder siblings tend to move far away to free-ride on the future care provision by their younger siblings, nevertheless, we cannot conclude from this simple model whether this is due to the sequential nature of location decisions among siblings, or that it occurs because elder and younger siblings have different costs and outside options. In the structural models below, we do not use these sibling structure variables; we instead explicitly model externality and sequential decision making among siblings.

[Insert Table 4: Improvement from Adding Interactions: 2010]

The first step of our structural modeling is a simple within-family correlation in the error terms, $\{\varepsilon_{i,h}\}_{i=1}^{I_h}$. The off-diagonal elements of household h's covariance matrix, Ω^h, are all equal to a correlation coefficient, $\rho \in (-1, 1)$. This correlation captures resemblance in the preferences of siblings, unobserved family characteristics that affect location decisions by siblings, and behavioral interaction among siblings. The result shown in Model [3] in Table 4 testifies a significant positive correlation in the error term. Compared with the baseline probit model, the model fits the data much better. The proportion of correctly predicted observations, which is defined based on the outcome with the highest fitted probability, also shows an improvement, especially for large families. There is no substantial change in the other coefficients.

Now we explicitly introduce externality, first with a constant externality, $u^{\alpha} = \alpha$. This term captures behavioral interdependence between the decisions of siblings. As shown in Model [4], we find a positive and significant estimate of α, illustrating the presence of a positive externality in the location decision of siblings. It leads to an even larger ρ. This makes sense for the following reason. The positive externality creates strategic substitutability in the decision of siblings. Without explicitly modeling this externality, the correlation in the error term needs to reflect this negative

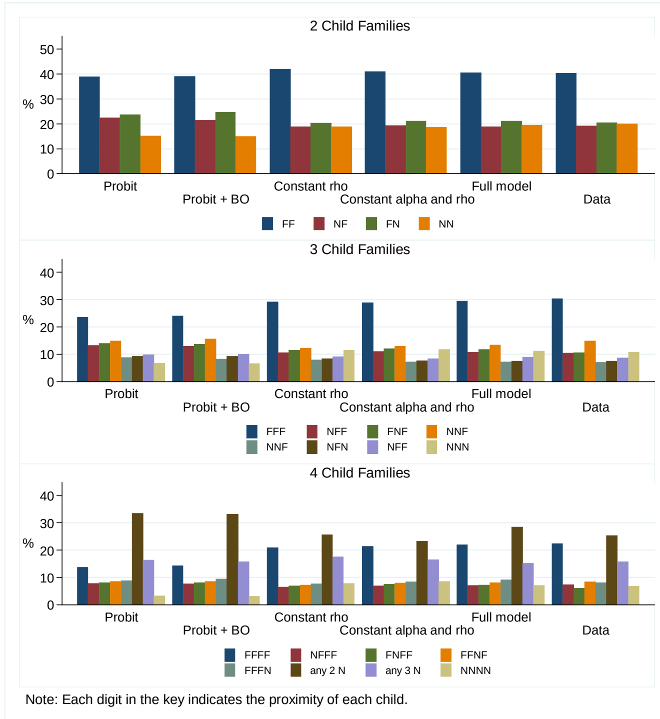

Figure 9: Predicted and Observed Location Patterns: 2010

behavioral correlation, resulting in a smaller estimate of ρ in the previous specification. Again, there is not much change in the other coefficients. To confirm the robustness of these results, we replicate Model [4] for other waves and find α and ρ always positive and highly significant (see Table 10 in Appendix B).

From the probit model to the model with α and ρ, the model fit improves over every step of elaboration. In terms of log L, incorporating correlation ρ contributes most, but incorporating externality α also shows a decent improvement. These improvements arise through a particular aspect. At the individual child level, the proportion of correct prediction shows no major improvement. Rather, the significant improvement is found at the family level. This is reasonable because both α and ρ model within-family interaction. The family-level fit is graphically illustrated in Figure 9. The two probit models are much worse in imitating the actual location patterns.

6.2 Model with Heterogeneous Externality

While the results so far find a strong indication of positive externality, the extent of externality is assumed to be constant over families. This is a strong restriction. To allow for potential heterogeneity in externality, we now parameterize u^α, u^γ, and ρ as specified in (2) and (5). When parameterizing these terms, we need to choose the set of covariates for each term. Including the full set of covariates in every term is impractical, because it makes precise identification of parameters difficult and increases the computational burden. Thus, we need a reasonably general but parsimonious specification. Admittedly, the choice of variables is rather arbitrary. Two considerations guide us to our final model. First we select variables that are likely to be relevant for each term based on the intention of each term. Variables in u^α are supposed to be the determinants of innate preferences for altruism, while variables in u^γ are supposed to affect the cost and benefit of cooperation. Second, we adjust the set of variables by attempting various specifications. We find variables that are always estimated with a large standard error and/or without statistical and economic significance. These variables are not included in our final specification. Regarding ρ, we allow the correlation between $\varepsilon_{i,h}$ and $\varepsilon_{j,h}$ to depend on the age and gender differences between children i and j.

Table 5 reports the result of the full model for 1998 and 2010. Compared with the model with constant α and ρ, the model fit is much improved both in terms of log likelihood and correct prediction, indicating the importance of heterogeneity in externality.[15] The coefficients in the u^β term are estimated less precisely than those in the constant externality model, but the sign and magnitude of each coefficient are fairly similar to the previous models. The variables that we also

[15] The LR test confirms that the improvement is significant at standard significance levels.

include in u^α and u^γ show somewhat different coefficients, which is not surprising. The majority of estimates for 1998 and 2010 tell us the same story, though there are modest differences between the two waves. Since our econometric framework is inherently cross-sectional, we do not make much attempt to interpret or investigate these differences and focus on findings that are robust across the two waves.

[Insert Table 5: Heterogeneous Externality: 1998 and 2010]

The estimates of u^α, u^γ, and ρ also confirm the importance of heterogeneity. The correlation in the error term is stronger for siblings of closer age and of the same sex than for others; those siblings tend to share similar preferences. The reduction in correlation due to gender difference is equivalent to the reduction due to an approximately 10 year age difference in 2010 and a 17 year age difference in 1998. The altruism term, u^α, varies across children and families. Altruism is strongest toward widowed mothers and weakest toward widowed fathers. Children with more education are more altruistic. In 2010, the positive estimates of α_1 and α_2 suggest that the utility gain from the u^α term increases in the number of children living near the parent. This effect, however, is not found in 1998.[16] The value of u^α is always positive. It ranges $[0.026, 0.522]$ in 2010 and $[0.133, 0.717]$ in 1998. The cooperation term, u^γ, also exhibits heterogeneity. It ranges $[-0.182, 0.233]$ in 2010 and $[0.008, 0.300]$ in 1998. Thus, though there tends to be a positive cooperation effect, it can take a negative value and its size is smaller than the altruism term. u^γ is larger for younger children, and for single children (only in 2010). One straightforward interpretation of this heterogeneity is that younger single siblings enjoy living near to or with each other. This story has little to do with the provision of care and attention. Alternatively, younger children tend to have less experience of care

[16] This may reflect a more direct link between the amount of care and the number of children living nearby in 2010. As the population ages, the care needs become higher, while coresidence and proximate living become less common.

provision and hence mutual assistance reduces the (actual and perceived) cost of providing care and attention. Another possible interpretation is that since we somewhat control the age of parents by $P\_age5064$ and P_age80p, the children's lower age means that the parents are relatively older for this cohort of children. For relatively older parents, young children may better understand the future uncertainty regarding care provision and prepare for the future in a more cooperative way. Having the third sibling nearby does not significantly change u^γ. The ranges of u^α and u^γ indicate that while in many families, significant externality influences the decision making of siblings, there are families in which externality is negligible. This heterogeneity is also pointed out in existing studies (Checkovich and Stern, 2002; Byrne et al, 2009; Knoef and Kooreman, 2011).

How does this heterogeneity affect the extent of inefficiency and strategic interaction? The key findings from the simulation section are: (1) joint-utility inefficiency increases with the absolute size of externality terms, u^α and u^γ; (2) the prisoner's dilemma and associated Pareto inefficiency arise when both u^α and u^γ are positive and large; (3) given a positive externality, both joint-utility and Pareto inefficiencies increase when u^β is small and children are similar, because coordination is more difficult due to larger conflicts of interest; and (4) the first-mover advantage is larger when u^β is not dominantly large and when the game exhibits a stronger strategic substitutability, that is, when u^α is large and u^γ is small.[17]

Externality (u^α and u^γ) is larger in a family with a widowed mother and with children who are younger (conditional on the parent's age) and have higher education. Among these families, coordination becomes difficult and inefficiency becomes larger when u^β is small and ρ is large—when the parent (widowed mother) is educated and does not own a home and when children are married and similar to each other. These factors affect the extent of the first-mover advantage the

[17] In three- and four-child families, larger values of α_1 and α_2 weaken strategic substitutability.

same way except for the effect of u^γ; the sequential distortion becomes larger when the positive externality from cooperation is smaller—when siblings are older (relative to the parent's age). Given the estimated range of u^α and u^γ, we can learn the magnitudes of inefficiency loss and the first-mover advantage from Figures 5 to 7. For the estimated ranges of u^α and u^γ, the first-mover advantage and Pareto inefficiency are much smaller than the efficiency loss in joint utility.

Several further remarks are worth making. First, we can now fully discuss why elder siblings tend to live far away from the parent. We find a negative age effect in u^β only in 1998; in the 1998 data, the private cost of living near the parent increases with age. This may capture the cohort effect over time. The older cohorts in 1998 may have been much less mobile with much more limited outside options, while cohorts currently in their working age are more mobile and tend to have outside options regardless of age. On the other hand, u^γ decreases with age. Compared with the age effect through u^β, this effect is equally important in 1998 and much more important in 2010. Younger siblings tend to live close to the parent because their utility is higher when they can cooperate with other siblings. These explanations have nothing to do with the first-mover advantage. The estimates in u^α and u^γ indicate the presence of strategic substitutability, and hence a first-mover advantage exists. As we further discuss below, however, the magnitude of the first-mover advantage is quite small.

Second, the probit regression shown in Table 3 finds that male children are more likely to live far from the parent, though this is not evident in the 1998 result. This is consistent with existing studies (e.g. Byrne et al, 2009; Compton and Pollak, 2009; Fontaine et al, 2009). According to the result of the full model in Table 5, though it is not very precisely estimated, this tendency is due to the gender difference in u^β, not in u^α. Male children are more likely to live far from parents not because they are not concerned about their parents, but because their private cost is larger, probably reflecting better outside opportunities for male children. Third, children's education shows more salient contrast. The probit result in Table 3 reveals a significant negative relationship between children's education and their propensity to live near their parents. The full model confirms that this negative effect arises completely through the private utility term, u^{β}, most likely reflecting the high opportunity costs to educated children of living near the parent. These children are actually concerned more about the well-being of their parents than less educated children. Less educated children are more likely to live near the parent not because they care more about their parents but because their utility gain from moving far away is limited.

6.3 Counterfactual Simulations

The estimated parameters reveal how $u^{\alpha}, u^{\beta}, u^{\gamma}$, and ρ vary across families and siblings, and guide us to counterfactual simulations to quantitatively illustrate how expected utility, efficiency loss, and the first-mover advantage vary across families.

We first evaluate the current utility level of each child, $u_{i,h}$, based on estimated parameters $\widehat{\theta}$ and data, $\left\{a^{o}_{i,h}, X_{i,h}\right\}_{i=1}^{I_h}$. Since we do not observe the values of $\varepsilon_{i,h}$, what we compute is the expected value of $u_{i,h}$ conditional on the data. It measures the expected utility gain of having an option to live near a parent, compared with the situation in which no child has such an option. As is the case with standard probit models, the level of utility itself has no meaningful interpretation. This is because the level of utility is not identified from observed discrete choice and so normalization is necessary; the random utility framework only concerns relative utility differentials across alternatives. Moreover, the comparison of the relative size of expected utility across individuals with different X requires caution, because it depends on how we normalize the discrete model: currently we normalize it as $u_{i,h} = 0$ when $\mathbf{a}_h = 0$. Nevertheless, the evaluated utility profile across families is informative in examining the welfare implications of externality.

Evaluating $u_{i,h}$ is not straightforward for several reasons. First, we do not observe $\varepsilon_{i,h}$ in data. If we knew the true values of $\varepsilon_{i,h}$, we could calculate the utility values, $u_{i,h}\left(\mathbf{a}_h^o, \mathbf{X}_h, \boldsymbol{\varepsilon}_h\right)$. But since $\varepsilon_{i,h}$ is not available in data, we need to infer the expected value of $u_{i,h}$. We can compute the expected utility value by taking the following integral over the domain of $\boldsymbol{\varepsilon}_h$ that rationalizes observed outcome, $\mathbf{a}_h^o$. By denoting this integration domain over the space of $\boldsymbol{\varepsilon}_h$ as $\Delta\left(\mathbf{a}_h^o\right)$,

$$\widehat{u}_{i,h} = \frac{1}{\Pr\left(\boldsymbol{\varepsilon}_h \in \Delta\left(\mathbf{a}_h^o\right)\right)} \int_{\boldsymbol{\varepsilon}_h \in \Delta\left(\mathbf{a}_h^o\right)} u_{i,h}\left(\mathbf{a}_h^o, \mathbf{X}_h, \boldsymbol{\varepsilon}_h\right) \phi\left(\boldsymbol{\varepsilon}_h\right) d\boldsymbol{\varepsilon}_h,$$

where $\phi\left(\boldsymbol{\varepsilon}_h\right)$ is the density function of $\boldsymbol{\varepsilon}_h$. Second, because this multidimensional integral does not have an analytical solution, a simulation method is necessary to numerically approximate this integral. Third, this simulation-based integration is complicated by the behavioral interaction among siblings. We can evaluate this integral and the probability in the denominator in exactly the same way that we simulate the likelihood function using the GHK simulator and subgame perfection.

We can also evaluate the two inefficiency measures for each family. If we were given $\left(\mathbf{X}_h, \boldsymbol{\varepsilon}_h\right)$ and θ, obtaining efficient location configurations and associated utility levels would be straightforward; comparing $\sum_{i=1}^{I_h} u_{i,h}\left(\mathbf{a}_h, \mathbf{X}_h, \boldsymbol{\varepsilon}_h\right)$ over all the possible location configurations, which is as many as $2^4 = 16$ for four-child families. However, because we do not observe $\boldsymbol{\varepsilon}_h$, we need to evaluate the expected utility values of each location configuration by the Monte Carlo integration to obtain the optimal solution for each simulation draw of $\boldsymbol{\varepsilon}_h$.

Table 6 presents these results by sibling structure. The first column reports the sample average of first children's expected utility, $\frac{1}{N}\sum_h \widehat{u}_{1,h}$, by sibling structure. The first number in the first column is the sample average of expected utility of sons who have no sibling. The last two columns in Table 6 report the sample average of Pareto and joint-utility inefficiencies at family level. Table 6 reveals several robust findings. First, utility levels increase with the number of siblings. This is due to positive externality. At the same time, for the same reason, both inefficiencies are larger for families with more siblings. With more siblings, the same level of externality has a larger consequence and the free-rider problem becomes larger. Second, there is no clear evidence of the first-mover advantage. Younger siblings sometimes receive a larger utility than the elder siblings. Third, joint-utility inefficiency is much larger than Pareto inefficiency, as illustrated in the simulation section. Fourth, both inefficiencies are the largest in son-only families and the smallest in daughter-only families.

[Insert Table 6: Expected Utility and Family Inefficiency by Sibling ...]

Though these results are informative, it is not easy to draw definitive conclusions on externality and strategic interaction, because observed covariates, $\mathbf{X}_h$, may be responsible for the above findings. To extract the pure first-mover advantage and the effects of externality, we conduct a counterfactual simulation on families with "identical" siblings. Specifically, we first choose family and child characteristics that yield the largest u^α and the smallest u^γ to illustrate the case of the largest strategic substitutability. The remaining covariates are set to the most common values in the sample. This leads to the following experiment setup: families with a white widowed mother aged between 65 and 79, with an ADL index of 2.747, without a stroke, with poor eyesight, with some college education, and living in an owned house in a metropolitan area. All children are aged 47 and married and have a university degree and one child. Based on this setup, we simulate the number of siblings and the gender composition. Except for the gender and birth order, siblings are identical; they are even at the same age so that we can evaluate the pure birth order effect. We also highlight the gender difference. Although the gender difference estimated in our full model is not large or significant, we investigate it because it has real world relevance and is still useful for intuitively illustrating the working of the model. This Monte Carlo simulation is much simpler than the previous simulations based on sample distribution. We simply generate the unobserved errors following the modeled data generating process with estimated parameter values, and use the errors to solve the game for the equilibrium and optimal location configurations.

Table 7 presents the results. The first column reports the expected utility of the first child and indicates several trends. First, everything else equal, the first child's utility increases with the number of siblings. This is due to externality. Second, everything else equal, female children in one-child and two-child families enjoy higher utility than male children, but in three- and four-child families, male children enjoy higher utility than female children. This is because though u^{β} is higher for females, the effect of male children's larger u^{α} dominates in larger families.

[Insert Table 7: First-Mover Advantage and Inefficiency by Sibling ...]

The next three columns show the utility difference between the eldest and younger siblings to see the birth order effect. The utility gap of the youngest children in any same-gender siblings, from the M-M case (two sons) to the F-F-F-F case (four daughters), shows robust evidence of the first-mover advantage. The expected utility of the youngest child is consistently smaller than that of the eldest child. The size of these utility gaps, however, is very small compared with the level of expected utility values, even for this experiment setup with a large strategic substitutability; thus, the first-mover advantage is effectively negligible in reality. Note that we do not see a clear advantage or disadvantage for middle children.[18]

A more stark finding from this table is the coordination among male and female siblings. Compare the first child's utility levels for the M-M and M-F families. The first son's utility is higher when the younger sibling is female. Also compare M-F and F-M families. In M-F families, the daughter's utility is lower than the son's, and in F-M families, the son's utility is higher than the sister's, despite his second-mover disadvantage. Similar gaps can also be found for four-child families. The reason driving these findings is that male children incur a larger private cost of living near the parent (a smaller u^β) and have a larger altruism effect (a larger u^α). The gender difference reduces the chance of conflicts within a family. The daughter lives near the parent regardless of her brother's decision and the son easily free-rides on his sister.[19]

The last two columns in Table 7 present simulated inefficiencies. Generally what we learn from here is quite similar to what is observed in Table 6. First, the magnitude of joint-utility inefficiency is much larger than that of Pareto inefficiency and first-mover advantages. If we compute per-child family efficiency loss, it is often worth almost 10% of expected utility values. Second, inefficiencies increase with the number of siblings. With more siblings, the same level of externality has a larger impact and coordination becomes more difficult. Holding the number of siblings fixed, the extent of inefficiency is the largest for all-son families and the smallest for all-daughter families, with mixed-gender families in between. Because males have a larger u^α and a smaller u^β, coordination is more difficult in all-son families than in all-daughter families.

With a positive externality, family inefficiency is effectively the under-provision of proximate

[18] Also note that the first-mover advantage does not increase with the number of siblings. This is because while the effect of altruism externality increases with the number of siblings, the effect of cooperation externality increases too, which neutralizes strategic substitutability.

[19] These gender utility gaps will be somewhat different if we change how we normalize the model, e.g. assuming $u_{i,h} = 0$ when $\mathbf{a}_h = 1$ instead of $\mathbf{a}_h = 0$. Nevertheless, this normalization is without loss of generality and it does not affect the main story.

living.[20] We can translate the above welfare inefficiency measures into a more interpretable measure of location patterns. Table 8 presents the probability that a parent has at least one child nearby. The left half of Table 8 presents the simulation results for the families with "identical" siblings, as described in the previous table, and the right half shows the results for the empirical distribution. In each part, columns 1 and 2 show the probability of having at least one child nearby for non-cooperative equilibrium and joint utility maximization, respectively. The third column shows the difference between these two, and the last column re-produces family inefficiency from Tables 6 and 7 for reference purposes. The table proves that this location pattern inefficiency measure follows the pattern of family inefficiency quite closely, and that the difference between the joint utility maximizing location pattern and the equilibrium one is large. Of the 2010 HRS families with multiple children, had families fully internalized externality and jointly maximized family utility, 17% more parents would have had at least one child nearby.[21]

[Insert Table 8: Probability to Have At Least One Child Near Parent]

Note that our framework does not include parents' welfare, though this is partly captured by the children's altruism term. Hence, the terms "inefficiency" and "under-provision" in this paper do not necessarily coincide with actual family inefficiency. If we assume, however, that children's proximate living increases parents' utility, our inefficiency measures are the lower bound of actual inefficiencies.[22]

[20] Strictly speaking, inefficiency and under-provision do not necessarily coincide. For example, in families with $u^{\alpha} = 0$ and $u^{\gamma} < 0$, over-provision is possible.

[21] Knoef and Kooreman (2011) also find a large implication of inefficiency in joint utility in a similar context.

[22] Strictly speaking, the bias also depends on the structure of ex-post bargaining and intergenerational transfers. Also, Maruyama (2012) discusses the possiblity that altruistic parents may incur disutility from placing a care burden on children.

6.4 Alternative Model: Joint Utility Maximization

As discussed earlier, our model assumes the non-cooperative behavior of siblings. The discrete, irreversible, and long-term nature of the location choice justifies the assumption to some extent, but siblings may be able to arrange enforceable side-payment transfers to achieve the highest joint utility possible, as discussed by Engers and Stern (2002). We examine this possibility by estimating a model of joint utility maximization. This model uses the same functional form specification as before, namely, (2), (3), and (4), and instead of individual utility maximization, we assume the following joint utility maximization:

$$\max_{\mathbf{a}_h \in \{0,1\}^{I_h}} \sum_{i}^{I_h} u_i \left(\mathbf{a}_h \right).$$

This family problem is estimated as a multinomial probit model. Because the multivariate normal distribution does not have a tractable analytical form, the estimation is based on the standard method of simulated likelihood.[23]

Table 9 compares the joint maximization and the non-cooperative models based on different functional form specifications. We find non-cooperative models better fit with data, indicating the presence of strategic interactions.[24]

[Insert Table 9: Test of Joint Decision Model]

[23] Börsch-Supan et al (1988) and Hoerger et al (1996) also apply multinomial probit models to study living arrangements.

[24] Engers and Stern (2002) conduct a similar model comparison in their framework of family long-term care decisions, and favor a game-theoretic model over a "collective" model.

7 Conclusion

This paper studies externality and strategic interaction among adult siblings regarding their location decision relative to their elderly parents, based on a sequential game-theoretic framework. We find a positive externality and strategic interaction. Siblings make location decisions non-cooperatively and proximate living with elderly parents is an under-provided public good. The impact of this public good problem is striking; of the 2010 HRS families with multiple children, had families fully internalized externality and jointly maximized family utility, 17% more parents would have had at least one child nearby.

The estimated inefficiency is mainly due to the public good problem; the sequential strategic behavior and the prisoner's dilemma only marginally account for the inefficiency. Many multi-child families can achieve a larger joint utility, by lowering some child(ren)'s utility. Historically in many countries, social norms and traditions have forced daughters to assume caregiving obligations (e.g. Holroyd, 2001; Silverstein et al, 2006), serving as an enforceable mechanism for families to achieve a larger joint utility. Without such a mechanism, improved gender equality and increased female labor force participation may have reduced the joint utility of families.

Our findings have implications for future research and public policies. Without taking the interdependence among siblings and the free-rider problem into consideration, research and policy developments may draw misleading conclusions. The first-mover advantage and the prisoner's dilemma, however, are of negligible empirical importance. In the coming decades in modern aging societies, on one hand, we anticipate some factors that will reduce the extent of externality and inefficiency in siblings' location choice, such as a smaller number of siblings, more efficient travel, and a more developed formal care sector. On the other hand, the younger generation is more educated and more mobile, and families are more likely to have a widowed mother: these factors contribute to inefficiency. One way to achieve the joint-utility-maximizing location pattern is to develop an enforceable compensation scheme from those who free-ride to those who provide care. Alternatively, policies that reduce the private costs of caring elderly parents also enhance social welfare. Policy options include a developed formal care sector, generous leave and travel supports for carers, and subsidies and tax benefits that facilitate proximate living with elderly parents and the relocation of elderly parents.

A The HRS Cohorts

To ensure the robustness of our cross-sectional framework, we compare the results from other waves in the HRS. The HRS starts in 1992. The first "original HRS" cohort is age 51-61 (born 1931-1941). At the baseline only non-institutionalized individuals are sampled, though the HRS follows those who move to nursing homes in later waves. In the following year, it interviews a much older cohort (born before 1923), the Study of Assets and Health Dynamics among the Oldest Old (AHEAD) cohort. In 1998, the AHEAD cohort is merged with the HRS respondents, and two new cohorts are added to the study: the Child of the Depression Age Cohort (CODA) (born 1924-1930) and the War Babies cohort (born 1942-1947). In 2004, the Early Baby Boomers (born 1948-1953) are added, and the Mid Baby Boomers (born 1954-1959) in 2010.

Figure 10 depicts the evolution of cohorts from 1998 to 2010 for the elderly parents who satisfy our sample selection criteria (1) – (8) as discussed in subsection 2.2. In 1998, our sample comprises mostly the AHEAD and the original HRS cohorts. Over time, the AHEAD cohort becomes smaller, mostly due to death, while the HRS cohort becomes dominant. The War Babies and the Early

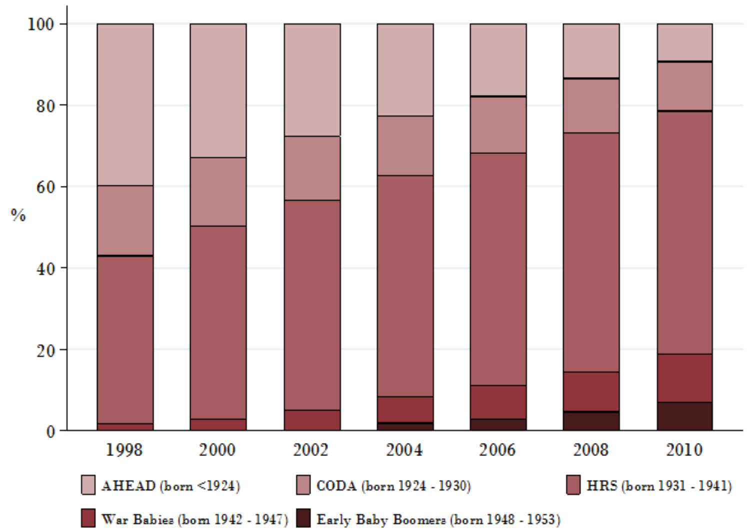

Figure 10: HRS Cohorts: 1998 - 2010

Baby Boomers make up relatively small shares of our sample because these younger parents tend to have younger children below 30. In 2010, an even younger cohort, the Mid Baby Boomers, is added to the HRS, but, as at the time we conduct this analysis, the 2010 public use data do not include this latest cohort. The missing latest cohort is unlikely to introduce significant bias to our results, given that our focus is on the parents of relatively mature adult children (at least 30 years of age). The representation of the new cohort will be small in our sample, similar to the Early Baby Boomers when they are added in 2004. Figure 10 also reveals the close resemblance of any two consecutive waves, due to the nature of longitudinal data. The 1998 and 2010 waves, however, are fairly apart from each other, and will provide a good robustness check.

B Robustness of Externality over Waves

[Insert Table 10: Robustness of Externality 1998 — 2010]

References

[1] Bernheim, B. D., A. Schleifer, and L. H. Summers (1985), "The Strategic Bequest Motive," *Journal of Political Economy*, 93: 1045-76.

[2] Berry, S. (1992), "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60(4): 889-917.

[3] Börsch-Supan, A., L. J. Kotlikoff, and J. Morris (1988) "The Dynamics of Living Arrangements of the Elderly," Working Paper 2787, NBER.

[4] Byrne, D., M. S. Goeree, B. Hiedemann, and S. Stern (2009), "Formal Home Health Care, Informal Care, and Family-Decision Making," *International Economic Review*, 50(4): 1205-42.

[5] Checkovich, T. and S. Stern (2002), "Shared Caregiving Responsibilities of Adult Siblings with Elderly Parents," *Journal of Human Resources*, 37: 441-478.

[6] Compton, J. and R. Pollak (2009), "Proximity and Coresidence of Adult Children and their Parents: Description and Correlates," Working Paper, WP2009-215, University of Michigan.

[7] Dostie, B. and P. T. Léger (2005), "The Living Arrangement Dynamics of Sick, Elderly Individuals," *Journal of Human Resources*, 40(4): 989-1014.

[8] Engers, M. and S. Stern (2002), "Long-Term Care and Family Bargaining," *International Economic Review*, 43(1): 73-114.

[9] Fontaine, R., A. Gramain, and J. Wittwer (2009), "Providing Care for an Elderly Parent: Interactions among Siblings?" *Health Economics*, 18(9): 1011-1029.

[10] Gal-Or, E. (1985), "First Mover and Second Mover Advantages," *International Economic Review*, 26(3), 649-653.

[11] Hank, K. (2007), "Proximity and Contacts between Older Parents and Their Children: A European Comparison," *Journal of Marriage and Family*, 69: 157-173.

[12] Hequembourg, A. and S. Brallier (2005), "Gendered Stories of Parental Caregiving among Siblings," *Journal of Aging Studies*, 19: 53-71.

[13] Hiedemann, B. and S. Stern (1999), "Strategic Play among Family Members When Making Long-Term Care Decisions," *Journal of Economic Behavior and Organization*, 40: 29-57.

[14] Hoerger, T. J., G. A. Picone, F. A. Sloan (1996), "Public Subsidies, Private Provision of Care and Living Arrangements of the Elderly," *Review of Economics and Statistics*, 78(3): 428-440.

[15] Holroyd, E. (2001), "Hong Kong Chinese Daughters' Intergenerational Caregiving Obligations: A Cultural Model Approach," *Social Science and Medicine*, 53: 1125-34.

[16] Holmlund H., H. Rainer, and T. Siedler (2009), "Geography of the Family: Comment," mimeo. Available http://www.tsiedler.net/pdf/articles/Geo_of_Family.pdf.

[17] Hotz, V. J., K. McGarry, and E. Wiemers (2010), "Living Arrangements of Mothers and Their Adult Children over the Life Course," mimeo.

[18] Johar, M. and S. Maruyama (2011), "Intergenerational Cohabitation in Modern Indonesia: Filial Support and Dependence," *Health Economics*, 20(S1): 87-94

[19] Knoef, M. and P. Kooreman (2011), "The Effects of Cooperation: A Structural Model of Siblings' Caregiving Interactions," University of Tilburg. Available: http://members.ziggo.nl/peterkooreman/ltc.pdf.

[20] Konrad, K. A., H. Künemund, K. E. Lommerud, and J. R. Robledo (2002), "Geography of the Family," *American Economic Review*, 92(4): 981-998.

[21] Lin I., N. Goldman, M. Weinstein, Y. Lin, T. Gorrindo, and T. Seeman (2005), "Gender Differences in Adult Children's Support of Their Parents in Taiwan," *Journal of Marriage and Family*, 65: 184-200.

[22] Maruyama, S. (2010), "Estimation of Finite Sequential Games," *UNSW Australian School of Business Research Paper*, 2010/ECON22, University of New South Wales.

[23] Maruyama, S. (2012), "Inter Vivos Health Transfers: Final Days of Japanese Elderly Parents," mimeo.

[24] McGarry, K. and R. Schoeni (1995), "Transfer Behavior in the Health and Retirement Study: Measurement and the Redistribution of Resources within the Family," *Journal of Human Resources*, 30: S184-S226.

[25] OECD (2005) *The OECD Health Project: Long-Term Care for Older People*. OECD, Paris.

[26] Pezzin, L. E., R. A. Pollak, and B. S. Schone (2007), "Efficiency in Family Bargaining: Living Arrangements and Caregiving Decisions of Adult Children and Disabled Elderly Parents," *CESifo Economic Studies*, 53(1): 69-96.

[27] Pezzin, L. E. and B. S. Schone (1999), "Intergenerational Household Formation, Female Labor Supply and Informal Caregiving: A Bargaining Approach," *Journal of Human Resources*, 34(3): 475-503.

[28] Sakudo, M. (2008), "Strategic Interactions between Parents and Daughters: Co-Residence, Marriage, and Intergenerational Transfers in Japan," mimeo, University of Chicago.

[29] Silverstein, M., D. Gans, and F. M. Yang (2006), "Intergenerational Support to Aging Parents: The Role of Norms and Needs," *Journal of Family Issues*, 27(8): 1068-84.

Table 1: Sibling Location Patterns by Birth Order: 2010

1 child families (*N*=799)

N of children living near	Location patterns
0	Far – 54.1%
1	Near – 45.9%

2 child families (*N*=1,809)

N of children living near	Location patterns		Total
0	FF – 40.4%		40.4%
1	NF – 19.2%	FN – 20.5%	39.7%
2	NN – 19.9%		19.9%

3 child families (*N*=1,311)

N of children living near	Location patterns			Total
0	FFF – 30.4%			30.4%
1	NFF – 10.5%	FNF – 10.6%	FFN – 14.8%	35.9%
2	NNF – 7.0%	NFN – 7.5%	FNN – 8.6%	23.1%
3	NNN – 10.7%			10.7%

4 child families (*N*=700)

N of children living near	Location patterns				Total
0	FFFF– 22.4%				22.4%
1	NFFF – 7.3%	FNFF – 6.0%	FFNF – 8.4%	FFFN – 8.1%	29.9%
2	NNFF – 3.6%	NFNF – 4.0%	NFFN– 4.4%	FNNF – 4.6%	25.3%
			FNFN– 3.6%	FFNN – 5.1%	
3	NNNF – 4.0%	NNFN – 4.9%	NFNN– 3.0%	FNNN – 3.9%	15.7%
4	NNNN– 6.7%				6.7%

Note: Each digit in the key indicates the proximity of each child, with the first digit representing the eldest child. E.g., "FFN" indicates that the elder siblings live far from the parent and the youngest child lives near the parent. "N" includes coresidence.

Table 2: Definition and Summary Statistics of Explanatory Variables

Variable	Definition	Mean (std dev) 2010	Mean (std dev) 1998
Parent			
P_father	=1 if male and does not live with a spouse	0.136	0.108
P_mother	=1 if female and does not live with a spouse	0.436	0.419
*P_age5064**	Parent's age <65	0.129	0.261
*P_age6579**	Parent's age 65-79 (reference group)	0.632	0.554
*P_age80p**	Parent's age 80+	0.239	0.185
P_white^	=1 if the parent is White	0.836	0.864
*P_ADL**	# activities of daily living (ADL) performed with difficulty (out of 6 tasks)	2.747 (2.33)	3.048 (2.36)
P_stroke#	=1 if self-report having a stroke	0.114	0.101
P_eye#	=1 if self-report having a poor eyesight	0.756	0.764
P_nodegree^	=1 if highest education is GED/high school or lower	0.164	0.253
P_somecollege#	=1 if highest education is some college (reference group)	0.558	0.552
P_college#	=1 if highest education is university	0.278	0.195
Metro1	=1 if lives in a metro area (reference group)	0.449	0.261
Metro2	=1 if lives in an area within 250 km from metro	0.216	0.175
Metro3	=1 if lives in an area more than 250 km from metro	0.300	0.176
Metro_missing	=1 if remoteness information is missing	---	0.388
House	=1 if own a residential house	0.689	0.714
Child			
C_age	Child's age	46.72 (7.96)	43.69 (8.18)
C_male	=1 if the child is male	0.508	0.509
C_spouse	=1 if the child lives with a spouse (partnered child)	0.707	0.681
C_nodegree	=1 if the child's highest education is high school or lower	0.070	0.088
C_somecollege	=1 if the child's highest education is some college (reference group) – include a small number of parents with missing information	0.592	0.587
C_college	=1 if the child's highest education is college or above	0.338	0.325
C_kids_spouse	# children of the child when the child is married	1.456 (1.54)	1.323 (1.49)
C_kids_single	# children of the child when the child is single	0.338 (0.90)	0.331 (0.91)
C_onlychild	=1 if the child is the only child	0.072	0.085
C_m_youngest	=1 if the child is male and the last child of siblings	0.173	0.168
C_f_youngest	=1 if the child is female and the last child of siblings	0.170	0.167

Note:

^ Both parents if a spouse is present

* Average if a spouse is present

Either one if a spouse is present

Standard deviation is reported in parentheses only for continuous variables

Note: ADL includes dressing, walking, bathing, eating, getting in and out of bed, and using the toilet.

Table 3: Preliminary Probit Regressions: 1998 – 2010

	1998 [1]	2000 [2]	2002 [3]	2004 [4]	2006 [5]	2008 [6]	2010 [7]
	Coeff	Coeff	Coeff	Coeff	Coeff	Coeff	Coeff
P_father	-0.108***	-0.078**	-0.071**	-0.054*	-0.091***	-0.064*	-0.060*
P_mother	0.039	0.056**	0.064***	0.122***	0.099***	0.135***	0.092***
P_age5064	-0.070**	0.021	0.015	-0.026	0.034	-0.025	-0.024
P_age80p	0.080**	0.080**	0.036	0.124***	0.085**	0.067**	0.064*
P_white	-0.070**	-0.119***	-0.044	-0.014	-0.045	-0.052*	-0.058*
P_ADL	-0.015***	-0.008*	-0.017***	-0.014***	-0.019***	-0.012**	-0.014***
P_stroke	0.046	0.077**	0.010	0.130***	0.114***	0.096***	-0.003
P_eye	-0.018	0.082***	0.056**	0.043*	0.013	0.067***	0.042
P_college	-0.206***	-0.112***	-0.182***	-0.195***	-0.159***	-0.100***	-0.145***
P_nodegree	0.155***	0.123***	0.191***	0.183***	0.110***	0.078***	0.111***
Metro2	-0.025	-0.060**	-0.048*	-0.023	-0.018	0.010	-0.028
Metro3	-0.103***	-0.133***	-0.116***	-0.078***	-0.065***	-0.089***	-0.042
House	0.091***	0.053**	0.075***	0.122***	0.070***	0.064***	0.133***
C_age	-0.013***	-0.011***	-0.007***	-0.013***	-0.009***	-0.010***	-0.007***
C_male	-0.009	0.005	-0.039*	-0.024	-0.073***	-0.085***	-0.042*
C_college	-0.357***	-0.357***	-0.361***	-0.348***	-0.369***	-0.375***	-0.360***
C_nodegree	-0.151***	-0.060	-0.015	-0.037	-0.013	0.013	0.023
C_spouse	-0.320***	-0.376***	-0.310***	-0.329***	-0.336***	-0.322***	-0.385***
C_kids_partner	0.030***	0.022**	0.025**	0.027***	0.031***	0.039***	0.049***
C_kids_single	-0.013	-0.037**	-0.020	-0.016	-0.049***	-0.051***	-0.018
Constant	0.739***	0.448***	0.403***	0.555***	0.551***	0.482***	0.352***
Observations	13,337	12,790	13,104	13,137	12,801	12,279	11,150
Pseudo R-sq	0.035	0.041	0.039	0.040	0.036	0.033	0.036
Log L	-8,720.43	-8,020.19	-8,501.74	-8,531.97	-8,302.67	-7,985.76	-7,205.55

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively. For 1998, a missing dummy for region is included.

Table 4: Improvement from Adding Interactions: 2010

	Probit [1]		Probit +birth order[2]		Constant ρ (α=0) [3]		Constant α and ρ[4]	
	coefficient	std dev	coefficient	std dev	coefficient	std dev	coefficient	std dev
P_father	-0.060*	0.035	-0.064*	0.035	-0.066	0.044	-0.073	0.045
P_mother	0.092***	0.026	0.083***	0.026	0.098***	0.033	0.098***	0.034
P_age5064	-0.024	0.038	-0.019	0.039	-0.028	0.045	-0.035	0.046
P_age80p	0.064*	0.034	0.043	0.035	0.079*	0.041	0.078*	0.042
P_white	-0.058*	0.031	-0.056*	0.031	-0.065*	0.039	-0.070*	0.040
P_ADL	-0.014***	0.005	-0.014***	0.005	-0.015**	0.006	-0.015**	0.006
P_stroke	-0.003	0.035	-0.003	0.035	0.010	0.045	0.013	0.046
P_eye	0.042	0.027	0.043	0.027	0.041	0.034	0.042	0.034
P_college	-0.145***	0.028	-0.145***	0.028	-0.144***	0.034	-0.153***	0.035
P_nodegree	0.111***	0.031	0.114***	0.031	0.113***	0.040	0.120***	0.041
Metro2	-0.028	0.028	-0.025	0.028	-0.028	0.036	-0.027	0.037
Metro3	-0.042	0.026	-0.043*	0.026	-0.044	0.033	-0.047	0.033
House	0.133***	0.025	0.134***	0.025	0.130***	0.032	0.132***	0.033
C_age	-0.007***	0.002	-0.005**	0.002	-0.007***	0.002	-0.007***	0.002
C_male	-0.042*	0.024	-0.032	0.030	-0.042*	0.023	-0.041*	0.023
C_college	-0.360***	0.027	-0.358***	0.027	-0.351***	0.028	-0.349***	0.028
C_nodegree	0.023	0.047	0.025	0.047	0.034	0.049	0.039	0.049
C_spouse	-0.385***	0.038	-0.383***	0.038	-0.385***	0.038	-0.381***	0.038
C_kids_partner	0.049***	0.010	0.050***	0.010	0.048***	0.010	0.047***	0.010
C_kids_single	-0.018	0.016	-0.016	0.016	-0.017	0.016	-0.017	0.015
C_onlychild			0.181***	0.048				
C_m_youngest			0.043	0.040				
C_f_youngest			0.071*	0.040				
Constant	0.352***	0.111	0.222*	0.127	0.385***	0.126	0.333***	0.126
α (constant)							0.096***	0.030
ρ (constant)					0.251***	0.018	0.324***	0.029
Log *L*	-7,205.55		-7,197.85		-7,100.91		-7,095.55	
% correct prediction								
All children	62.65%		62.52%		61.65%		61.67%	
All families	37.65%		37.76%		38.93%		39.01%	
1 child families	59.32%		60.08%		58.45%		59.95%	
2 child families	40.80%		40.24%		41.74%		41.68%	
3 child families	29.60%		30.74%		31.73%		31.27%	
4 child families	19.86%		19.00%		22.86%		22.71%	

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively.

Table 5: Heterogeneous Externality: 1998 and 2010

	2010 [1]		1998 [2]	
	coefficient	std dev	coefficient	std dev
P_father	0.098	0.107	-0.102*	0.057
P_mother	0.067	0.062	-0.035	0.043
P_age5064	-0.019	0.049	-0.060	0.036
P_age80p	0.051	0.044	0.048	0.042
P_white	-0.080*	0.041	-0.067*	0.038
P_ADL	-0.016**	0.006	-0.015***	0.006
P_stroke	0.012	0.047	0.055	0.045
P_eye	0.046	0.035	-0.016	0.032
P_college	-0.220***	0.073	-0.223***	0.048
P_nodegree	0.096	0.070	0.113***	0.043
Metro2	-0.029	0.038	-0.024	0.041
Metro3	-0.050	0.034	-0.108***	0.041
House	0.136***	0.034	0.091***	0.030
C_age	0.000	0.003	-0.008***	0.003
C_male	-0.077	0.057	-0.007	0.032
C_ college	-0.396***	0.067	-0.413***	0.039
C_nodegree	0.040	0.097	-0.076	0.054
C_spouse	-0.332***	0.044	-0.335***	0.039
C_kids_partner	0.048***	0.010	0.031***	0.009
C_kids_single	-0.014	0.015	-0.014	0.014
Constant	-0.174	0.248	0.280	0.231
α_0 (constant)	0.265	0.170	0.303*	0.156
P_father	-0.231*	0.116	-0.021	0.061
P_mother	0.042	0.072	0.157***	0.046
P_college	0.082	0.082	0.003	0.052
P_nodegree	0.043	0.087	0.140***	0.051
C_male	0.057	0.075	-0.007	0.045
C_college	0.076	0.085	0.117**	0.051
C_nodegree	-0.008	0.134	-0.142*	0.079
α_1	0.491***	0.187	0.048	0.134
α_2	0.221	0.177	-0.010	0.095
γ_0 (constant)	0.553***	0.182	0.528**	0.210
C_age	-0.010***	0.003	-0.008***	0.003
C_spouse	-0.085*	0.048	0.012	0.047
γ_1	0.219	0.369	-0.160	0.267
ρ (constant)	0.509***	0.045	0.455***	0.038
Dage	-0.010**	0.004	-0.006	0.004
Dsex	-0.102***	0.030	-0.104***	0.028
Observations	11,150		13,337	
Log *L*	-7,066.48		-8,560.35	
% correct prediction				
All children	62.25%		61.19%	
All families	39.34%		38.70%	
1 child families	60.70%		60.34%	
2 child families	41.96%		40.65%	
3 child families	31.20%		29.62%	
4 child families	23.43%		21.84%	

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively. For 1998, a missing dummy for region is included.

Table 6: Expected Utility and Family Inefficiency by Sibling Structure: Sample Average

Sibling structure:	U1	U2	U3	U4	Pareto Inefficiency	Family Inefficiency
M	0.3356				0.0000	0.0000
F	0.3811				0.0000	0.0000
M – M	0.4308	0.4317			0.0039	0.0441
M – F	0.4317	0.4422			0.0033	0.0368
F – M	0.3825	0.4078			0.0034	0.0386
F – F	0.4253	0.4223			0.0030	0.0325
M – M – M	0.4907	0.4894	0.5064		0.0334	0.1413
F – F – F	0.4439	0.4233	0.4453		0.0220	0.1135
M – M – M – M	0.4533	0.4322	0.4238	0.4634	0.0801	0.3649
M – M – F – F	0.5108	0.4753	0.4711	0.5231	0.0501	0.2552
F – F – M – M	0.6100	0.6281	0.7143	0.7114	0.0353	0.1713
F – F – F – F	0.5472	0.5741	0.5883	0.6078	0.0287	0.1809

Note: M – F indicates two-child families with an elder brother and a younger sister.

Table 7: First-Mover Advantage and Inefficiency by Sibling Structure: Simulation for a Family with Large α

Sibling structure:	U1	U2 – U1	U3 – U1	U4 – U1	Pareto Inefficiency	Family Inefficiency
M	0.2417				0.0000	0.0000
F	0.2512				0.0000	0.0000
M – M	0.3444	-0.0057			0.0032	0.0476
M – F	0.3607	-0.0178			0.0027	0.0420
F – M	0.3500	0.0051			0.0025	0.0421
F – F	0.3461	-0.0035			0.0023	0.0365
M – M – M	0.4238	-0.0065	-0.0068		0.0164	0.1883
F – F – F	0.4059	0.0023	-0.0089		0.0121	0.1449
M – M – M – M	0.4738	-0.0003	-0.0018	-0.0053	0.0332	0.3694
M – M – F – F	0.4850	0.0028	-0.0290	-0.0342	0.0289	0.3291
F – F – M – M	0.4544	-0.0012	0.0273	0.0298	0.0276	0.3282
F – F – F – F	0.4480	0.0003	0.0027	-0.0031	0.0237	0.2916

Note: M – F indicates two-child families with an elder brother and a younger sister. The setup in the experiment: Families with a white widowed mother of age between65 and 79, with an ADL index of 2.747, without a stroke, with poor eyesight, with some college education, and living in an owned house in a metropolitan area. All children are aged 47 and married and have a university degree and one child. Monte Carlo simulation with 20,000 random draws.

Table 8: Probability of Having at Least One Child Near Parent

	Families with large α				Empirical distribution			
Sibling structure:	Nash equil.	Family optimum	Under-provision	Family inefficiency	Nash equil.	Family optimum	Under-provision	Family inefficiency
M	35.5%	35.5%	0.0%	0.0000	43.3%	43.3%	0.0%	0.0000
F	36.8%	36.8%	0.0%	0.0000	48.5%	48.5%	0.0%	0.0000
M – M	50.6%	68.8%	18.2%	0.0476	57.5%	74.1%	16.6%	0.0441
M – F	53.5%	70.8%	17.3%	0.0420	62.2%	76.1%	13.9%	0.0368
F – M	53.4%	70.7%	17.3%	0.0421	57.9%	72.9%	15.1%	0.0386
F – F	52.4%	68.4%	16.0%	0.0365	60.7%	73.7%	13.1%	0.0325
M – M – M	59.7%	89.7%	30.0%	0.1883	65.9%	89.1%	23.2%	0.1413
F – F – F	60.9%	87.6%	26.7%	0.1449	61.7%	83.7%	22.0%	0.1135
M – M – M – M	65.8%	97.3%	31.5%	0.3694	59.5%	95.8%	36.2%	0.3649
M – M – F – F	68.0%	97.4%	29.4%	0.3291	73.9%	96.2%	22.3%	0.2552
F – F – M – M	68.6%	97.9%	29.3%	0.3282	80.4%	95.8%	15.4%	0.1713
F – F – F – F	66.0%	95.9%	29.9%	0.2916	83.7%	96.6%	12.8%	0.1809
Average (N_i>1)					66.3%	83.4%	17.0%	0.0975
Overall average					62.8%	76.9%	14.1%	0.0807

Note: M – F indicates families with an elder brother and a younger sister. The setup in the large α experiment (same as table 7): Families with a white widowed mother aged between 65 and 79, with an ADL index of 2.747, without a stroke, with poor eyesight, with some college education, and living in an owned house in a metropolitan area. All children are aged 47 and married and have a university degree and one child. Monte Carlo simulation with 20,000 random draws.

Table 9: Test of Joint Decision Model

Functional form	Solution	Log *L*	# Parameter	AIC
$\alpha=\rho=0$	Independent maximization	-7,205.55	21	14,453.1
Constant α and ρ	Joint maximization	-7,245.28	23	14,536.6
Constant α and ρ	Non-cooperative	-7,095.55	23	14,237.1
Full model	Joint maximization	-7,205.13	38	14,486.3
Full model	Non-cooperative	-7,066.48	38	14,209.0

Note: Based on 11,150 observations in 2010.

Table 10: Robustness of Externality: 1998 – 2010

	1998 [1]	2000 [2]	2002 [3]	2004 [4]	2006 [5]	2008 [6]	2010 [7]
	Coeff	Coeff	Coeff	Coeff	Coeff	Coeff	Coeff
P_father	-0.116***	-0.091*	-0.080*	-0.062	-0.101**	-0.075*	-0.073
P_mother	0.047	0.066**	0.077**	0.129***	0.112***	0.139***	0.098***
P_age5064	-0.070**	0.018	0.009	-0.029	0.017	-0.033	-0.035
P_age80p	0.066	0.080*	0.029	0.116***	0.082**	0.066	0.078*
P_white	-0.070*	-0.139***	-0.064*	-0.026	-0.055	-0.071*	-0.070*
P_ADL	-0.015***	-0.008	-0.019***	-0.016***	-0.021***	-0.014**	-0.015**
P_stroke	0.054	0.085*	0.021	0.130***	0.125***	0.102**	0.013
P_eye	-0.017	0.089***	0.053*	0.049	0.013	0.080**	0.042
P_college	-0.223***	-0.125***	-0.191***	-0.211***	-0.174***	-0.119***	-0.153***
P_nodegree	0.173***	0.133***	0.202***	0.203***	0.127***	0.089**	0.120***
Metro2	-0.026	-0.062*	-0.051	-0.021	-0.022	0.004	-0.027
Metro3	-0.106***	-0.138***	-0.124***	-0.086***	-0.076**	-0.105***	-0.047
House	0.090***	0.047	0.065**	0.118***	0.057*	0.058*	0.132***
C_age	-0.013***	-0.011***	-0.006***	-0.012***	-0.009***	-0.008***	-0.007***
C_male	-0.015	-0.000	-0.046**	-0.033	-0.074***	-0.087***	-0.041*
C_college	-0.352***	-0.345***	-0.350***	-0.329***	-0.358***	-0.362***	-0.349***
C_nodegree	-0.141***	-0.059	0.000	-0.027	-0.003	0.028	0.039
C_spouse	-0.327***	-0.393***	-0.323***	-0.327***	-0.333***	-0.320***	-0.381***
C_kids_partner	0.031***	0.022**	0.026***	0.026***	0.032***	0.036***	0.047***
C_kids_single	-0.014	-0.038***	-0.026*	-0.017	-0.050***	-0.051***	-0.017
Constant	0.646***	0.387***	0.347***	0.456***	0.468***	0.384***	0.333***
α (constant)	0.183***	0.172***	0.140***	0.164***	0.180***	0.175***	0.096***
ρ (constant)	0.346***	0.416***	0.361***	0.390***	0.382***	0.368***	0.324***
Observations	13,337	12,790	13,104	13,137	12,801	12,279	11,150
Log L	-8,601.07	-7,848.92	-8,356.87	-8,367.29	-8,160.60	-7,857.51	-7,095.55
% correct prediction							
All children	60.87%	65.11%	61.49%	61.32%	61.23%	61.11%	61.67%
All families	37.81%	43.08%	39.12%	38.79%	37.95%	37.56%	39.01%

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively. For 1998, a missing dummy for region is included.