Australian School of Business Research Paper No. 2013 ECON 06

Do Siblings Free-Ride in "Being There" for Parents?

Shiko Maruyama
Meliyanni Johar

Do Siblings Free-Ride in "Being There" for Parents?*

Shiko Maruyama[†]and Meliyanni Johar

March 21, 2013

Abstract

When siblings are concerned for the well-being of their elderly parents, the costs of caregiving and long-term commitment create a free-rider problem. If siblings living near their parents can share the costs, this positive externality exacerbates the under-provision of proximate living. Location decisions allow siblings to make a commitment to not provide long-term support for parents, and if decisions are made in birth order, elder siblings may enjoy the first-mover advantage. To quantify these effects, we study siblings' location decisions relative to parents by estimating a sequential participation game that features rich heterogeneity. We find moderate altruism and cooperation in the US that imply: (1) limited strategic behavior: more than 90% of children have a dominant strategy; and (2) non-negligible free-riding: of the families with multiple children, had siblings fully internalized externality and jointly maximized their utility, 18.3% more parents would have had at least one child nearby.

*Earlier versions of the paper were circulated under the title "Externality and Strategic Interaction in the Location Choice of Siblings under Altruism toward Parents." The authors gratefully acknowledge financial support from the Australian Research Council's *Discovery Projects* funding scheme (project number DP110100773).

[†]Corresponding author. E-mail: s.maruyama@unsw.edu.au

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. In a family with multiple adult children, altruism toward the elderly parent and the cost of caregiving result in a textbook public good problem. Children feel comfortable if they know their parents are well treated and well taken care of, and the cost of caregiving creates an incentive to free-ride on their siblings.

This public good problem is particularly highlighted when we consider siblings' location decisions. The opportunity cost of living near or with the parent and forgoing opportunities elsewhere is much less documented but no less important than other caregiving burdens. The discrete nature of location choice and associated non-negligible relocation costs make efficient bargaining challenging. Furthermore, in the course of location decisions, there exists a potential commitment device — birth order. The eldest children may enjoy the first-mover advantage by choosing to move away from their parents once they finish their schooling, before their younger siblings. Consistent with this argument, Konrad et al (2002) find that in Germany, elder siblings are more likely to move far away from their parents than their younger siblings.

The goal of this paper is to quantify this public good problem and the first-mover advantage for the first time in the literature, by studying the location decisions of adult siblings relative to their elderly parents. Based on American families in the Health and Retirement Study (HRS), we estimate a sequential participation game played by siblings.

In an attempt to quantify the free-rider problem, it is important to incorporate and distinguish another type of externality that is highly likely in reality — cooperation. Siblings living near parents may be able to share the costs, and shared caregiving is widely observed (Matthews and

Rosner 1988; Checkovich and Stern, 2002). While cooperation offsets the coordination difficulty caused by the free-rider problem, a positive externality resulting from cooperation exacerbates the under-provision of proximate living. A negative cooperation effect is also possible and implies an excessive provision of proximate living, similar to the standard entry game. An example is the bequest motive argument by Bernheim et al (1985), in which the presence of another sibling reduces transfers from the parent. When altruism toward parents is present, positive cooperation induces more families to be in the *prisoners' dilemma* situation, while negative cooperation leads to a larger first-mover advantage.

Our empirical framework addresses this potentially highly complex nature of the location game, relying upon the idea that all these externalities and strategic interactions can be summarized by three structural parameters: *altruism*, *private cost*, and *cooperation*. We then introduce rich heterogeneity in these three terms so that our empirical model can represent a wide range of participation games. As a result, our empirical framework sheds light on (1) the degree of externality, specifically altruism toward parents and cooperation among siblings, (2) the associated under-provision or over-provision of proximate living, (3) the game structure and equilibrium characteristics (e.g. coordination vs. anti-coordination games), (4) the extent of the first-mover advantage, and (5) how externality and inefficiency vary across families. Estimation relies on the maximum simulated likelihood. After the preference parameters of children are recovered, counterfactual simulations are conducted. We also estimate a joint-utility maximization model and a private-information model to confirm the validity of the main model.

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, around 80% of the hours of care are provided informally (OECD, 2005). Despite declining intergenerational coresidence and the increased 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, frequent visits, and mental and emotional support, 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.

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. All existing studies of empirical games in this literature assume the simultaneous move of siblings. Second, we are the first to apply an empirical game to the location decisions among siblings, rather than the informal care arrangement decision. The location choice problem more clearly highlights the free-rider problem and strategic interactions due to its discrete, irreversible, and long-term nature. Moreover, studying the location decision is important because the location pattern is a critical determinant of formal and informal care arrangements (Checkovich and Stern, 2002; and Engers and Stern, 2002).

This paper also makes two contributions to the literature of empirical games. First, our model features rich heterogeneity in the two distinct externalities. As a result, different players face participation games of different equilibrium characteristics (e.g. coordination and anti-coordination games). This is particularly important in the environments studied, because we expect a high level of heterogeneity in families' preferences and behaviors. This approach enables us to draw inferences on the share of families in the situation of prisoners' dilemma, families achieving joint-utility optimum, and families with a large first-mover advantage. Second, this paper is one of very few empirical analyses that allow us to study the first-mover advantage, preemption, and commitment device by incorporating sequential decision making. Most of the previous sequential games focus on extremely simple cases, such as two-player games. The only exception we are aware of is Schmidt-Dengler (2006), who studies the timing game of MRI adoption by hospitals in a fairly general setup and finds a significant but small preemption effect. Compared to his model, the rich heterogeneity in our model allows us to capture a wider variety of participation games.

The findings are summarized as follows. We find that the location game played by American siblings is characterized by altruism and cooperation of moderate size relative to private costs that have two implications. First, there is limited strategic behavior. More than 90% of children have a dominant strategy. The first-mover advantage is almost negligible because of the positive cooperation effect: among two-child families, reversing the birth order changes the location outcomes of only 1.9% of two-child families. 2.0% of multi-child families suffer from prisoners' dilemma. Second, altruism and positive cooperation lead to a non-negligible under-provision of proximate living due to free-riding. 28.8% of multi-child families result in location configurations that are not joint-utility optimal. Of the families with multiple children, had siblings fully internalized externality and jointly maximized utility sum, 18.3% more parents would have had at least one child living nearby. We also find substantial heterogeneity across families. Under-provision is more severe if children have a stronger altruism toward their parents, particularly in a family with a less educated single mother (widowed and non-widowed) with poor health and younger children. Lastly, we find that the non-cooperative model fits the data considerably better than the joint-utility maximization model. Our findings suggest that future research should direct its attention more toward externality, the free-rider problem, and the under-provision of care and attention rather than strategic interactions such as the first-mover advantage and prisoners' dilemma.

2 Related Literature

There are myriad economics and non-economics studies on intergenerational coresidence and colocation between elderly parents and their children (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), but few investigate the non-cooperative decision making of families, and none quantifies the free-rider problem among siblings.[1]

Regarding informal care arrangements, a small but tangible body of literature applies the non-cooperative game-theoretic framework to study interactions among siblings (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 or her own utility and the equilibrium arrangement is fully solved 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 using European multi-country data. Except for Byrne et al (2009), these studies find support for interdependence in caregiving

[1] 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.

decisions among siblings. Knoef and Kooreman (2011) argue 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 decision as given. We contribute to this empirical literature by being the first to apply the game-theoretic framework to the location decision among siblings.

Given the complexity of care and living arrangements, one model cannot capture all possible aspects of family decision making. The existing structural studies 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 from these relevant features to concentrate on modeling strategic interaction and externality. Our study should therefore be regarded as a complement to existing 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 first-born children are found to be more likely to live far 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 a younger sibling to locate closer to the parent as the primary caregiver. However, as we discuss below, the observed birth-order asymmetry may be explained by observed characteristics, and Konrad et al's (2002) nonstructural approach does not provide much understanding of the mechanism behind the observed asymmetry.

3 Data and Descriptive Results

3.1 Population

The data are drawn from the Health and Retirement Study (HRS), a nationally representative biannual longitudinal survey of Americans over 50. The HRS took its current form in 1998 by merging with the Study of Assets and Health Dynamics among the Oldest Old (AHEAD). Since then, the HRS has added two new cohorts, in 2004 and 2010. It 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.

To make the analysis tractable, we take a cross-sectional approach: we build a game-theoretic model to explain the cross-sectional variation of siblings' location patterns, abstracting from dynamic aspects other than sequential decision in birth order. We combine the three HRS waves in 1998, 2004, and 2010, and construct our "cross-sectional" data as follows. First, we choose family observations from HRS 1998 that meet the sample selection criteria explained below. Next, we add families from HRS 2004 that (1) meet the criteria and (2) are not yet included in our data. We then add the HRS 2010 observations, repeating the same procedure. A family thus never appears twice in our data set. We take this approach because pooling three waves considerably increases the sample size, and including waves that are twelve years apart ensures the robustness of our results. We also estimate our models using each wave separately. As reported in Johar and Maruyama (2012), our basic results are robust regarding the choice of waves.

Our sample consists of individuals over 50: (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 four children; (5) who have no step or foster children; (6) whose youngest child is aged 35 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, 3% of the elderly population live in nursing homes and fewer than 7% have no child. We limit the number of children to four to limit computational burdens. This group represents 75% of parents. For our research question, we expect to learn little from adding very large families.

We focus on relatively older children because the moves of younger children are often temporary; for example they may relate to post-graduate education. The location configuration of those above 35 is more likely to involve serious long-term commitment. We find that lowering this limit to age 30 does not affect our main results. We also set the maximum age of children because our model concerns where children set up their own families. We focus our study on biological children to the respondent parent, to avoid complications from potentially different family preferences in relation to non-biological children.[2] Finally, we exclude same age children because sequential decision making is one of our main interests and our estimation method utilizes birth order.

For this sample of parents, we create a child-level data set. The spousal information is retained as explanatory variables. Our final data consist of 18,647 child observations in 7,670 families, of which 55.0%, 24.9%, and 20.0% is from the HRS waves 1998, 2004, and 2010, respectively.

3.2 Location Patterns of Siblings

The location of the children relative to the parent defines our dependent variable. We group "living with the parent" and "living close to the parent" together and refer to this as living *near*

[2]Due to data limitations, our sample does not include parents who have a child whose relationship is biological only to the current spouse of the respondent.

the parent. While coresidence is becoming less common, shared caregiving is observed in a non-negligible proportion of families (Matthews and Rosner 1988; Checkovich and Stern, 2002). Siblings living nearby also contribute to the family by other means—by frequent visits and as a backup in case of caregiving burnout of the primary caregiver. 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).

Table 1 presents the location patterns of siblings in our data. The top panel shows that 48.7% of only children live far from their parents. Elderly parents with two children, shown in the second panel, are most likely to have one child nearby (43.1%) and least likely to have both of them nearby (17.4%). The probability of having no child nearby decreases with the number of children: parents of four children are much less likely to live without any child nearby (20.5%). We also compute the theoretical share of each location configuration under the assumption that each child makes decisions independently. This benchmark tells us what a standard probit model cannot capture and what our model needs to explain. The shares are computed based on p =40.4% – the share of children living near parents in our entire sample, and reported in the last column. Comparing the last two columns highlights two empirical regularities. First, only children are more likely to choose to live nearby, perhaps because they have no one to free-ride on. Second, in families with two or more children, siblings' decisions are positively correlated. This may be because they share similar preferences and environments or because of the positive coordination effect.

[Insert Table 1: Sibling Location Configurations by Birth Order]

Conditional on one child living near the parent, two-child families have two possible location configurations: (near, far) and (far, near), where the first element in parentheses denotes the first child's location. The two-child family panel shows that the former is less frequent than the latter. The three- and four-child family panels prove the robustness of this birth-order asymmetry, across the number of siblings and across the number of siblings living near the parent: in all rows with multiple possible location configurations, the rightmost cell has the largest share. This 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-documented fact that elder children tend to have more education.[3] Hence, how much of the observed birth-order asymmetry is attributed to the first-mover advantage is an empirical question.

3.3 Explanatory Variables

We use both parents' and children's characteristics. Parental characteristics include demographics (age, sex, marital status, and ethnicity), education, health status, location type (urban or rural), and housing. For children, we use age, sex, education, marital status, and information on grandchildren. Table 2 provides the definitions of the explanatory variables and their summary statistics. Parental health status is constructed as the first factor from factor analysis consisting of self-assessed health index, Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) scores, and previous diagnoses of diabetes, hypertension, and stroke. We choose these items because they tend to be persistent in time and occur in relatively many observations. We also attempt other health conditions in our specification, and find that they do not change the

[3] In our sample, the share of those who have a university degree is 37.0% and 35.6% for the first and second children in two-child families, and 36.2%, 32.0%, and 31.3% for the first, second, and third children in three-child families.

results. When a respondent parent is married, the data of the couple are aggregated.

Our assumption in using parents' residential location and housing status is that children's migration decisions determine the location configuration, not parents' migration. Although elderly parents sometimes relocate closer to their children, our calculation based on HRS 2010 reveals that more than 80% of new coresidence is formed by children moving in with the parent(s). Removing the location and housing variables does not affect our main results.

The majority of the parents in the sample own a house and are single, with widowed mothers being the most common. The majority of children are married. The mean ages of parents and children are 72 and 45, respectively.

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

4 The Model

4.1 Environment

We consider a game played by children. Our goal is to describe the observed cross-sectional snapshot of location configurations of families by explicitly modeling strategic interactions among siblings. Each child chooses whether to live close to their parent(s). To make our analysis tractable, we do not distinguish between 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 perfect-information sequential game in which each child sequentially makes a once-and-for-all location decision. 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 the interaction among siblings, whereas it does not assume that parents are passive and play no role. In reality, parents may influence children's payoff function by promising compensation for informal care in the future (e.g. a bequeathable house). Family bargaining and intergenerational transfers are implicit in our payoff function and our coefficient estimates should be interpreted in a reduced-form way.[4]

Second, modeling location choice as a once-and-for-all decision abstracts from the dynamic aspects of location choice except for the birth order sequence. Location choice dynamics caused by events and decisions in the later life, such as changes in the family structure, parents' health deterioration, and sibling bargaining regarding the informal care arrangement, is beyond the scope of this study, as it has been for most previous studies. In this sense, our utility function should be interpreted as indirect utility.

Third, we rely on the non-cooperative framework. An alternative is a model of joint-utility maximization, which we estimate and test against our non-cooperative framework. Fourth, ex-post bargaining and side payments *among* siblings are beyond the scope of our discrete setup. Large relocation costs justify this approach to some extent. Alternatively, our estimates of externality and strategic interaction can be regarded as their lower bound estimates, because in general, side-payments neutralize externality and strategic interaction.

Fifth, 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.[5] We also estimate an incomplete-information simultaneous-move game.

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

[5] 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.

4.2 Preferences

Denote the utility of child i by $u_{i,h}(a_{i,h}, a_{-i,h})$, where $a_{-i,h} \in \{0,1\}^{I_h - 1}$ is the choices of child i's siblings. In the rest of the paper, subscript $-i$ indicates a vector that contains the values of all siblings except for child i, and the family subscript, h, is omitted when no ambiguity arises. Given a_{-i}, child i's problem is, after dropping subscript h, written as

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

We further assume that child i's utility depends only on a_i and the number of siblings who choose to live near the parent irrespective of which siblings.[6] Let $N = \sum_k a_k$ denote the number of siblings who choose to live near the parent. The utility levels when child i lives far from the parent and near the parent are specified as follows:

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

Utility flow consists of three structural parameters, $u_i^{\alpha}(N), u_i^{\beta}$, and $u_i^{\gamma}(N)$. The first parameter, $u_i^{\alpha}(N)$, 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 has a child nearby. 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, intergenerational proximate living is a public good with a positive externality, and child i *free-rides* on child j if child i lives far and child j lives near the parent. $u_i^{\alpha}(N)$ may

[6] Relaxing this restriction is conceptually straightforward but computationally challenging.

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 parameter, u_i^{β}, captures child i's *private cost* (or benefit) from living near the parent that are independent of a_{-i}. For example, it captures caregiving burdens, opportunity costs, monetary transfers to/from parents, housing benefits in the case of coresidence, attachment to the location, and the consumption value of time child i shares with the parent.

The third parameter, $u_i^{\gamma}(N)$, is child i's private costs or benefits that depend on a_{-i}. This *cooperation* parameter 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 may be negative and decreasing in N. An 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. We normalize this term as $u_i^{\gamma}(1) = 0$ without loss of generality, that is, when child i is the only child near the parent, child i's utility is $u_i^{\alpha}(1) + u_i^{\beta}$.[7]

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

[7] Although we name the three parameters *altruism*, *private cost*, and *cooperation* and we attempt to interpret the results accordingly, such interpretation requires caution. For example, our altruism term, $u_i^{\alpha}(N)$, only captures the additively separable part as our "altruism". In other words, equation (1) is how we define the altruism in this study. As long as we interpret the results in such a way, how we label the three parameters does not undermine the generality of our behavioral model.

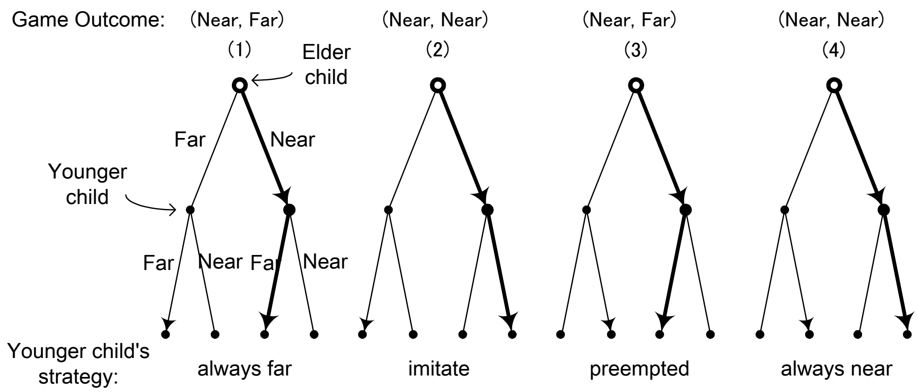

Figure 1: Strategies and Outcomes in Extensive-Form Presentation

from individually deviating from their equilibrium strategy *in every subgame*. Every finite game with perfect information has a pure-strategy SPNE (Zermelo's theorem).[8][9]

The sequential nature of the game is illustrated in the extensive-form representation in Figure 1. 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". Given the payoffs at each terminal node, we can find the SPNE outcome and strategies by solving the choice problem at each decision node sequentially from the youngest child to the eldest child (backward induction). Note that in Figure 1, if the first child lives nearby, two strategies of the second child, "always far" and "preempted", lead to the same game outcome — (Near, Far), because the difference between "always far" and "preempted" lies only in the unobservable off-the-equilibrium path. In estimation, we exploit this one-to-many mapping structure.

To examine the desirability of an equilibrium outcome, we use two efficiency measures: Pareto efficiency and efficiency in joint utility. Even if a game has a unique SPNE, it may have a Pareto-

[8] In this paper, we only consider pure strategies. In our perfect-information setup, mixed strategies are irrelevant because every decision node has a choice strictly better than the other.

[9] For Zermelo's theorem, see Mas-Colell et al. (1995), page 272.

improving (non-equilibrium) outcome, which constitutes well-known "prisoners' dilemma". Efficiency in joint utility, or Kaldor-Hicks efficiency, concerns the sum of siblings' utility. Though this criterion does not guarantee Pareto improvement, it is sensible to study this efficiency measure because it provides families or policy makers implications for implementable compensation schemes.[10]

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 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.[11] Example 2 exhibits the prisoners' dilemma. (Near, Near) is no longer an equilibrium, but it remains Pareto dominating and hence creates Pareto inefficiency in the equilibrium. 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 compensation is possible, these efficient outcomes will be chosen. Assuming a constant altruism (i.e. $u^{\alpha}(N=1) = u^{\alpha}(N=2)$), it is easy to show that the payoff matrix in Example 1 corresponds

[10] 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 study should be interpreted as such. If we assume that children's proximate living increases parents' utility, our inefficiency measures are the lower bound of family inefficiency.

[11] Though we do not discuss here, there is a normal-form representation of the sequential game. For example, the sequential game in Figure 1 is represented by a two by four payoff matrix.

to $\left(u_i^a, u_i^\beta, u_i^\gamma\right) = (1, -2, 3)$. 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. In this way, altruism, private cost, and cooperation in our model govern the game structure in each family.

A negative u_i^γ leads to an anti-coordination game, as is typical in entry games. Example 4 assumes $\left(u_i^a, u_i^\beta, u_i^\gamma\right) = (2, -1, -2)$, and has two Nash equilibria, (Near, Far) and (Far, Near). A smaller u_i^γ leads to a larger first-mover advantage. When sequence is introduced, the SPNE is (Far, Near), and child 1 enjoys a higher utility than child 2. Example 5 shows a rather rare but interesting case. Its normal form has a unique Nash equilibrium (Near, Far), in which child 1 plays a dominant strategy. However, the SPNE is (Far, Near), in which child 1 receives a higher utility by not playing the normal-form dominant strategy. The decision order provides child 1 with the "commitment device" and hence the first-mover advantage.

Example 4:

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

Example 5:

	$a_2 = 1$	$a_2 = 0$
$a_1 = 1$	$(0.5, 0.25)$	$(0.2, 0.26)$
$a_1 = 0$	$(0.4, 0.01)$	$(0, 0)$

4.4 Theoretical Predictions

Theoretical predictions can be derived and confirmed by a simulation study, which we have conducted in Johar and Maruyama (2012). The main results in symmetric two-player games are summarized as follows. First, joint-utility inefficiency increases with the *absolute* size of externality terms, u^α and u^γ. Both positive and negative values of u^γ enlarge inefficiency. The under-provision of proximate living results from positive values of u^α and u^γ, because children do not take into consideration the potential positive externalities to the other siblings. Similarly, if $u^\gamma < 0$, excessive participation may occur, creating a setting similar to the standard entry game.[12] When there is no externality ($u^{\alpha} = u^{\gamma} = 0$), the SPNE outcome maximizes the joint profit.

Second, the *prisoners' dilemma* case only appears when $u^{\alpha} > 0$, $u^{\gamma} > 0$, and $u^{\beta} < 0$, i.e. when cooperation increases payoffs but the incentive to free-ride exists. Its associated Pareto inefficiency increases as u^{α} and u^{γ} both become large.

Third, the size of the 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. Consider child 1's utility in a two-child family:

$$\begin{array}{c} u_1\left(a_1=1, a_2=1\right)=u_1^{\alpha}+u_1^{\beta}+u_1^{\gamma}, \quad u_1\left(a_1=1, a_2=0\right)=u_1^{\alpha}+u_1^{\beta}, \\ u_1\left(a_1=0, a_2=1\right)=u_1^{\alpha}, \quad u_1\left(a_1=0, a_2=0\right)=0. \end{array} \tag{2}$$

Then strategic substitutability in our two-player setup can be studied based on

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

Analogous to Gal-Or's (1985) argument, when the payoff function exhibits *decreasing difference* ($-u_1^{\alpha}+u_1^{\gamma}<0$), it implies strategic substitutability and we observe a larger first-mover advantage. If cooperation benefits siblings ($u^{\gamma} > 0$), it reduces the size of the first-mover advantage. *Strategic complements* (or a *supermoduler game*) may also result from a small u^{α} and/or large u^{γ}. In our symmetric binary setup, however, the second-mover advantage never appears, because strategic complementarity degenerates the game into the choice between (Near, Near) and (Far, Far) and

[12] Both positive u^{α} and negative u^{γ} create strategic substitutability, but the former leads to under-participation and the latter to excessive-participation. This makes our setting different from the standard entry game.

at the same time, the first mover is never worse off. Decreasing difference is also necessary for anti-coordination games such as Example 4 above.

In summary, if we find $u^{\alpha} > 0$ and $u^{\gamma} > 0$, it suggests: positive externality and free-riding among siblings; the under-provision of proximate living; possible prisoners' dilemma; and, if u^{γ} and u^{α} are of similar size, a small first-mover advantage. Finally, the extent of these externalities and distortions depends on the size of u^{α} and u^{γ} relative to the size of u^{β}. If the absolute value of u^{β} is dominantly large, the family is more likely to achieve the joint-utility optimal outcome.

5 Estimation

5.1 Unobserved Error Term

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

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

The unobserved error term is assumed to be distributed as a normal distribution independent of $\left(u_i^{\alpha}, u_i^{\beta}, u_i^{\gamma}\right)$. Under the assumption of perfect information, $\boldsymbol{\varepsilon}$ is unobservable to an econometrician but is observed by the siblings. The normality assumption implies that the game almost surely has a unique equilibrium, because ties occur with probability measure zero.[13] We can solve for the unique equilibrium by backward induction for any given $\left(u_i^{\alpha}, u_i^{\beta}, u_i^{\gamma}\right)$ and $\boldsymbol{\varepsilon}$.

As with the standard random utility models, the level of utility is not identified. Assuming the

[13] Here we use the term *almost surely* rather than *generically* because from the player's point of view, the payoff function is deterministic, unlike games for which game theorists use the term *generically*.

same variance for every child, 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 $\rho_{i,j} \in (-1,1)$, which we parameterize as

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

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

5.2 Specifying Functional Forms

For estimation, we also need to specify the functional forms of the three components, $u_i^{\alpha}(N), u_i^{\beta}$, and $u_i^{\gamma}(N)$. Let $X_i^{\alpha}, X_i^{\beta}$, and X_i^{γ} be vectors of covariates observable to the econometrician, each of which includes a constant term. In this paper, we report the results of the following four models. The simplest model, Model [1], imposes $u_i^{\alpha}(N) = u_i^{\gamma}(N) = 0, u_i^{\beta} = X_i^{\beta}\beta$, and $\rho_{i,j} = 0$. In this model, the decisions of siblings have no interdependency, and the econometric model degenerates to a standard binary probit model. Model [2] allows $\rho_{i,j}$ to be some constant, ρ_0, so that the preferences of siblings may correlate. Model [3] introduces externality in the most parsimonious way: $u_i^{\alpha}(N) = \alpha_0,\ u_i^{\beta} = X_i^{\beta}\beta,\ u_i^{\gamma}(N) = 0$, and $\rho_{i,j} = \rho_0$. Model [4] allows externality to vary in a flexible way so that the degree of externalities may depend on N and the characteristics of parents and children. Specifically:

$$\begin{aligned} u_i^{\alpha}(N) &= I[N \geq 1] \cdot \exp\{X_i^{\alpha}\alpha_0 + \alpha_1 \cdot I[N \geq 2] + \alpha_2 \cdot I[N \geq 3]\}, \\ u_i^{\beta} &= X_i^{\beta}\beta, \quad \text{and} \\ u_i^{\gamma}(N) &= X_i^{\gamma}\gamma_0 \cdot (I[N \geq 2] + \gamma_1 \cdot (N-2) \cdot I[N \geq 3]), \end{aligned} \tag{6}$$

where α_1, α_2, and γ_1 are scalar parameters, and α_0, β, and γ_0 are vectors of coefficient parameters, which allow preference heterogeneity based on observables. In our behavioral model, a negative value of $u_i^{\alpha}(N)$ has no meaningful interpretation. After we estimate Model [3] and confirm a positive estimate of α_0, we introduce heterogeneity in this term using the exponential function so that its value is always positive. As discussed below, we have attempted many alternative specifications to (6), and the main results are found to be robust.

5.3 Identification

To understand how our structural parameters are identified, take a simple model of two-child families with no heterogeneity as an example: $\left(u_i^{\alpha}(N),\ u_i^{\beta}, u_i^{\gamma}(N)\right) = (\alpha_0, \beta_0, \gamma_0)$ and $\rho_{i,j} = 0$. First, consider the choice problem of child 2 after he observes that child 1 chooses to live near the parent. This binary choice problem compares $u_2(a_2 = 1, a_1 = 1) = \alpha_0 + \beta_0 + \gamma_0 + \varepsilon_2$ and $u_2(a_2 = 0, a_1 = 1) = \alpha_0$, and thus allows us to identify $\beta_0 + \gamma_0$. Similarly, when child 1 chooses to live far, we identify $\alpha_0 + \beta_0$. From these two values, the degree of strategic substitutability, $\alpha_0 - \gamma_0$, (and hence the degree of the first-mover advantage) is determined. When we assume no cooperation effect (i.e. $\gamma_0 = 0$), then the identification of α_0 and β_0 follows. Otherwise, three parameter values are indeterminate.

The rest of identification relies on sequential interaction. To illustrate this point, consider the following two families: (1) free-riding siblings, $(\alpha_0, \beta_0, \gamma_0) = (2, -2, 0)$ and (2) siblings hating each other: $(\alpha_0, \beta_0, \gamma_0) = (0, 0, -2)$. In both cases, $\alpha_0 + \beta_0 = 0$, $\beta_0 + \gamma_0 = -2$, and $\alpha_0 - \gamma_0 = 2$, thus studying the choice problem of child 2 cannot distinguish between these two types. Under the payoff function with decreasing difference, child 2 has three strategies depending on the value of ε_2: "always far", "preempted", and "always near", as shown in Figure 1. The last step of identification is achieved by studying child 1's choice problem when child 2 takes the strategy "preempted", that is, comparing $u_1 (a_1 = 1, a_2 = 0) = \alpha_0 + \beta_0 + \varepsilon_1$ and $u_1 (a_1 = 0, a_2 = 1) = \alpha_0$, and thus identifying β_0. If we observe that child 1 almost always chooses to live far when child 2 takes the "preempted" strategy, it implies a larger α_0 and a smaller β_0, i.e. siblings with free-riding. In the second type families, we will observe child 1 choosing "near" and "far" with the same probability. Put differently, the size of the birth-order asymmetry *given the size of the first-mover advantage* provides essential information in identifying the three parameters separately.

As shown here, the identification of this simple game is achieved solely by studying the two-child families, although the use of functional form assumptions and information from families with different numbers of siblings further assists the identification. Including three- and four-child families also allows for the identification of models with richer heterogeneity.

5.4 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 $a_h^o \in \{0, 1\}^{I_h}$.

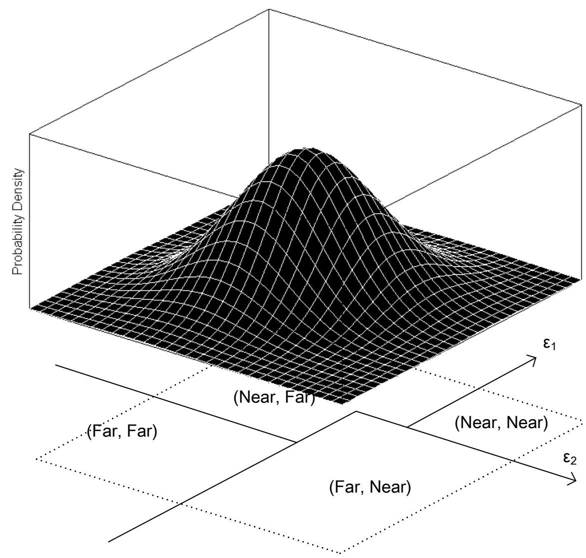

Figure 2: Relationship between Error Terms and Location Outcome

The log-likelihood function is written as

$$\widehat{\theta}_{ML} = \arg\max_{\theta} \left\{ \frac{1}{H} \sum_{h}^{H} \ln \Pr_{\rho} \left[a_h^o = a_h^*(\mathbf{X}_h, \boldsymbol{\varepsilon}_h;\ \alpha, \beta, \gamma) \right] \right\}, \tag{7}$$

where θ is the vector of the model parameters, $(\alpha, \beta, \gamma, \rho)$, and $\mathbf{X}$ is the union of $X_i^{\alpha}, X_i^{\beta}$, and X_i^{γ}. The intuition behind this likelihood function is that, conditional on $\mathbf{X}$ and (α, β, γ), the location configuration is determined by $\boldsymbol{\varepsilon}$, and the probability of a location configuration is computed based on the distribution of $\boldsymbol{\varepsilon}$. Figure 2 depicts this relationship in a two-child family example. The asymmetry around the center part in Figure 2 is due to sequential strategic interaction.

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 correlate among siblings and externality makes the decisions of siblings interdependent. When the dimension of ε_h becomes large (i.e. more than two), computationally demanding numerical approximation, such as the quadrature method, is impractical. To overcome this computation problem, we use the Monte Carlo integration method developed by Maruyama (2010).

5.5 Monte Carlo Integration

Maruyama (2010) develops the Monte Carlo integration method applicable to finite 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. First, the estimation relies on MSL assisted by the Geweke-Hajivassiliou-Keane (GHK) simulator, the most popular solution for approximating 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 SPNE profiles that rationalize the observed equilibrium outcome. In the sequential game framework, the econometrician does not observe the underlying SPNE, because an equilibrium strategy consists of a complete contingent plan, which includes off-the-equilibrium-path strategies as unobserved counterfactuals. Hence, there may exist different realizations of unobservables that lead to different subgame-perfect equilibria but generate an observationally equivalent game outcome. Figure 3 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.

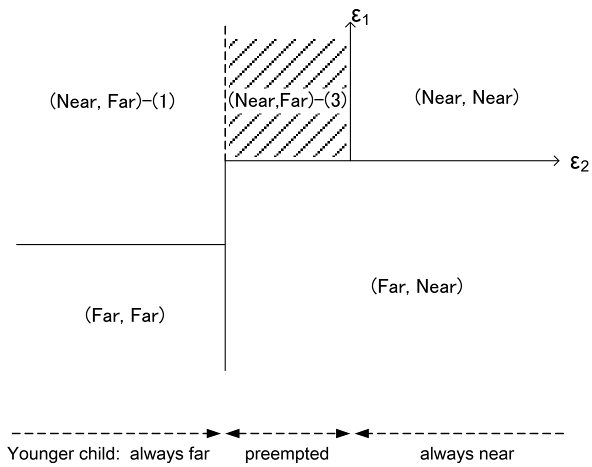

Figure 3: Dividing Observed Location Outcome into Strategy Profiles

The use of subgame perfection resolves this non-rectangular domain problem. In the example in Figure 3, the non-rectangular integration domain for (Near, Far) consists of two rectangular regions that correspond to two sets of SPNE, labeled (1) and (3), which correspond to (1) and (3) in the extensive form in Figure 1. Maruyama (2010) proves that the separate evaluation of the 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. After computing the probability of each SPNE that rationalizes the observed outcome, the econometrician then obtains the probability of the observed outcome by summing the probabilities of those SPNEs, and the use of maximum likelihood follows.

6 Results

6.1 Probit Results

Before presenting advanced models, it is useful to summarize the results from a simple probit model, which serves as a benchmark for complex models. In addition to its reduced-form interpretation, the model offers a simple random-utility-model interpretation under the following assumptions: each child makes their location decision independently; their decision has no implications for other children; and their unobserved preference component is distributed i.i.d. normal.

The results are reported in Column [1] of Table 3. Parents who have a child living nearby tend to be older, less educated, and less healthy widowed parents who live in their owned home in urban areas. Proximate living is less likely for white parents and single but non-widowed fathers. Child variables are also relevant. Unmarried children, particularly single daughters, are likely to live near their parents. Married children are less likely to do so, but the presence of their children slightly offsets this effect (grandparenting effect). Education moves children away from the parents. These findings are consistent with previous studies (e.g. Checkovich and Stern, 2002; Compton and Pollak, 2009; Byrne et al, 2009). Proximate living is less likely for children who are older after controlling for parental age. This partly captures the higher tendency of elder siblings to live far from the parents.

[Insert Table 3: Estimated Parameters]

6.2 Models with Interactions among Siblings

The first step to build interdependence among siblings is to model a correlation in the error terms, $\{\varepsilon_{i,h}\}_{i=1}^{I_h}$. Model [2] has a covariance matrix, Ω^h, whose off-diagonal elements are all equal to a constant, $\rho_0 \in (-1, 1)$, which captures resemblance in the preferences of siblings, shared environments, and a certain behavioral interaction among siblings. The result shown in Column [2] of Table 3 testifies a significant positive correlation in the error term.

Now we explicitly introduce externality, first by a constant externality, $u_i^{\alpha}(N) = \alpha_0$. As shown in Column [3] of Table 3, we find a positive and significant estimate of α_0.[14] To confirm the robustness of this result, we estimate Model [3] separately using each wave of HRS from 1998 to 2010 and we find α and ρ always positive and highly significant.

When we compare Models [1] - [3], while there is no substantial change in coefficient estimates, the goodness-of-fit improves over every step of elaboration. In terms of $\log L$, a decent improvement comes from incorporating externality α, but incorporating correlation ρ contributes most. The proportion of correctly predicted observations, which is defined based on the location configuration with the highest predicted probability, also shows improvement in the model fit. Although the proportion of correct prediction becomes slightly worse at the individual child level, a significant improvement is found at the family level.

6.3 Model with Heterogeneous Externality

Model [3] assumes a restrictive form of externality — a constant altruism toward parents. To allow for another form of externality, cooperation , as well as heterogeneity in externality, we now parameterize $u_i^{\alpha}, u_i^{\gamma}$, and $\rho_{i,j}$ as specified in (5) and (6), by introducing covariates in each term. Given the positive significant estimate of α_0 in Model [3], it is reasonable to parameterize u_i^{α} in the exponential function so that u_i^{α} is always positive. Including the full set of covariates in every term

[14] This model leads to an even larger ρ for the following reason. The positive externality creates strategic substitutability in the decisions of siblings. Without explicitly modeling this externality, the correlation in the error term needs to reflect this negative behavioral correlation, resulting in a smaller estimate of ρ in the previous specification.

is impractical, because it makes precise identification of parameters difficult and increases computational burden. We thus need a reasonably general yet parsimonious specification. Admittedly, the choice of variables is arbitrary. Two guidelines lead us to our final model. First is the intended behavioral interpretation of each term. Variables in u_i^α are supposed to be the determinants of innate altruism, while variables in u_i^γ should affect the cost and benefit of cooperation. Second, we adjust the sets of variables by attempting various specifications. Variables that are always estimated with a large standard error and/or without statistical and economic significance are not included. We find our main results are reasonably robust across these modifications. Regarding correlation within siblings, we allow $\rho_{i,j}$ to depend on the age and gender differences between children i and j.

Column [4] of Table 3 reports the result of the full model. Compared to Model [3], the model fit is improved both in terms of log likelihood and correct prediction, indicating the importance of heterogeneity in externality. The LR test confirms that the improvement is significant at standard significance levels. Figure 4 compares the predicted distribution of location configurations of Models [1] - [4] with the actual distribution in data, illustrating step-by-step improvement in model prediction.

Correlation in the error term is stronger for siblings of closer age and of the same sex than for other siblings, indicating that those siblings share preferences and environments to a greater extent. The altruism parameter, u_i^α, varies across children and families. Altruism is strongest toward less educated single mothers (widowed or not) with poor health. α_1 and α_2 measure utility components in the u_i^α term that increase with the number of children living near the parent, but their estimates are small and insignificant. This implies that what is important to children is whether at least one child lives near the parent. Based on the distribution of X_i^α, the value

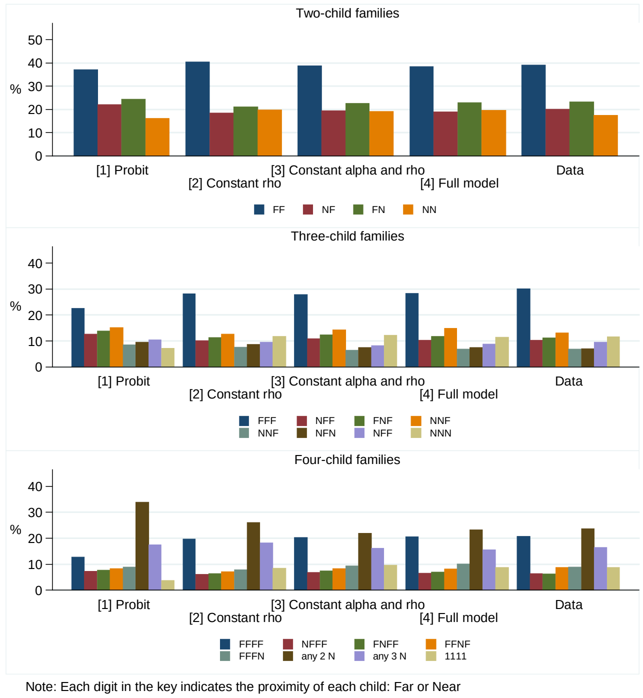

Figure 4: Predicted and Observed Location Configurations

of u_i^{α} ranges $[0.120, 1.370]$ with its mean 0.377.[15] The cooperation parameter, u_i^{γ}, also exhibits heterogeneity, ranging $[-0.046, 0.361]$ with its mean 0.199. The size of u_i^{γ} is overall smaller than u_i^{α}. The ranges of u_i^{α} and u_i^{γ} also indicate that a few families effectively show no altruism and cooperation. u_i^{γ} is larger for younger children. One interpretation of this heterogeneity is that younger siblings enjoy living close to each other. This interpretation has little to do with the provision of care and attention. Alternatively, younger siblings tend to have less experience of care provision and hence mutual assistance reduces the (actual and perceived) cost of providing care and attention. Similar to α_1 and α_2, the estimate of γ_1 indicates that having the third sibling nearby has no significant effect on u_i^{γ}. Thus, externality and inefficiency do not distort the behavior of

[15] We also estimate a model with a linearly parameterized u_i^{α}, instead of in the exponential function. Its results imply that u_i^{α} sometimes takes a small negative value, though the vast majority of children have a positive u_i^{α}. We find no substantial difference between these two models in model fit and main findings.

families in which more than two siblings choose to live near parents. We attempt a number of alternative functional specifications of u_i^α and u_i^γ other than (6), and we consistently find our main results robust.

Heterogeneity in u_i^α and u_i^γ determines the extent of inefficiency and strategic interaction in each family. Inefficiency becomes larger in a family with larger u_i^α and u_i^γ, i.e. a family with a less educated single mother with poor health and relatively younger children. The prisoners' dilemma is more likely in these families. The first-mover advantage, on the other hand, becomes larger when u_i^α and is large and u_i^γ is small. For relatively older children, the value of cooperation is limited, and if their parent is a less educated non-widowed single mother with poor health, the incentive of free-riding is large, and the elder child has a large first-mover advantage.

The range of $X_i^\beta\beta$ is $[-2.193, 0.861]$ with its mean -0.545.[16] Given that the variance of ε_i is unity, the ranges of the three preference parameters suggest that although the two externalities are not negligible, the private costs, $u_i^\beta + \varepsilon_i$, is the primary determinant of the location patterns. The coefficients in the u_i^β term of Model [4] are estimated less precisely than previous models, but the sign and magnitude of each coefficient are largely similar to the previous models. Exceptions are the variables also included in u_i^α and u_i^γ. These variables offer additional insight. While Models [1]-[3] find that parents of poor health are more likely to have their children nearby, in Model [4], this effect in u_i^β becomes smaller and we find that poor health significantly increases u_i^α. This implies that poor parental health induces intergenerational proximity both (1) because poor parental health increases children's net utility of living near the parent and (2) because children are more concerned about the well-being of those parents. Children's utility from living near the parent increases with parental poor health despite the expected larger cost of care provision, probably because children

[16] The distributions of u_i^α and u_i^β are skewed because of the skewed distribution of parental health.

value sharing the time with parents with shorter life expectancy.

Children's education shows another contrast. Previous models reveal 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_i^{β}, most likely reflecting high opportunity costs to educated children of living near the parent. The estimated coefficients in u_i^{α} show no evidence that educated children are less concerned about the well-being of their parents than less educated children.

Lastly, why do elder siblings tend to live far from the parent? Our coefficient estimates offer some explanation. We find a negative age effect both in u_i^{β} and u_i^{γ}; the private cost of living near the parent increases with age, and an additional sibling near the parent benefits elderly siblings less. Both of these effects contribute to the lower tendency of elder siblings to live near their parents, and these effects have nothing to do with the first-mover advantage; elder siblings live far because they are older. At the same time, the significant estimates of u_i^{α} and u_i^{γ} indicate the existence of the sequential strategic interaction. In the next section, we quantify how much of the birth-order asymmetry in location choice can be explained by the first-mover advantage.

7 Counterfactual Simulations

7.1 Method

The estimated parameters reveal how $u_i^{\alpha}, u_i^{\beta}, u_i^{\gamma}$, and $\rho_{i,j}$ vary across families and siblings, and guide us to counterfactual simulations to quantitatively illustrate how the game structure and game outcomes vary across families under different game settings, such as a simultaneous game and joint-utility optimization.

In the counterfactual exercises, we simulate location configurations under certain assumptions, based on estimated parameters $\widehat{\theta}$ and data, $\left\{a_{i,h}^{o}, X_{i,h}\right\}_{i=1}^{I_h}$. This simulation is not straightforward for several reasons. First, if we knew the true values of $\varepsilon_{i,h}$, solving for the equilibrium or optimal location configurations would be trivial, but we do not observe $\varepsilon_{i,h}$ in data. We thus rely on Monte Carlo simulations, in which we generate the simulated values of error terms that rationalize the observed location configuration. For example, we can compute the probability that the siblings in family h result in location configuration $\widetilde{\mathbf{a}}_h$ 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 $\Delta\left(\mathbf{a}_h^o\right)$,

$$\Pr\left(\widetilde{\mathbf{a}}_h\right) = \frac{1}{\Pr\left(\boldsymbol{\varepsilon}_h \in \Delta\left(\mathbf{a}_h^o\right)\right)} \int_{\boldsymbol{\varepsilon}_h \in \Delta\left(\mathbf{a}_h^o\right)} I\left[\widetilde{\mathbf{a}}_h = \mathbf{a}_h^*\left(\mathbf{X}_h, \boldsymbol{\varepsilon}_h\right)\right] \phi\left(\boldsymbol{\varepsilon}_h\right) d\boldsymbol{\varepsilon}_h,$$

where $\phi\left(\boldsymbol{\varepsilon}_h\right)$ is the density function of $\boldsymbol{\varepsilon}_h$, and $\mathbf{a}_h^*\left(\mathbf{X}_h, \boldsymbol{\varepsilon}_h\right)$ is a function for the solution under a particular setup. 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 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.

7.2 Joint-Utility Optimal Location Configuration

We first consider joint-utility inefficiency. Table 4 shows the relationship between the observed SPNE location patterns and the family-optimal location patterns in three-child families for illustration purposes (similar patterns are observed for two- and four-child families). If having no child near the parent leads to the highest joint utility (the first column), it always occurs as an SPNE outcome. In this case, the positive externality is so small compared to large private costs that externality plays no role. If the joint-utility optimum is achieved by one or two children living nearby, the SPNE outcome may or may not lead to the same location pattern, and if the SPNE location pattern is different from the joint-utility optimal one, it almost always results in the under-provision of proximate living, though over-provision is possible when u_i^α is close to zero and $u_i^\gamma < 0$. The last column shows that when having three children nearby is joint-utility optimal, there is no distortion. This finding comes from the negative estimates of α_2 and γ_1.

[Insert Table 4: Observed and Family-Optimal Location Configurations in Three-Child Families]

Table 5 presents this joint-utility inefficiency by family size. Many families whose optimal number of children living near the parent is one or more in reality have no child living nearby. This gap between the SPNE and the joint-utility optimum increases with family size, because positive externality is shared by more children. The last row in the table shows that, of the families with multiple children, had siblings fully internalized externality and collectively maximized utility sum, 18.3% more parents (= 32.5% − 14.2%) would have had at least one child nearby.[17]

[Insert Table 5: Observed and Family-Optimal Location Configurations by Family Size]

7.3 Normal-Form Game Structure

Payoff matrices, or the normal-form representations of games, provide useful information to understand the nature of the games played by American siblings. Table 6 characterizes (simultaneous)

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

payoff matrices of two-child families by the observed SPNE location pattern. The top panel in Table 6 reports whether siblings have dominant strategies in the payoff matrix. In 86.2% of two-child families, both children have a dominant strategy. This reflects that for the majority of children, the size of u_i^β is so large that u_i^α and u_i^γ do not influence their decisions. It is trivial to show that, when every child has a dominant strategy, the simultaneous game has a unique equilibrium, and the equilibrium outcome is always achieved as an SPNE. Hence, the fact that the vast majority of children have a dominant strategy suggests limited strategic behavior. The table also shows that when we observe (Far, Far) or (Near, Near) in data, it almost always implies that both children in those families have a dominant strategy. The last column of Table 6 reports a simulation in which we double u_i^α for every family. The share of families in which both children have a dominant strategy reduces to 62.7%. A larger externality induces strategic behavior to a greater extent.

[Insert Table 6: Characteristics of Simultaneous Normal-Form Games in Two-Child Families]

The bottom panel in Table 6 characterizes the Nash equilibrium of the simultaneous game. This table shows limited strategic behavior even more clearly. More than 99% of two-child families have a unique simultaneous equilibrium and it is rare to have no equilibrium or multiple equilibria. In most cases, the unique equilibrium in the simultaneous game actually occurs as an SPNE outcome. The only non-negligible gap between the normal-form equilibrium outcome and the SPNE outcome is found among the families that choose (Far, Near). This group includes not only families whose normal-form equilibrium is (Far, Near) but also families whose normal-form equilibrium is (Near, Far) and families with two equilibria that consist of (Far, Near) and (Near, Far). This gap suggests the presence of the first-mover advantage.

7.4 Efficiency Type

The observed SPNE location patterns can be classified into three groups: (1) joint-utility optimal; (2) joint-utility suboptimal but Pareto efficient; and (3) prisoners' dilemma, that is, there is a non-SPNE location configuration that is Pareto-dominating. Table 7 presents the distribution of these three groups by family size. It also shows the distributions under different externality parameter values (Panels [2]-[4]). Panel [1] shows that the prisoners' dilemma is observed only for 2.0% of the multi-child families, but that its presence increases with family size: 2.7% of four-child families suffer from prisoners' dilemma. While 98.0% of multi-child families achieve Pareto efficiency, many of them do not achieve the joint-utility optimum. This joint-utility inefficiency is particularly large in three- and four-child families: only 65.6% of them achieve the joint-utility optimum. The simulation results reported in Panels [2]-[4] confirm the theoretical predictions. Larger u_i^α and u_i^γ lead to larger joint-utility inefficiency and Pareto inefficiency. u_i^α explains a larger part of joint-utility inefficiency than u_i^γ, while a large value of u_i^γ is necessary for prisoners' dilemma to occur.

[Insert Table 7: Efficiency Type by Family Size]

Table 8 compares the efficiency types by the observed location configuration (the top panel) and by the joint-utility optimal location configuration (the bottom panel) in two-child families. The first number in each cell represents its column share and the second number its row share. The table illustrates how prisoners' dilemma occurs. 70.5% of families in the prisoners' dilemma situation have no one near the parent despite the fact that (Near, Near) is Pareto dominating. The remaining 29.5% have the second child near the parent, although (Near, Far) is Pareto dominating. Joint-utility inefficiency occurs in a similar way. As discussed earlier, when (Far, Far) is joint-utility optimal, a family can achieve it as an SPNE outcome. When we observe (Near, Near) in data, it is always joint-utility optimum, while when we observe (Far, Far), it is joint-utility efficient only for 61.5% of those families.

[Insert Table 8: Location Patterns and Efficiency Type in Two-Child Families]

7.5 First-Mover Advantage

One way to quantify the first-mover advantage in our setup is to study the equilibrium outcome that arises when we reverse the decision order. Table 9 reports the results of the reverse-order simulations for two-child families. The top panel compares the location configurations in the observed SPNE and in the reverse-order SPNE. The bottom panel investigates how reversing the order alters the utility of each child. Overall, the table clearly states that the sequential interaction is negligible. The reverse order only affects 1.9% of two-child families. When it does affect a family, it is almost always the case that the reverse order changes the SPNE outcome from (Far, Near) to (Near, Far), increasing the second child's utility and decreasing the first child's utility. The joint utility may or may not increase. If we double altruism, the share of families with the first-mover advantage increases to 9.3%.

Konrad et al (2002) argue that the birth-order asymmetry in the location configuration supports the first-mover advantage hypothesis. An interesting question is how much of the birth-order asymmetry in the location configuration can be explained by the first-mover advantage. The numbers of two-child families that result in (Far, Near) and (Near, Far) are 658 and 564, respectively. Of the 658 families with (Far, Near), 7.8%, or 51 families are affected by the reversed order. If we assume that removing the first-mover advantage affects the half of 51 families switching their location configuration from (Far, Near) to (Near, Far), we would observe 632 and 590 for (Far, Near) and (Near, Far) without the first-mover advantage. Hence, even though the first-mover advantage implied by our estimates is small, because the birth-order asymmetry in data is also small, the first-mover advantage explains roughly half of the asymmetry. The rest of the asymmetry is explained by observable characteristics.

[Insert Table 9: Reverse-Order SPNE in Two-Child Families]

8 Robustness and Validity of Results

8.1 Sensitivity Check

We have attempted various population selection criteria and functional forms, and the main findings are fairly robust. In this subsection, we discuss selected robustness tests that are critical to the interpretation of our results. The detailed results of these tests are reported in the Appendix.

Measuring Decision Order If the decision order we impose in estimation contains measurement error, estimated strategic effect may be biased toward zero. We expect that birth order is recorded with little measurement error, but birth order may not necessarily coincide with the actual order of location decisions among siblings. There may be a number of temporary moves when siblings are in their twenties, such as attending college elsewhere, and some of those moves may become permanent. It is not uncommon that younger siblings make a permanent move before their elder siblings complete their post-graduate education. The question then is how well birth order approximates the true decision order. The degree of measurement error in decision order determines the size of bias. Maruyama (2010) conducts a Monte Carlo experiment by applying the same estimation method as this study for a sequential entry game, and reports that such bias tends to be marginal if decision order is correctly specified in more than 90% of game observations.

One way to investigate potential bias resulting from misspecifed order is to estimate the same models excluding siblings of similar age. In this way, birth order reflects the true decision order more accurately and the strategic effect will be estimated more precisely. Specifically, we exclude families that have a pair of siblings whose age difference is only one year and estimate the same models. With this additional restriction, most of our results are unaffected,[18] and we find the same when we further increase the minimum age difference to three years. Thus, measurement error in decision order is unlikely to cause significant bias in our findings. We exclude only same age children in the full model, because this approach provides significantly more observations and it guarantees that our results are largely nationally representative.

Conceptually, decision order in our model is a broader notion than a mere timing of migration, involving any credible commitment related to a permanent move, such as the choice of occupation and spouse, that determines the possibility of the child moving away from the parent in the future. Hence, although it is not uncommon for younger siblings to make a permanent move before elder siblings complete a post-graduate degree, it does not necessarily contradict the use of birth order.

Are Only Children Special? We include one-child families in our sample because they aid identification. However, the results may be biased if only children considerably differ from children with siblings (after controlling for observable characteristics). To address this concern, we estimate our models without one-child families. We find that excluding one-child families makes parameter estimates less robust regarding specification choice. Standard errors tend to be larger with slightly

[18] We find slightly larger estimates of externality as well as a smaller proportion of families experiencing inefficient family location both in the sense of Kaldor-Hicks and Pareto. These two findings arise at the same time because while the sample of siblings with a larger age gap leads to a larger estimate of strategic effect, such siblings tend to have more diverse characteristics than siblings of similar age. When players differ to a greater extent, they are more likely to have a dominant strategy and game outcomes depend less on strategic interaction; hence we find smaller inefficiency.

worse goodness-of-fit measures. However, while these findings suggest an important role for one-child families in estimation, the results do not show any distinct systematic difference and the results are overall consistent with our main results, indicating that our results are highly unlikely to be artifact generated by the distinct nature of only children.

Potential Bias Due to the Cross-Sectional Approach To quantify sequential strategic interaction in a tractable yet intuitive manner, this study takes the cross-sectional approach that bases our empirical analysis on cross-sectional variation of data and abstracts from dynamic aspects of siblings' location decisions except for birth order. For our estimates to be meaningful and credible, our empirical framework must be approximately consistent with the underlying data generating process in reality. In particular, the explanatory variables used in estimation are taken from information recorded many years after children make their location decisions. The results can be interpreted in a way that is consistent with our behavioral model if all our explanatory variables are either observed or accurately predicted at the time children make decisions. For this reason, we select exogenous variables that we expect to be either time-invariant or reasonably stable and predictable in the long run. Rigorously speaking, however, our structural parameters may be biased because many of our explanatory variables are time-varying for a number of reasons. For example, a child's location decision might have a long-term effect on our explanatory variables such as parental health (reverse causality). Location and spouse might be determined at the same time (simultaneity). A child might have responded to recent parental health decline many years after the child has left the parent (misspecification of the time frame), and current variables might have accumulated stochastic errors since the child makes the decision and thus they may lead to downward bias even if the child's prediction is not biased (measurement error).

To address these concerns, we estimate a simplified model that excludes health and marital variables, which may be endogenous events in later life. We find that the results of the simplified model are overall consistent with the full model, despite its significantly worsened model fit. Counterfactual simulation results are also similar between the two models. This finding provides some assurance that our main findings are not driven by the time inconsistency due to those time-variant variables.[19]

8.2 Alternative Behavioral Assumptions

While our key findings are robust regarding the data selection criteria and functional form assumptions, our discussion so far is based on the assumption of the perfect information sequential game. To investigate how appropriate this behavioral assumption is, we discuss three alternative models.

Collective Maximization First, we examine the assumption of the non-cooperative decision making. This assumption is to some extent justified by the discrete, irreversible, and long-term nature of the location choice, 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 identical functional form specification as our full model, namely, (3), (4), and (6), and assumes the following joint-utility maximization:

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

[19] A more conservative view for our approach is that even if our cross-sectional approach does not lead to precise estimates, it is an empirical model exercise that focuses not on the precision of estimates but on finding models with new features that better fit the data. It is not uncommon that empirical game-theoretic analysis of an inherently dynamic subject starts with the cross-sectional framework. The structural econometric modeling of firms' entry, for example, starts with the framework to study a cross-sectional snapshot of market structures.

We estimate this model by using the multinomial probit framework. Because the multivariate normal distribution does not have an analytical form, the estimation is based on the method of simulated likelihood with the GHK simulator.

Incomplete-Information Game To examine how crucial the perfect-information assumption is, we estimate an incomplete-information model, maintaining the same functional form specifications as before. In this setup, each child makes a decision simultaneously by maximizing expected utility based on the privately observed value of ε_i, the distribution of ε_{-i} (conditional on ε_i), and "conjectures" of the other siblings' strategies. Conjectures about the other siblings' behaviors underlie the utility maximization because they affects one's expected utility. Child i's strategy, or decision rule, is denoted as $a_i(\varepsilon_i)$, and effectively it is a threshold value of ε_i above which child i chooses "near" or $a_i = 1$. A strategy profile in family h, $\{a_i(\varepsilon_i)\}_{i=1,\ldots,I^h}$, constitutes a Nash equilibrium if:

$$a_i^e(\varepsilon_i) = \arg\max_{a\in\{0,1\}} E_{\varepsilon_{-i}}\left[u_i\left(a, \{a_k^e(\varepsilon_k)\}_{k\neq i}, \varepsilon_i\right)\right] \text{ , for } i = 1, \ldots, I^h. \tag{8}$$

We estimate this multivariate probit model by the method of simulated likelihood.

The procedure for constructing the simulated likelihood consists of three key algorithms. The first is an algorithm to obtain the optimal strategy of child i, $a_i^*(\varepsilon_i)$, given the strategies of the siblings, $\{a_k(\varepsilon_k)\}_{k\neq i}$, by evaluating the net expected utility gain of choosing "near". For incomplete-information games, previous studies typically assume a normal or logit distribution independent across players, but our error terms are correlated within siblings, and hence, the optimal strategy, $a_i^*(\varepsilon_i)$, needs to be obtained from a conditional normal distribution that incorporates the correlation parameters. When child i has more than one sibling, the expectation is evaluated numerically by the GHK probit simulator.[20] The second algorithm obtains the equilibrium strategy profile, $\{a_i^e(\varepsilon_i)\}_{i=1,\ldots,I^h} \equiv \mathbf{a}_h^e$. This algorithm consists of a numerical iteration loop that nests the first algorithm inside, and solves the equilibrium strategy profile as a fixed point in (8).[21] We find that this numerical iteration procedure is well-behaved as long as parameter values are not far from reasonable values. Because the mapping defined by (8), $f : \mathbf{a}^t \to \mathbf{a}^{t+1}$, is a continuous mapping from $\mathbb{R}^I$ to $\mathbb{R}^I$, the existence of a fixed point is guaranteed by Brouwer's fixed point theorem. Although the uniqueness of the equilibrium depends on model parameters, as long as f is decreasing or moderately increasing (derivatives less than one) at any point of $\mathbb{R}^I$, it is trivial to show the uniqueness. In our model, uniqueness is guaranteed under the condition that the positive cooperation effect does not overwhelmingly dominate the altruism effect to the extent that the game exhibits strong strategic complementarity at some point on $\mathbb{R}^I$. The results of the perfect-information model indicate that this condition is very unlikely to be violated in our setup. The third algorithm, based on the equilibrium strategy profile obtained by the above algorithms, computes the likelihood value. The algorithm conducts Monte Carlo integration over a multivariate normal distribution of dimension I^h, taking the correlation of ε_i into account and using the GHK simulator.

Sequential Game with Reversed Order The difference between the perfect-information sequential game and the incomplete-information simultaneous game may result from both the information structure and the timing of decisions. In an attempt to disentangle these two effects, we estimate a perfect-information sequential game with the reversed order, that is, we estimate our

[20] Because the value of ε_i affects the net utility gain not only as the additive error term but also through the conditional distribution of $\boldsymbol{\varepsilon}_{-i}$, the optimal decision rule (the optimal threshold value for ε_i) does not have an analytical-form solution. Thus the optimal strategy is solved by numerical iteration using the fact that the expected net utility gain is a continuous increasing function of ε_i within the region of parameter values of our interest.

[21] We start the estimation with $a_i^1(\varepsilon_i)$, the threshold values of ε_i that makes near and far indifferent under the standard binary probit model. Every time the likelihood value improves, the previously saved initial point for the numerical iteration is replaced with the new strategy profile.

preferred model under the artificial assumption that the youngest child makes the decision first and the eldest last. This experiment allows us to examine how crucial our decision order assumption is.[22]

Model Fit Comparison Table 10 compares the goodness-of-fit of six alternative models with different behavioral assumptions: independent maximization under no externalities (Model [1]), the non-cooperative perfect-information sequential model (Models [3] and [4]), joint maximization, the non-cooperative private-information model, and the non-cooperative perfect-information sequential model with reverse decision order. The last four columns compare different behavioral assumptions based on the same functional form assumption as the full model (Model [4]). The table reports three comparison measures: the log likelihood values, the Akaike information criterion, and the percentage of correct prediction.

[Insert Table 10: Comparison of Alternative Behavioral Assumptions]

Overall, the comparison supports the use of the non-cooperative sequential framework. The joint-maximization model shows worse goodness-of-fit than the non-cooperative models, indicating the presence of conflicting self-interest.[23] The private-information model fits better with data than the joint-decision model, but not as well as the perfect-information sequential model. Between these two lies the model with reverse decision order, supporting the use of both the perfect information framework and the birth order.[24]

[22] A more direct approach would be to estimate a perfect-information simultaneous game. We do not attempt it because its estimation is not trivial due to the multiplicity of equilibrium.

[23] 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.

[24] We also conduct the same comparison using simpler specifications and find that it does not affect our conclusion.

9 Conclusion

We study externality and strategic interaction among adult siblings regarding their location decision relative to their elderly parents, estimating a rich sequential participation game that exceeds the scope of previous studies. We find positive externality and strategic interaction. Siblings make location decisions non-cooperatively and proximate living with elderly parents is an under-provided public good. While the size of strategic behavior is limited, the impact of the public good problem is striking; of the families with multiple children, 18.3% more parents would have had at least one child nearby had siblings fully internalized externality and jointly maximized family utility.

The complex nature of the subject requires us to employ a tractable framework: we rely on the cross-sectional approach and do not explicitly model parental utility. We conduct a number of model comparisons, however, and our parameter estimates consistently support the significant role of the non-cooperative behavior of siblings, the empirical relevance of externality, and the empirically limited role of sequential interaction, largely for the first time in the literature. Validating our results under a more general setup is left for future research.

The most direct way to achieve the joint-utility optimum is to develop a mechanism that forces families to do so. Historically in many countries, social norms and traditions have forced daughters to assume caregiving obligations (e.g. Holroyd, 2001; Silverstein et al, 2006), which serves as an enforceable mechanism for families to achieve a larger joint utility. In modern societies, however, improved gender equality and increased female labor force participation may have reduced the joint utility of families. The joint-utility optimum can also be achieved by a transfer scheme from those who free-ride to those who provide care, but this option may be practically difficult. Parents can utilize inheritance to enforce such a transfer, although this is not available for parents in need, for whom we identify a severe free-riding problem. Further, this within-family transfer may not be effective where there is the law of legitim—a statutory fraction of the decedent's gross estate from which the decedent cannot disinherit his next-of-kin. Free-riding is thus likely to be more severe in jurisdictions that have legitim, such as Scotland, Louisiana until recently, and Japan. In general, policies that reduce the private costs of caring for elderly parents, such as tax benefits for carers, increases proximate living, but if the costs of such policies are financed by equally taxing other children, their social welfare effect is ambiguous. The welfare effect of public support for parents is similarly ambiguous, depending on families' preferences (altruism) and how such policies are financed.

If the free-riding problem identified in this study is not taken into consideration in future research, misleading conclusions may be drawn. Future research should direct its attention toward externality, the free-riding problem, and the under-provision of care and attention rather than to strategic interactions such as the first-mover advantage, given that more than 90% of children have a dominant strategy.

A Results of Selected Robustness Tests

[Insert Table A1: Robustness of Results]

References

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

[2] Börsch-Supan, A., L. J. Kotlikoff, and J. Morris (1988) "The Dynamics of Living Arrangements

of the Elderly," Working Paper 2787, NBER.

[3] 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.

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

[5] 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.

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

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

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

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

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

[11] 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.

[12] 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.

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

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

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

[16] Johar, M. and S. Maruyama (2012), "Externality and Strategic Interaction in the Location Choice of Siblings under Altruism toward Parents," *UNSW Australian School of Business Research Paper*, 2012/ECON15, University of New South Wales.

[17] 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.

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

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

[20] Mas-Colell, A., M. D. Whinston, and J. Green (1995) *Microeconomic Theory*. Oxford University Press.

[21] 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.

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

[23] 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.

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

[25] Schmidt-Dengler, P. (2006), "The Timing of New Technology Adoption: The Case of MRI," mimeo.

[26] 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 Configurations by Birth Order with Shares Implied under Independence (p=0.404)

One-child families (*N*=1,493)

N of children living near	Observed location configurations	Total	Implied shares under independence (p=0.404)
0	Far: 48.7%	48.7%	59.6%
1	Near: 51.3%	51.3%	40.4%

Two-child families (*N*=2,840)

N of children living near	Observed location configurations	Total	Implied shares under independence (p=0.404)
0	FF: 39.6%	39.6%	35.5%
1	NF: 19.9% ; FN: 23.2%	43.1%	48.2%
2	NN: 17.4%	17.4%	16.3%

Three-child families (*N*=2,054)

N of children living near	Observed location configurations	Total	Implied shares under independence (p=0.404)
0	FFF: 30.3%	30.1%	21.2%
1	NFF: 10.2% ; FNF: 11.1% ; FFN: 13.0%	34.3%	43.1%
2	NNF: 7.3% ; NFN: 7.1% FNN: 9.4%	23.8%	29.2%
3	NNN: 11.7%	11.7%	6.6%

Four-child families (*N*=1,283)

N of children living near	Observed location configurations	Total	Implied shares u/ independence (p=0.404)
0	FFFF: 20.5%	20.5%	12.6%
1	NFFF: 6.7% ; FNFF: 6.3% ; FFNF: 8.5% ; FFFN: 8.9%	30.4%	34.2%
2	NNFF: 3.1% ; NFNF: 3.4% ; NFFN: 4.7% ; FNNF: 3.3% ; FNFN: 4.9% ; FFNN: 4.8%	24.2%	34.8%
3	NNNF: 3.7% ; NNFN: 4.4% ; NFNN: 3.8% ; FNNN: 4.4%	16.3%	15.7%
4	NNNN: 8.7%	8.7%	2.7%

Note: Each digit in the key indicates the proximity of each child to their parents, either far or near, with the first digit representing the eldest child. E.g., "FFN" indicates that in a three-child family, the two elder siblings choose far and the youngest child chooses near. "N" includes coresidence. As a benchmark, the last column shows the share for each number of children living near computed under the assumption that each child independently chooses near with probability 0.404 (=overall average).

Table 2: Definition and Summary Statistics of Variables

Variable	Definition	Mean	S.D.
Outcome			
Near	=1 if the child lives with or within 10 miles from the parent	0.404	0.491
Parent			
P_father_widow	=1 if only father and he is a widow	0.062	0.242
P_father_nonwidow	=1 if only father and he is not a widow (married/separated/divorced/other status)	0.057	0.232
P_mother_widow	=1 if only mother and she is a widow	0.302	0.459
P_mother_nonwidow	=1 if only mother and she is not a widow (married/separated/divorced/other status)	0.132	0.339
P_cohab	=1 if the parent lives with a partner (regardless of marital status)	0.447	0.497
*P_age**	Parent's age	71.939	7.576
*P_white^*	=1 if race is white	0.838	0.369
*P_health**	The first factor from factor analysis consisting of self-assessed health index, ADL and IADL scores (functional limitations expected to last more than 3 months), and three indicator variables for ever being diagnosed with diabetes, hypertension, and stroke.	-0.033	0.782
P_College#	=1 if highest education is college or post college	0.197	0.398
P_SomeCollege#	=1 if highest education is some college (13 – 15 years of formal education)	0.209	0.407
P_HighSchool#	=1 if highest education is high school (reference group - include 15 observations of parents with missing education)	0.354	0.478
*P_<HighSchool^*	=1 if less than 12 years of formal education	0.239	0.427
Geo_HighPop	=1 if lives in a metro area of 1 million population /more (reference group)	0.441	0.476
Geo_MedPop	=1 if lives in a metro area of 250,000 to 1 million population	0.250	0.433
Geo_LowPop	=1 if lives in a metro area of fewer than 250,000 population or non-metro area	0.283	0.450
Geo_missing	=1 if geographical information is missing	0.026	0.160
House	=1 if owns a residential house	0.698	0.459
Child			
C_age	Child's age	44.775	6.863
C_male_single	=1 if the child is a male and single	0.151	0.358
C_female_single	=1 if the child is a female and single (reference group)	0.154	0.360
C_male_partner	=1 if the child is a male and lives with a partner	0.357	0.479
C_female_partner	=1 if the child is a female and lives with a partner	0.339	0.473
C_College	=1 if the child’s highest education is college or post college	0.324	0.468
C_SomeCollege	=1 if the child’s highest education is some college (13–15 yrs of formal education)	0.212	0.408
C_HighSchool	=1 if the child’s highest education is high school or lower (reference group)	0.345	0.475
C_EducMiss	=1 if the child’s formal education is missing/unknown by parents	0.119	0.324
C_kids_partner †	The number of children of the child when the child is married	1.403	1.522
C_kids_single	The number of children of the child when the child is single	0.352	0.937
Wave			
Wave1998	=1 if the data is from wave 1998 (reference group)	0.550	0.497
Wave2004	=1 if the data is from wave 2004	0.249	0.433
Wave2010	=1 if the data is from wave 2010	0.200	0.400

Note:

^ Both parents if a spouse/partner is present.

* Average if a spouse/partner is present.

The one with higher education if a spouse/partner is present.

† Information about grandchildren in the 1998 wave is missing for observations in the AHEAD cohorts. We use information from the next HRS wave in 2000.

Table 3: Estimated Parameters

	Probit [1]		Constant ρ; u^{α}=0 [2]		Constant u^{α} and ρ [3]		Full model [4]	
	coefficient	s.e.	Coefficient	s.e.	coefficient	s.e.	coefficient	s.e.
P_father_widow	0.101***	0.036	0.104**	0.046	0.109**	0.048	0.093	0.060
P_father_nonwidow	-0.339***	0.038	-0.351***	0.047	-0.377***	0.049	-0.318***	0.065
P_mother_widow	0.132***	0.022	0.132***	0.028	0.135***	0.029	0.066*	0.039
P_mother_nonwidow	-0.037	0.029	-0.021	0.035	-0.023	0.037	-0.137**	0.054
P_age	0.005**	0.002	0.005**	0.002	0.005**	0.002	0.003	0.002
P_white	-0.080***	0.024	-0.087***	0.030	-0.090***	0.031	-0.092***	0.032
P_health	-0.061***	0.012	-0.065***	0.015	-0.069***	0.016	-0.048**	0.020
P_College	-0.227***	0.026	-0.227***	0.032	-0.246***	0.033	-0.254***	0.043
P_SomeCollege	-0.076***	0.024	-0.072**	0.030	-0.074**	0.031	-0.094**	0.039
P_<HighSchool	0.060**	0.023	0.068**	0.030	0.080***	0.031	0.046	0.038
Geo_MedPop	-0.009	0.022	-0.009	0.027	-0.007	0.028	-0.007	0.029
Geo_LowPop	-0.091***	0.021	-0.091***	0.026	-0.095***	0.027	-0.095***	0.028
House	0.093***	0.020	0.087***	0.025	0.091***	0.026	0.096***	0.026
C_age	-0.014***	0.002	-0.015***	0.002	-0.014***	0.002	-0.008***	0.003
C_male_single	-0.137***	0.034	-0.139***	0.034	-0.134***	0.033	0.023	0.100
C_male_partner	-0.375***	0.037	-0.386***	0.036	-0.378***	0.035	-0.249**	0.100
C_female_partner	-0.374***	0.037	-0.376***	0.037	-0.367***	0.036	-0.341***	0.045
C_College	-0.406***	0.025	-0.396***	0.026	-0.393***	0.026	-0.423***	0.035
C_SomeCollege	-0.070***	0.026	-0.066**	0.026	-0.069***	0.026	-0.071**	0.034
C_kids_partner	0.021***	0.008	0.021***	0.008	0.021***	0.008	0.021***	0.008
α_0 (=u_i^{α} in model [3] and a constant term in $\log u_i^{\alpha}$ in [4])					0.171***	0.023	-0.951***	0.364
P_father_widow							0.107	0.178
P_father_nonwidow							-0.308	0.317
P_mother_widow							0.329**	0.135
P_mother_nonwidow							0.481***	0.170
P_health							-0.111**	0.051
P_College							0.067	0.117
P_SomeCollege							0.119	0.109
P_<HighSchool							0.205*	0.108
C_male							-0.328	0.222
C_College							0.155	0.108
C_SomeCollege							0.009	0.113
α_1 (additional term in u_i^{α} when more than one child lives near)							0.048	0.144
α_2 (additional term in u_i^{α} for the third and fourth child living near)							-0.038	0.105
γ_0 (constant term in the cooperation term u_i^{γ})							0.628***	0.201
C_age							-0.008***	0.003
C_male_single							-0.178	0.117
C_male_partner							-0.124	0.111
C_female_partner							-0.050	0.057
γ_1 (additional term in u_i^{γ} when two siblings join child *i*)							-0.058	0.174
ρ_0 (constant term in ρ)			0.238***	0.014	0.361***	0.021	0.476***	0.035
Dage							-0.008**	0.003
Dsex							-0.114***	0.024
Log *L*	-11,951.04		-11,788.79		-11,759.61		-11,693.94	
% correct prediction								
All children	62.50%		61.40%		61.58%		61.95%	
All families	38.37%		38.71%		39.14%		39.62%	
1-child families	57.13%		57.33%		58.94%		59.95%	
2-child families	43.03%		41.83%		42.18%		43.06%	
3-child families	29.70%		31.60%		31.65%		31.35%	
4-child families	20.11%		21.51%		21.36%		21.59%	

Note: *N*=18,647. *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively. The top half reports the coefficients of the $u_i^{\beta}(= X_i^{\beta}\beta)$ term, followed by the coefficients in $u_i^{\alpha}, u_i^{\gamma}$, and ρ. The u_i^{α} term in the full model [4] is specified in the exponential function as in Eq.(5). For all models the u_i^{β} term includes the following unreported variables: a constant term, *Geo_missing, C_EducMiss, C_kids_single, Wave2004*, and *Wave2010*. Model [4] also includes *Wave2004, Wave2010*, and *C_EducMiss* in the u_i^{α} term.

Table 4: Observed and Family-Optimal Location Configurations in Three-Child Families

Number of children living near the parent in SPNE (observed location configurations):	Number of children living near the parent in the joint-utility optimal location configuration: Nobody near	1	2	3
Nobody near	100%	46.6%	1.9%	0.0%
1 child	0.0%	53.4%	26.7%	0.0%
2 children	0.0%	0.0%	67.3%	0.0%
3 children	0.0%	0.0%	4.2%	100.0%

Note: Results are based on empirical distribution with Monte Carlo simulation for the error terms with 1,000 random draws. Three-child families are chosen for illustration purposes. Similar patterns of under-provision are observed for two-child and four-child families (available upon request).

Table 5: Observed and Family-Optimal Location Configurations by Family Size

	Number of children living near the parent in SPNE (observed location configurations)					Number of children living near the parent in the joint-utility optimal location configuration				
Family size:	Nobody near	1	2	3	4	Nobody near	1	2	3	4
1-child family	48.7%	51.3%				48.7%	51.3%			
2-child family	39.6%	43.0%	17.4%			24.4%	51.0%	24.6%		
3-child family	30.3%	34.3%	23.8%	11.7%		8.0%	46.5%	35.3%	10.2%	
4-child family	20.5%	30.4%	24.1%	16.3%	8.7%	1.9%	36.5%	39.6%	14.1%	7.9%
Overall average	35.7%	40.2%	16.8%	5.9%	1.5%	21.0%	47.4%	25.2%	5.1%	1.3%
Average ($N_i \geq 2$)	32.5%	37.5%	20.9%	7.3%	1.8%	14.2%	46.5%	31.3%	6.3%	1.6%

Note: The last row shows average numbers over multi-child families. Results are based on empirical distribution with Monte Carlo simulation for the error terms with 1,000 random draws.

Table 6: Characteristics of Simultaneous Normal-Form Games in Two-Child Families

	Location patterns in SPNE (observed location configurations)					
Who has dominant strategy:	(Far–Far)	(Far–Near)	(Near–Far)	(Near–Near)	Total	Total when u_i^{α}×2.0
Both children	99.5%	66.5%	71.1%	99.2%	86.2%	62.7%
Only 1st child	0.3%	16.0%	9.9%	0.5%	5.9%	14.9%
Only 2nd child	0.5%	14.0%	19.0%	0.3%	7.1%	16.3%
Neither	0.0%	3.5%	<0.1%	<0.1%	0.8%	6.1%

	Location patterns in SPNE (observed location configurations)					
Equilibrium patterns in simultaneous normal-form games:	(Far–Far)	(Far–Near)	(Near–Far)	(Near–Near)	Total	Total when u_i^{α}×2.0
No normal-form equilibrium	0.0%	<0.1%	0.0%	<0.1%	<0.1%	<0.1%
Unique equil. (Far–Far)	100.0%	0.0%	0.0%	0.1%	39.6%	23.9%
Unique equil. (Far–Near)	0.0%	94.6%	0.0%	0.0%	21.9%	27.5%
Unique equil. (Near–Far)	0.0%	2.0%	99.9%	0.0%	20.3%	24.6%
Unique equil. (Near–Near)	0.0%	0.0%	0.0%	99.9%	17.3%	18.0%
Two equil. (coordination)	0.0%	0.0%	0.0%	<0.1%	<0.1%	<0.1%
Two equil. (anti-coordination)	0.0%	3.5%	<0.1%	0.0%	0.8%	6.1%

Note: An event that occurs for less than 0.1% of the population is denoted as "<0.1%". Two equil. (coordination) means multiple equilibrium that consists of (Near-Near) and (Far-Far), and Two equil. (anti-coordination) means (Near-Far) and (Far-Near). Results are based on empirical distribution with Monte Carlo simulation for the error terms with 1,000 random draws.

Table 7: Efficiency Type by Family Size

Family size:	[1] Based on estimated distribution of u_i^{α} and u_i^{γ} Joint-utility optimal	Joint-utility suboptimal but Pareto efficient	Prisoners' dilemma	[2] Based on u_i^{α}×2.0 and u_i^{γ}×1.0 Joint-utility optimal	Joint-utility suboptimal but Pareto efficient	Prisoners' dilemma
1-child family	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%
2-child family	76.7%	21.7%	1.6%	71.8%	25.9%	2.3%
3-child family	65.6%	32.0%	2.4%	66.9%	30.0%	3.1%
4-child family	65.6%	31.7%	2.7%	66.3%	30.3%	3.4%
Overall Average	76.8%	21.6%	1.6%	75.1%	22.7%	2.3%
Average among multi-child families, ($N_i \geq 2$)	71.2%	26.8%	2.0%	69.0%	28.1%	2.8%

Family size:	[3] Based on u_i^{α}×1.0 and u_i^{γ}×0.0 Joint-utility optimal	Joint-utility suboptimal but Pareto efficient	Prisoners' dilemma	[4] Based on u_i^{α}×2.0 and u_i^{γ}×2.0 Joint-utility optimal	Joint-utility suboptimal but Pareto efficient	Prisoners' dilemma
1-child family	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%
2-child family	82.8%	17.0%	0.1%	65.6%	30.0%	4.4%
3-child family	73.5%	26.3%	0.2%	61.7%	32.9%	5.4%
4-child family	74.3%	25.4%	0.3%	63.1%	32.0%	4.9%
Overall Average	82.2%	17.6%	0.1%	70.8%	25.3%	3.9%
Average among multi-child families, ($N_i \geq 2$)	77.9%	21.9%	0.2%	63.8%	31.4%	4.8%

Note: Panels [2]-[4] report the results of simulations under different externality parameter values. E.g., in Table 2, the value of u_i^{α} is multiplied by 2.0 for every observation. A joint-utility optimal location configuration is a location arrangement that maximizes the sum of children's utility. Prisoners' dilemma means a location configuration that has another Pareto-dominating location configuration. The last row shows average numbers over multi-child families. Results are based on empirical distribution with Monte Carlo simulation for the error terms with 1,000 random draws.

Table 8: Location Patterns and Efficiency Type in Two-Child Families

	Efficiency type: Joint-utility optimal	Joint-utility suboptimal but Pareto efficient	Prisoners' dilemma	Total
Location patterns in SPNE (observed location configurations)				
(Far – Far)	31.5% / 61.5%	67.3% / 35.8%	70.5% / 2.6%	39.6% / 100%
(Far – Near)	23.6% / 78.7%	21.3% / 19.4%	29.5% / 1.9%	23.2% / 100%
(Near – Far)	22.6% / 87.9%	11.4% / 12.1%	0.0% / 0.0%	19.9% / 100%
(Near – Near)	22.4% / 100%	0.0% / 0.0%	0.0% / 0.0%	17.4% / 100%
Total	100% / 77.4%	100% / 21.1%	100% / 1.5%	100% / 100%
The joint-utility optimal location configuration				
(Far – Far)	31.5% / 100%	0.0% / 0.0%	0.0% / 0.0%	24.4% / 100%
(Far – Near)	23.6% / 69.3%	36.7% / 29.4%	24.2% / 1.4%	26.3% / 100%
(Near – Far)	22.6% / 70.8%	31.9% / 27.3%	32.8% / 2.0%	24.7% / 100%
(Near – Near)	22.4% / 70.5%	31.4% / 26.9%	43.0% / 2.6%	24.6% / 100%
Total	100% / 77.4%	100% / 21.1%	100% / 1.5%	100% / 100%

Note: The first number in each cell represents its column share and the second number its row share. A joint-utility optimal location configuration is a location arrangement that maximizes the sum of children's utility. Results are based on empirical distribution with Monte Carlo simulation for the error terms with 1,000 random draws.

Table 9: Reverse-Order SPNE in Two-Child Families

Location patterns in Reverse-order SPNE:	Location patterns in the subgame-perfect Nash equilibrium (observed location configurations) (Far–Far) N=1,125: 39.6%	(Far–Near) N=658: 23.2%	(Near–Far) N=564: 19.9%	(Near–Near) N=493: 17.4%	Total
(Far–Far)	99.9%	0.0%	0.0%	0.1%	39.6%
(Far–Near)	0.0%	92.2%	0.0%	0.0%	21.4%
(Near–Far)	0.0%	7.8%	100%	<0.1%	21.7%
(Near–Near)	<0.1%	<0.1%	0.0%	99.9%	17.4%

Utility changes in reverse-order SPNE:	Location patterns in the subgame-perfect Nash equilibrium (observed location configurations) (Far–Far) N=1,125	(Far–Near) N=658	(Near–Far) N=564	(Near–Near) N=493	Total	Total when $u_i^{\alpha}\times 2.0$
No change	99.9%	92.2%	100%	99.9%	98.2%	90.8%
1st child (−); 2nd (+); total (−)	0.0%	3.7%	0.0%	0.0%	0.9%	4.8%
1st child (−); 2nd (+); total (+)	0.0%	4.1%	0.0%	0.0%	1.0%	4.5%

Note: Events that occur for less than 0.1% of the population are denoted as "<0.1%". Although we do not report it here because it is extremely rare, the first child's utility may increase in a reverse-order SPNE. Also the second child's utility may decrease, but these two events never occur at the same time (no second-mover advantage). Results are based on empirical distribution with Monte Carlo simulation for the error terms with 1,000 random draws.

Table 10: Comparison of Alternative Behavioral Assumptions

Behavioral assumption:	Independent maximization	Non-cooperative, sequential (preferred model)		Joint maximization	Non-cooperative, private information	Non-cooperative, reverse-order sequential
Functional form assumption:	u^{α}=u^{γ}=ρ =0 (Model [1])	u^{α}, ρ constant; u^{γ}=0 (Model [3])	heterogeneous externalities (Model [4])	heterogeneous externalities	heterogeneous externalities	heterogeneous externalities
Log L	-11,951.0	-11,759.6	-11,693.9	-11,957.7	-11,727.8	-11,711.4
Number of parameters	26	28	52	52	52	52
AIC	23,954.1	23,575.2	23,491.9	24,019.5	23,559.5	23,526.9
% correct prediction:						
All children	62.50%	61.58%	61.95%	61.31%	61.91%	61.94%
All families	38.37%	39.14%	39.62%	38.63%	39.35%	39.66%
1-child families	57.13%	58.94%	59.95%	58.20%	61.29%	60.55%
2-child families	43.03%	42.18%	43.06%	42.25%	41.87%	42.61%
3-child families	29.70%	31.65%	31.35%	30.77%	31.30%	31.69%
4-child families	20.11%	21.36%	21.59%	20.42%	21.12%	21.59%

Note: Based on 18,647 child observations in 7,670 families. When u^{α}=u^{γ}=ρ =0, there is no externality or dependency among siblings, and independent utility maximization and joint utility maximization coincide. AIC stands for the Akaike information criterion. The percentage of correct prediction is based on the predicted location outcome for each family observation that is defined as the location configuration with the highest implied probability among all possible location configurations.

Table A1: Robustness of Results

	Full model [4]	Without siblings of age difference < 2 years [5]	Multi-child families [6]	Simplified model without potentially endogenous variables [7]	
P_father_widow	0.093	0.093	0.076		
P_father_nonwidow	-0.318***	-0.338***	-0.263***		
P_mother_widow	0.066*	0.077	0.041		
P_mother_nonwidow	-0.137**	-0.172***	-0.176***		
P_age	0.003	0.000	0.003	*P_age*	0.006***
P_white	-0.092***	-0.086**	-0.057*	*P_white*	-0.150***
P_health	-0.048**	-0.073***	-0.034		
P_College	-0.254***	-0.262***	-0.241***	*P_College*	-0.274***
P_SomeCollege	-0.094**	-0.092**	-0.104	*P_SomeCollege*	-0.132***
P_<HighSchool	0.046	0.038	0.014	*P_<HighSchool*	0.038
Geo_MedPop	-0.007*	-0.024	-0.004	*Geo_MedPop*	-0.002
Geo_LowPop	-0.095***	-0.097***	-0.097***	*Geo_LowPop*	-0.100***
House	0.096***	0.100***	0.105***	*House*	0.088***
C_age	-0.008***	-0.005	-0.011***	*C_age*	-0.005**
C_male_single	0.023	-0.072	0.218	*C_male*	-0.117
C_male_partner	-0.249**	-0.370***	-0.040		
C_female_partner	-0.341***	-0.345***	-0.337***		
C_College	-0.423***	-0.438***	-0.441***	*C_College*	-0.422***
C_SomeCollege	-0.071**	-0.058	-0.066	*C_SomeCollege*	-0.064
C_kids_partner	0.021***	0.018**	0.025***		
α_0 (constant term)	-0.951***	-1.129***	-0.495	α_0	-0.338**
P_father_widow	0.107	0.178	0.091		
P_father_nonwidow	-0.308	-0.199	-0.365		
P_mother_widow	0.329**	0.280**	0.297**		
P_mother_nonwidow	0.481***	0.553***	0.344**		
P_health	-0.111**	-0.104**	-0.097*		
P_College	0.067*	0.126	0.043	*P_College*	0.011
P_SomeCollege	0.119	0.171	0.132	*P_SomeCollege*	0.213
P_<HighSchool	0.205*	0.360***	0.155	*P_<HighSchool*	0.374**
C_male	-0.328	-0.045	-0.622**	*C_male*	0.166
C_College	0.155	0.145	0.086	*C_College*	0.077
C_SomeCollege	0.009	-0.047	-0.019	*C_SomeCollege*	-0.053
α_1	0.048	-0.142	0.214**	α_1	0.252
α_2	-0.038	-0.267	0.073	α_2	0.074
γ_0 (constant term)	0.628***	0.710***	0.701***	γ_0	0.498***
C_age	-0.008***	-0.011***	-0.005*	*C_age*	-0.011***
C_male_single	-0.178	-0.103	-0.281*	*C_male*	0.078
C_male_partner	-0.124	0.038	-0.249*		
C_female_partner	-0.050	-0.013	-0.064		
γ_1	-0.058	-0.197	0.311*	γ_1	-0.130
ρ_0 (constant term)	0.476***	0.490***	0.364***	ρ_0	0.511***
Dage	-0.008**	-0.003	-0.009***	*Dage*	-0.008
Dsex	-0.114***	-0.108***	-0.122***	*Dsex*	-0.094***
N of child observations	18,467	13,029	16,974		18,467
Log *L*	-11,693.94	-8,241.08	-10,694.73		-11,846.95
% correct prediction					
All children	61.95%	62.22%	62.07%		60.87%
All families	39.62%	42.58%	34.66%		38.14%
1-child families	59.95%	60.35%	NA		57.33%
2-child families	43.06%	42.49%	42.71%		40.88%
3-child families	31.35%	32.09%	31.69%		30.92%
4-child families	21.59%	23.49%	21.59%		21.28%

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively. See Eq. (5) for the functional specification. All models include in the u_i^β term: a constant term, *Geo_missing, C_EducMiss, C_kids_single* (except for model [4]), *Wave2004,* and *Wave2010*. The u_i^α term also includes *C_EducMiss, Wave2004,* and *Wave2010.*