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

Shiko Maruyama† and Meliyanni Johar

November 18, 2015

Abstract

There is a potential free-rider problem when several siblings consider future provision of care for their elderly parents. Siblings can commit to not providing long-term support by living far away. If location decisions are made by birth order, older siblings may enjoy a first-mover advantage. We study siblings' location decisions relative to their parents by estimating a sequential participation game for US data. We find: (1) limited strategic behavior: in two-child families, more than 92% of children have a dominant strategy; and (2) a non-negligible public good problem: in families with multiple children, 18.3% more parents would have had at least one child living nearby had location decisions been made cooperatively.

*We thank Chris Bidner, Zvi Eckstein, Denzil Fiebig, Makoto Hanazono, Susumu Imai, Michael Keane, Kevin Lang, Robert Porter, Nobue Suzuki, John Wooders, two anonymous referees, and the editor, Petra Todd, for insightful comments and suggestions that have improved the presentation and analysis. Research support from the Australian Research Council's Discovery Projects funding scheme (project number DP110100773) and from the ARC Centre of Excellence in Population Ageing Research (CEPAR) is also gratefully acknowledged. Earlier versions of the paper were circulated under the title "Externality and Strategic Interaction in the Location Choice of Siblings under Altruism toward Parents."

†Corresponding author. E-mail: shiko.maruyama@uts.edu.au

1 Introduction

The burden of caring for elderly parents has been well-documented (e.g., Ettner (1996), van den Berg, Brouwer, and Koopmanschap (2004), Bolin, Lindgren, and Lundborg (2008), and Lilly, Laporte, and Coyte (2010)). When several siblings consider providing care for their elderly parents, altruism toward the parents and the cost of caregiving result in a textbook public good problem. The more altruistic the siblings are, the stronger is their incentive to free-ride on each other because a stronger altruism implies a larger positive externality of caregiving.

This public good problem is particularly plausible when we consider siblings' location decisions. The opportunity cost of living near the parent may be substantial, although it is not widely documented in the literature. The discrete nature of location choice and associated relocation costs make efficient bargaining difficult. Furthermore, there exists a potential commitment device arising from birth order: the oldest child may enjoy a first-mover advantage by moving far away as soon as schooling is completed. Consistent with this argument, Konrad, Künemund, Lommerud, and Robledo (2002) find that in Germany, older siblings are more likely to move far away from their parents than younger siblings.

We quantify this free-rider problem and first-mover advantage for the first time in the family care literature by studying siblings' location decisions relative to their elderly parents. We build a game-theoretic econometric model to explain cross-sectional variation in the patterns of sibling location in the US. It is a perfect-information participation game in which, by birth order, each sibling sequentially makes a once-and-for-all location decision whether to live close to their parent. Although this approach abstracts from dynamics in location decisions except for birth order (as in most previous studies), our model instead features rich heterogeneity and encompasses a wide variety of participation games. Consequently, our analysis allows us to discover (1) the degree and nature of externality, (2) the associated under-provision or over-provision of proximate living, (3) the game structure and equilibrium characteristics, (4) the size of the first-mover advantage, and (5) how externality and inefficiency vary across families. To confirm the validity of our model, we also estimate a private-information model and a cooperative model in which siblings maximize their utility sum.

The key innovation in our empirical framework relies on the fact that a wide range of participation games can be summarized by three structural parameters: *altruism*, *private cost*, and *cooperation*. The "cooperation" term captures another likely source of positive externality of proximate living, the so-called "synergy effect": siblings living near parents may be able to cooperate and provide care more efficiently. In fact, shared caregiving is widely observed (see, e.g., Matthews and Rosner (1988) and Checkovich and Stern (2002)). By modeling altruism and cooperation together and by introducing heterogeneity in the three structural parameters, we can incorporate a broad range of participation games and identify the games played by American families.

Informal care still plays an important role in aging societies, despite a trend toward formal care. The OECD (2005) reports that around 80% of the hours of care for the elderly with a disability or severe medical condition are provided informally. Despite declining intergenerational coresidence, the majority of adult Americans still live within 25 miles of their mothers (see Compton and Pollak (2013)). 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 (see Matthews and Rosner (1988) and Bonsang (2009)). A good understanding of adult children's location decisions, hence, serves as an important step in designing public policies to promote the well-being of families in aging societies. In particular, by quantifying the extent of the public good problem, externality, and strategic behavior in the location decisions of families and by examining Pareto optimality, this study offers useful insights into who should be supported, subsidized, or taxed in order to achieve higher family welfare.

The results are summarized as follows. First, the location game played by American siblings is characterized by moderate altruism and cooperation. This implies very limited strategic behavior. In two-child families, more than 92% of children have a dominant strategy. The first-mover advantage is almost negligible: reversing birth order affects only 1.9% of two-child families. Second, however, there is non-negligible under-provision of proximate living due to free-riding. In multi-child families, 28.8% end up in location configurations that are not joint-utility optimal. Most typical in this case is the situation in which no child lives near the parent and no Pareto improving location configuration exists; however, the siblings can achieve higher joint utility if one of them lives near the parent. In families with multiple children, 18.3% more parents than actually observed in data would have had at least one child living nearby had location decisions been made cooperatively than as actually observed in data. Third, we find substantial heterogeneity across families. The under-provision of proximate living is more severe if children exhibit strong altruism toward their parents, particularly in a family with a single mother with limited education, poor health, and younger children. Lastly, we find that the non-cooperative model fits the data considerably better than the joint-utility maximization model.

This paper also contributes to the empirical literature on games. First, our model features rich heterogeneity in the two sources of externality. Consequently, different players face participation games with different equilibrium characteristics (e.g. coordination and anti-coordination games). This enables us to draw inferences about the shares of families in the *prisoners' dilemma* situation, families achieving the joint-utility optimum, and families with a large first-mover advantage. Second, this paper is one of very few empirical analyses to study the first-mover advantage, preemption, and commitment in sequential decision making. Most prior empirical studies examine extremely simple cases, such as two-player games, with two exceptions. Schmidt-Dengler (2006) studies the timing game of MRI adoption by hospitals in a fairly general setup and finds a significant but small preemption effect. Stern (2014), probably the closest work to ours, studies the patterns of sibling location in a private-information sequential framework. The models in these two studies, however, lack the rich heterogeneity of our model that allows us to capture a wider variety of participation games.

2 Related Literature

A small but tangible body of literature applies a non-cooperative game-theoretic framework to study interactions among siblings with respect to informal care arrangements (see Hiedemann and Stern (1999), Checkovich and Stern (2002), Engers and Stern (2002), Byrne, Goeree, Hiedemann, and Stern (2009), and Knoef and Kooreman (2011)). In these models, each family member acts to maximize his/her own utility, and the equilibrium arrangement is 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, Goeree, Hiedemann, and Stern (2009) enrich these studies by also modeling consumption, financial transfers for formal home care, and labor supply. These studies use US data, whereas Knoef and Kooreman (2011) estimate a model using European multi-country data. With the exception of Byrne, Goeree, Hiedemann, and Stern (2009), the results of these studies all indicate interdependence in caregiving decisions among siblings. For example, Knoef and Kooreman (2011) argue that if siblings engage in joint utility maximization, 50% more informal care will be provided to parents, and the cost to the children will increase to a much lesser extent. All these structural studies employ a game-theoretic framework to explain across-family variations in care arrangements, taking families' location decisions as given.

We advance the literature on informal care in two ways. First, we are one of the first to apply a game-theoretic framework to the location decisions of siblings, rather than the informal care arrangement decision. Studying the location decision is important because the location pattern is a critical determinant of formal and informal care arrangements (see Checkovich and Stern (2002), Engers and Stern (2002), Bonsang (2009), and Hiedemann, Sovinsky, and Stern (2014)). There are myriad economics and non-economics studies on coresidence and co-location between elderly parents and their children (e.g., Börsch-Supan, Kotlikoff, and Morris (1988), Dostie and Léger (2005), Hank (2007), Fontaine, Gramain, and Wittwer (2009), Hotz, McGarry, and Wiemers (2010), Johar and Maruyama (2011), Compton and Pollak (2013), Johar and Maruyama (2014), and Maruyama (2015)), but few investigate the non-cooperative decision of family living arrangements,[1] and none quantifies the free-rider problem among siblings, although the discrete and long-term nature of location decisions may reinforce the free-riding and strategic behavior involved in the coordination of caregiving among siblings.

Second, we are the first to develop an econometric model that captures the sequential aspect of decision making among siblings and to quantify its empirical importance. All studies with a game-theoretic econometric model in this literature assume that siblings make decisions simultaneously.

[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 using a bargaining model of coresidence, monetary transfers, and marriage. The study on living arrangements by Hoerger, Picone, and Sloan (1996) allows multiple children to contribute to caregiving, based on a single family utility function.

Our study builds on the nonstructural study by Konrad, Künemund, Lommerud, and Robledo (2002), who estimate an ordered logit model of children's distance from the parent with child-level data of two-child families drawn from the German Aging Survey. They find that first-born children are more likely to live far from their parents than their younger siblings, and 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, observed birth-order asymmetry may simply be explained by the observed characteristics of siblings. To the best of our knowledge, Stern (2014) is the only empirical work that studies the strategic location decision of families with more than two siblings. The unique feature of his work is that his model allows for private information in each child's location preferences, in addition to common-knowledge factors that are unobservable to the econometrician. In one sense, Stern's (2014) setup is more general than ours because it allows for two types of unobservable factors, but this feature itself makes the identification of model parameters challenging, and hence requires strong model restrictions: in particular, the private term in each sibling's location preferences follows an independent normal distribution; what we call altruism has limited heterogeneity; and externality due to cooperation is not allowed. Consequently, Stern's (2014) setup lacks the rich heterogeneity of our approach, and the source of identification and the generalizability of the results need further investigation.

Given the complexity of care and living arrangements, one model does not capture all the possible aspects of family decision making. Prior studies utilize various 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 sequential interaction and externality. Our study should therefore be regarded as a complement to existing studies.

3 Data and Descriptive Results

3.1 Data

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, and has since added two new cohorts in 2004 and 2010. It tracks the health, wealth, and well-being of elderly individuals and their spouses. The HRS also questions respondents about the demographics and location of all their children.

To make our econometric model tractable, we take a cross-sectional approach and abstract from dynamic aspects other than sequential decision by birth order. We combine the three HRS waves in 1998, 2004, and 2010, and construct our "cross-sectional" sample 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 included in our HRS 1998 sample. We then add families from HRS 2010, repeating the same procedure. Each family thus appears only once in our sample. We pool the three waves to increase the sample size and secure time variation (twelve years apart). As reported in Johar and Maruyama (2012), our basic results are unaffected by the choice of survey 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 35 years old or older and whose oldest child is younger than 65; (7) whose oldest 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 restrict the number of children to four to limit the computational burden. For the purpose of our research, 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 move for postgraduate 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 focuses on where children set up their own families, thus moves around retirement age should be excluded. Lastly, we exclude families with same-age children because we utilize birth order.

From this sample of parents, we create a child-level data set. 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* the parent. Although coresidence is becoming less common, shared caregiving by siblings living nearby is commonly observed (e.g., Matthews and Rosner (1988) and Checkovich and Stern (2002)). Siblings living nearby also contribute to the family by other means — by frequent visits and as a backup in the case of primary caregiver burnout.[2] Due to the design of the HRS, proximity is

[2] We focus on the binary setup for the ease of computation and interpretation, following the majority of the literature. One of the referees suggests a three-alternative setup rather than a binary setup, referring to Compton

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) and Byrne, Goeree, Hiedemann, and Stern (2009)).[3]

Table 1 presents the location patterns of siblings in our sample by the number of children in a family. The top panel shows that 48.7% of only children live far from their parents. The second panel shows that elderly parents with two children are most likely to have one child nearby (43.1%) and least likely to have both of them nearby (17.4%). Naturally, the probability of having at least one child nearby increases with the number of children: parents of four children are least likely to live with no child nearby (20.5%) compared to parents with fewer children. Table 1 also reports the detailed location configurations by birth order. Each possible configuration is denoted by the sequence of "F" and "N", indicating each sibling's decision from the oldest to the youngest. For example, "FFN" indicates the configuration of a three-child family where only the youngest child lives near the parent. In the last column, we report the theoretical share of each location configuration under the independence assumption. We compute these shares by using the overall propensity of living near parents (p =40.4%) under the assumption that there is no externality and each child makes a decision independently.

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

Table 1 highlights five empirical regularities that our econometric model needs to address. First, only children are more likely to choose to live nearby compared to children with siblings,

and Pollak's (2013) work, which finds qualitative difference between those who choose proximate living and those who choose coresidence. To examine the validity of the binary setup, we estimated several nested logit models that allow for two different nesting structures: (1) siblings make a decision between "living far" vs. "living near/coresidence" and (2) siblings make a decision between "separate living (living far/near)" vs. "coresidence." We found no strong evidence for the latter, that is, adult children do not appear to decide whether to leave the parents before they choose distance. This finding supports our binary aggregation.

[3] Using the National Survey of Families and Households, Compton and Pollak (2013) report that the median distance between a married couple and the husband's (wife's) mother is 25 (20) miles. Their finding suggests that a substantial proportion of children whom we categorize as "living far" live within 30 miles of their parent.

perhaps because only children have no one on which to free-ride. Second, in multi-child families, the location decisions of siblings are correlated: we observe polar cases such as "FFF" and "NNN" more frequently than theoretically implied by shares under independence. This correlation may arise as a result of (1) similarity in siblings' preferences, (2) similarity in siblings' characteristics, or (3) cooperation between siblings. These three possibilities are distinguished in our econometric model.

The third empirical regularity is birth-order asymmetry. Conditional on one child living near the parent, two-child families have two possible location configurations: "NF" and "FN". Table 1 shows that "NF" is less frequently found than "FN". The three- and four-child family panels show the robustness of this birth-order asymmetry: 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, Künemund, Lommerud, and Robledo's (2002) argument of first-mover advantage. However, this may simply reflect systematic difference between older and younger siblings. For example, it is a well-documented fact that older children tend to have more education than their younger siblings (see, e.g., Davis (1997), Sulloway (2007), and Booth and Kee (2009)). In our sample, the share of those who have a university degree is 37.0% and 35.6% respectively for the first and second children in two-child families, and in three-child families, it is respectively 36.2%, 32.0%, and 31.3% for the first, second, and third children. How much of the observed birth-order asymmetry is attributed to first-mover advantage is an empirical question.

The fourth empirical regularity concerns how a younger child responds to an older sibling's location decision. In the panel of two-child families, for example, conditional on the first child moving far from the parent, 63.1% (= 0.396/(0.396+0.232)) of second children choose "F", whereas conditional on the first child staying near the parent, 53.4% of second children choose "F". On the surface, this appears inconsistent with the free-rider problem, in which the second child is more likely to leave the parent when the first child remains near the parent. However, these numbers may simply capture the above-mentioned similarities in preferences and characteristics. Whether free-riding behavior exists needs to be examined after we have controlled for correlation.

Lastly, among the polar cases such as "FFF" and "NNN", everyone-far location patterns show larger differences between observed and predicted shares than everyone-near location patterns. This distortion toward everyone-far location patterns is not explained by either correlation or the cooperation effect. It instead suggests that free-riding behavior under altruism leads to under-provision of proximate living.

3.3 Explanatory Variables

We use the characteristics of both parents and children. Table 2 provides the definitions of the explanatory variables and their summary statistics. The parental variables, which are always named with prefix "$P\_$", include demographics (age, sex, marital/cohabitation status, and ethnicity), education, health status, location type (urban or rural), and housing status. For the child variables, which always have prefix "$C\_$", we use age, sex, education, marital status, and information on grandchildren. Parental health status is constructed as the first factor from a factor analysis that includes: (1) self-assessed health index; (2) Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) scores; and (3) previous diagnoses of diabetes, hypertension, and stroke. These diagnoses are chosen because they tend to be persistent and are relatively common among the elderly. When a respondent parent is married, the health data of the couple are averaged to reduce the computational burden of the full model while keeping its interpretation simple.[4]

[4] Our specification tests based on various binary probit models of the location decision suggest that how we include and aggregate the health indexes of married parents (e.g. separate indexes, adding interaction term, the worst of the

Our assumption about parental location and housing status is that the location configuration is determined by children's migration, not parental migration. This approach is justified by the fact that, 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 parents.[5]

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

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.

4 The Model

4.1 Environment

We consider a game played by children. Each child chooses whether to live close to their parent. To make our analysis tractable, "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 oldest 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,

two, etc.) does not affect our main results.

[5] Removing the location and housing variables does not affect our main results.

but it does not mean that parents are passive and play no role. Parents may influence children's payoff function by promising compensation for informal care in the future. Family bargaining and intergenerational transfers are implicit in our payoff function and our coefficient estimates should be interpreted in a reduced-form way.[6]

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 later life, such as changes in the family structure and the deterioration of parents' health, is beyond the scope of this study, as it has been for most previous studies. Our utility function should be interpreted as the present discounted value of future utility.

Third, we rely on a non-cooperative framework. An alternative is a model of joint-utility maximization, which we also 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 most studies of empirical games assume incomplete information, the perfect-information assumption is reasonable in the family setting because family members know each other well.[7] To verify this assumption, we also estimate an incomplete-information simultaneous game.

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

[7] The informal care literature uses both approaches: Byrne, Goeree, Hiedemann, and Stern (2009) assume a complete information game, whereas Engers and Stern (2002) and Stern (2014) 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 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.[8] Let $N = \sum_{k=1}^{I} a_k$ 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, regardless of which child that is. 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, proximate living has a positive externality, and child i *free-rides* on child j if child i lives far and child j lives near the parent. Altruism, $u_i^{\alpha}(N)$, may be an increasing function of N if the number of children living nearby relates to the

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

amount of care and attention given to the parent.

The next parameter, u_i^β, captures child i's *private cost* (or benefit) from living near the parent irrespective of the other children's decision, a_{-i}. This term includes not only caregiving burdens but also any other net utility or monetary gain/loss from living near, such as opportunity costs, housing benefits in the case of coresidence, attachment to the location, grandchild care from parents, and the consumption value of the time child i shares with the parent. When living far away, a child often provides financial assistance instead of informal care (as discussed by Antman (2012)). This is also a part of the net private cost term.

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 because siblings can share the costs of looking after parents. This term, however, becomes negative under the bequest motive hypothesis discussed in Bernheim, Schleifer, and Summers (1985) — the presence of another sibling taking care of the parent reduces transfers from the parent. The cooperation term, $u_i^\gamma(N)$, may also capture the benefit of proximate living that is unrelated to the parent; for example, children may enjoy living close to each other and they may provide childcare to their nephews and nieces.[9] 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 $u_i^\alpha(1) + u_i^\beta$.

4.3 Equilibrium and Efficiency Benchmarks

Siblings make location decisions by birth order. Their 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

[9] In reality, siblings living near one another may enjoy this benefit even if they do not live with their parent. In our setup, the utility flow, $u_i^\gamma(N)$, only occurs when siblings live near the parent.

decision *at every decision node* (thus note the difference between a_i and s_i). A subgame-perfect Nash equilibrium (SPNE) is obtained when no child expects to gain from individually deviating from their equilibrium strategy *in every subgame*. Every finite game with perfect information has a pure-strategy SPNE (Zermelo's theorem).[10] In this study, we only consider pure strategies. In our perfect-information setup, mixed strategies are irrelevant because every decision node has a choice that is strictly better than the other.

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 second child has two decision nodes, the choice set of the second child comprises four strategies, which we refer to as *always far*, *imitate*, *preempted*, and *always near*, as shown in Figure 1. *Preempted*, for example, refers to the second child's strategy of staying near the parent only when the older sibling moves away. Given the payoffs at each terminal node, we can find the SPNE outcome and strategies by sequentially solving the choice problem at each decision node from the youngest child to the oldest 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 many-to-one mapping structure.

To examine the desirability of an equilibrium outcome, we use two efficiency measures: (1) Pareto efficiency and (2) efficiency in joint utility. Even if a game has a unique SPNE, it may have a Pareto-improving (non-equilibrium) outcome, which constitutes the well-known *prisoners' dilemma*. Efficiency in joint utility, or Kaldor-Hicks efficiency, concerns the sum of siblings' utility. Although this criterion does not guarantee a Pareto improvement, it is sensible to study this

[10] For Zermelo's theorem, see Mas-Colell, Whinston, and Green (1995), page 272.

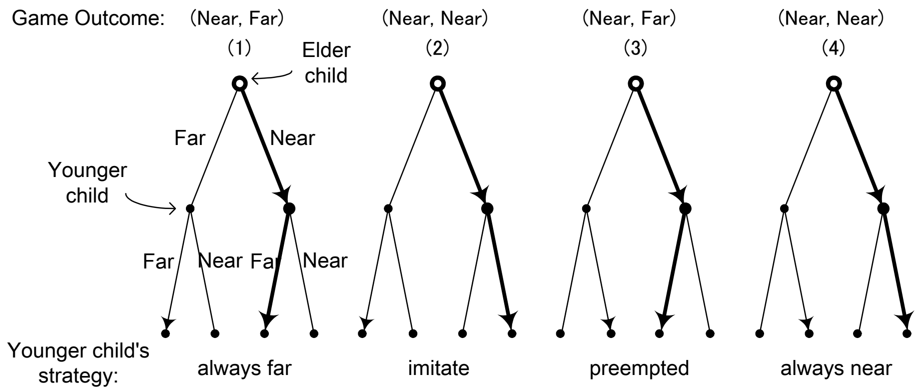

Figure 1: Strategies and Outcomes in Extensive-Form Presentation

efficiency measure because it has implications for implementable compensation schemes.[11]

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.[12] Example 2 shows 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 greater joint utility at (Near, Far) or (Far, Near) — at the expense of either sibling's disutility. If compensation

[11] Note that our framework does not include parents' welfare, although this may be partly captured by the children's altruism parameter. The terms "inefficiency" and "under-provision" in this study should be interpreted as such. If proximate living increases parental utility, our inefficiency measures are the lower bound of family inefficiency.

[12] Although we do not discuss it here, there is a normal-form representation corresponding to the sequential game.

is possible, these efficient outcomes will be chosen.

Altruism, private cost, and cooperation in our model govern the game structure in each family. For example, assuming a constant altruism (i.e. $u^{\alpha}(N=1) = u^{\alpha}(N=2)$), the payoff matrix in Example 1 corresponds to $\left(u_i^{a}, u_i^{\beta}, u_i^{\gamma}\right) = (1, -2, 3)$.[13] 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. A negative cooperation leads to an anti-coordination game (also known as a congestion 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 the sequence is introduced, the SPNE is (Far, Near), and child 1 enjoys 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 higher utility by not playing the normal-form dominant strategy. The decision order provides child 1 with a commitment device and hence a 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

The main theoretical predictions in symmetric two-player games are summarized as follows. First, joint-utility inefficiency increases with the *absolute* size of the two sources of externality — altruism, u^{α}, and cooperation, u^{γ}. Both positive and negative values of u^{γ} enlarge inefficiency. The under-

[13] To see this, use $u_1(1,1) = u^{\alpha} + u^{\beta} + u^{\gamma} = 2$, $u_1(1,0) = u^{\alpha} + u^{\beta} = -1$, and $u_1(0,1) = u^{\alpha} = 1$.

provision of proximate living results from positive values of u^α and u^γ because children do not take into consideration positive externality to other siblings. Similarly, if $u^\gamma < 0$, excessive participation may occur, creating a setting similar to the standard entry game.[14] When there is no externality ($u^\alpha = u^\gamma = 0$), the SPNE outcome maximizes joint utility.

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 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:

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

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 *supermodular game*) may also result from a small u^α and/or large u^γ. In

[14] 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.

our symmetric binary setup, however, a second-mover advantage never appears because strategic complementarity degenerates the game into the choice between (Near, Near) and (Far, Far) and 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 $u^{\gamma} > 0$, this 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 externality and distortion 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 optimum.

5 Estimation

5.1 Random Term

To match the model with the data, we need an individual-specific random term. We assume an additive random term, ε_i, that affects utility from 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 random term is assumed to follow a normal distribution independent of $\left(u_i^{\alpha}, u_i^{\beta}, u_i^{\gamma}\right)$. Under the assumption of perfect information, ε_i is unobservable to an econometrician but is observed by child i's siblings. The normality assumption implies that the game almost surely has a unique equilibrium because ties occur with probability measure zero.[15]

[15] 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.*

As with standard random-utility models, the level of utility is not identified. Assuming the same variance for every child, we normalize the variance of $\varepsilon_{i,h}$ to one. Formally,

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

where Ω^h is an $I_h \times I_h$ 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

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

where θ^{ρ} is a vector of parameters and $X^{\rho}_{i,j}$ is a set of relational variables between children i and j, such as their age difference.

5.2 Specifying Functional Forms

For estimation, we also need to specify the functional forms of $u^{\alpha}_i(N), u^{\beta}_i$, and $u^{\gamma}_i(N)$. Let $X^{\alpha}_i, X^{\beta}_i$, and X^{γ}_i be vectors of covariates observable to the econometrician including a constant term. Below, we report the results of the following four specifications. Specification [1] imposes $u^{\alpha}_i(N) = u^{\gamma}_i(N) = 0, u^{\beta}_i = X^{\beta}_i\beta$, and $\rho_{i,j} = 0$. This specification implies no interdependency between siblings, and the model degenerates to a standard binary probit model. Specification [2] allows $\rho_{i,j}$ to be some constant, ρ_0, so that the preferences of siblings may correlate. Specification [3] introduces externality in the most parsimonious way: $u^{\alpha}_i(N) = \alpha_0$, $u^{\beta}_i = X^{\beta}_i\beta$, $u^{\gamma}_i(N) = 0$, and $\rho_{i,j} = \rho_0$. Specification [4] allows externality to vary depending on N and the covariates.

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 setup, a negative value of $u_i^{\alpha}(N)$ has no sensible interpretation because it implies a situation in which a child receives disutility if one of the children lives near the parent, irrespective of which child that is. After we estimate Specification [3] and confirm a positive estimate of α_0, we introduce heterogeneity in this term using the exponential function to guarantee positive values. 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 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 who 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 $\alpha_0 + \beta_0$. These two values determine the degree of strategic substitutability, $\alpha_0 - \gamma_0$. When we assume no cooperation effect (i.e. $\gamma_0 = 0$), the identification of α_0 and β_0 follows.

If $\gamma_0 \neq 0$, the rest of the 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 who hate each other: $(\alpha_0, \beta_0, \gamma_0) = (0, 0, -2)$. Both types of family result in $\alpha_0 + \beta_0 = 0$, $\beta_0 + \gamma_0 = -2$, and $\alpha_0 - \gamma_0 = 2$, thus these two family types are indistinguishable when studying the choice problem of child 2. In this particular example, the payoff function exhibits decreasing difference ($-\alpha_0 + \gamma_0 = -2 < 0$). Under decreasing difference, child 2 never chooses the *imitate* strategy because for *imitate* to be optimal, the net gain of living near must be larger when child 1 chooses near than when child 1 chooses far. Thus, we will observe one of the other three strategies, *always far*, *preempted*, and *always near*, depending on the value of ε_2. The last step of the identification is achieved by studying child 1's choice problem when child 2 takes the strategy *preempted* by 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 of family, we will observe child 1 choosing "near" and "far" with the same probability. In other words, the size of the birth-order asymmetry *given the size of the first-mover advantage* provides essential information for separately identifying the three parameters.

The identification of ρ also relies on sequential interaction. Positive correlation between the location decisions of siblings can be generated by positive ρ and positive γ_0. Negative correlation between the location decisions of siblings can be generated by negative ρ and positive α. We can nevertheless identify ρ because of the fact that any correlation generated by ρ is unrelated to the sequence, whereas externalities caused by α and γ imply sequential interaction.

5.4 Method of Simulated Likelihood

The estimation relies on the maximum likelihood (ML) estimation in which the game is solved for an equilibrium outcome, a_h^*. Denote the observed family location configuration as $a_h^o \in \{0,1\}^{I_h}$. The ML problem 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 the likelihood function is that, given $\mathbf{X}$ and (α, β, γ), the location configuration is determined by $\boldsymbol{\varepsilon}$, and hence the distribution of $\boldsymbol{\varepsilon}$ determines the probability of a location configuration.

The probability term in (7) does not have an analytical form due to multidimensional integrals over the $\boldsymbol{\varepsilon}_h$ space. When the dimension of $\boldsymbol{\varepsilon}_h$ is more than two, computationally demanding numerical approximation, such as the quadrature method, is impractical. For high-dimensional integration, the maximum simulated likelihood (MSL) method, which utilizes Monte Carlo integration, has been developed in the literature. The most straightforward simulator for MSL is the crude frequency simulator. Given the model parameters, data, and the assumed distribution of $\boldsymbol{\varepsilon}_h$, the procedure takes a large number of random draws. For each random draw $\widetilde{\boldsymbol{\varepsilon}}_h$, an equilibrium location configuration, $\widetilde{a}_h^*$, is solved by backward induction. The probability in (7) is then obtained based on how many times the predicted equilibrium outcome coincides with the observed outcome out of the number of simulation draws. Although this simulator provides a consistent estimate of the probability, it is inefficient and requires a large number of simulation draws, and the estimation of our model is particularly computationally demanding because the game has to be solved for each simulation draw. To overcome this computation problem, we use the Monte Carlo integration method developed by Maruyama (2014).

5.5 Monte Carlo Integration with GHK Simulator

Maruyama (2014) develops the Monte Carlo integration method applicable to finite sequential games with perfect information, in which each player makes a decision by publicly-known exogenous decision order. The proposed method relies on two ideas. First, the MSL procedure utilizes the Geweke-Hajivassiliou-Keane (GHK) simulator, the most popular solution for approximating high-dimensional truncated integrals in 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 among truncation thresholds, which undermines the ground of the 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. Different realizations of unobservables that lead to different subgame-perfect equilibria but generate an observationally equivalent game outcome may therefore exist.

Figure 2 visualizes this point. The integration domain of $(\varepsilon_1, \varepsilon_2)$ that leads to the location outcome, (Near, Far), is not rectangular due to the strategic interaction between the two children, and hence the standard GHK simulator breaks down for this domain. The use of subgame perfection resolves this non-rectangular domain problem. 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 (2014) proves that the separate evaluation of the likelihood contribution for each subgame-perfect strategy profile makes it possible 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.

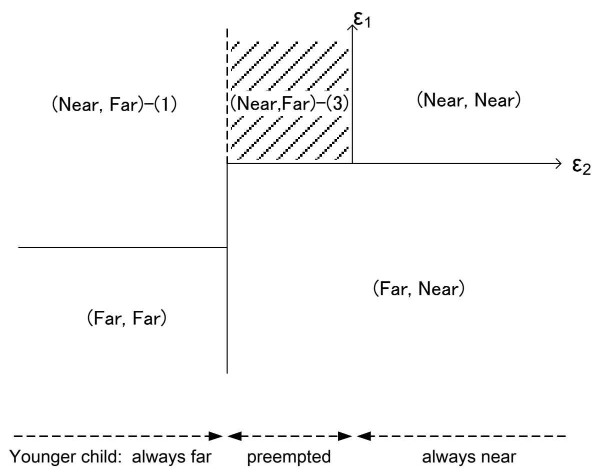

Figure 2: Dividing Observed Location Outcome into Strategy Profiles

Based on this logic, we obtain $\Pr_\rho \left[a_h^o = a_h^*(\mathbf{X}_h, \boldsymbol{\varepsilon}_h;\ \alpha, \beta, \gamma)\right]$ in (7) as follows. First, the complete list of SPNE profiles that rationalize the observed location configuration, a_h^o, is identified. Second, for each SPNE on the list, its associated probability is computed by the standard GHK simulator. In applying the GHK simulator, we sequentially calculate the truncation thresholds, i.e. the interval within which each child's random term, $\varepsilon_{i,h}$, has to fall for the SPNE to be realized. Lastly, we obtain the probability of the observed location configuration by summing the probabilities of all the SPNE profiles on the list.

6 Results

6.1 Probit Results

It is useful to first summarize the results from a simple probit model, which serves as a benchmark for extended specifications. In addition to its reduced-form interpretation, the probit specification offers a simple random-utility-model interpretation under the following assumptions: each child makes his/her location decision independently; his/her decision has no implications for the other children; and each child's unobserved preference component is distributed as an i.i.d. normal distribution.

The results are reported in Column [1] of Table 3. Parents who have a child living nearby tend to be old widowed parents with limited education and poor health who live in their owned home in an urban area. Proximate living is less likely for white parents and single but non-widowed fathers. Child variables are also relevant. Proximate living is less likely for older children (after controlling for parental age). Married children are less likely to live near their parents than single children. This is especially the case for daughters, probably because married daughters are more likely to live near their parents-in-law than married sons. This marriage effect is slightly offset by the presence of their children (grandparenting effect). Education moves children away from their parents; both *C_College* and *C_SomeCollege* have negative and significant coefficient estimates. These findings are consistent with Checkovich and Stern (2002), Byrne, Goeree, Hiedemann, and Stern (2009), and Compton and Pollak (2013).

[Insert Table 3: Estimated Parameters]

6.2 Specifications with Interactions among Siblings

The first step to building interdependence among siblings is to introduce correlation in the random term, $\{\varepsilon_{i,h}\}_{i=1}^{I_h}$. Specification [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 between siblings. The results shown in Column [2] of Table 3 testify to a significant positive correlation in the random term.

Now we explicitly introduce externality, first by including a constant altruism, $u_i^{\alpha}(N) = \alpha_0$. As shown in Column [3], we find a positive and significant estimate of α_0.[16] To confirm the robustness of this result, we estimate Specification [3] separately using each wave of the HRS from 1998 to 2010, and we find that the estimates of α and ρ are always positive and highly significant.

Whereas there is no substantial change in coefficient estimates when we compare Specifications [1] - [3], the goodness-of-fit improves over every step of elaboration. In terms of log L, a decent improvement results from incorporating externality α, but introducing correlation ρ makes the largest contribution. The proportion of correctly predicted observations, which are defined on the basis of location configuration with the highest predicted probability, also shows improvement. Although the simple probit model performs the best in predicting at the child level, it performs the worst at the family level because it ignores similarities and interactions among siblings.

6.3 Specification with Heterogeneous Externality

To introduce cooperation and allow for 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. Including the full set

[16] This specification leads to an even larger ρ than Specification [2] because altruism creates strategic substitutability, and omitting altruism forces the correlation term to capture this negative behavioral correlation, resulting in a smaller estimate of ρ.

of covariates in every term is impractical because it makes the model substantially flexible and makes the precise identification of parameters significantly difficult. We thus need a reasonably general yet parsimonious specification. Two guidelines have led us to our preferred specification. The first is the behavioral interpretation of each term: variables in u_i^α are supposed to be the determinants of innate altruism, and variables in u_i^γ should affect the cost and benefit of cooperation. Second, we adjust the sets of covariates by attempting various specifications. We exclude covariates whose coefficient parameters are always estimated with a large standard error and/or without statistical and economic significance.[17] We find our main results are reasonably robust across these modifications. Regarding correlation between 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 results of the full model. Compared to Specification [3], the goodness-of-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 3 compares the predicted distributions of the location configurations of Specifications [1] - [4] with the actual distribution in data, illustrating a step-by-step improvement in model prediction.

Correlation in the random term is stronger for siblings who are closer in age and of the same sex than for other siblings, indicating similarity in their preferences and environments. The altruism parameter, u_i^α, varies across children and families. According to the statistically significant coefficients in u_i^α, altruism is the strongest toward single mothers with limited education and poor health. The estimates of α_1 and α_2 are small and insignificant, indicating that what is important

[17] For example, including parental health in all three terms makes identification and convergence quite unreliable, and thus we take a conservative approach and not to include it in u_i^γ.

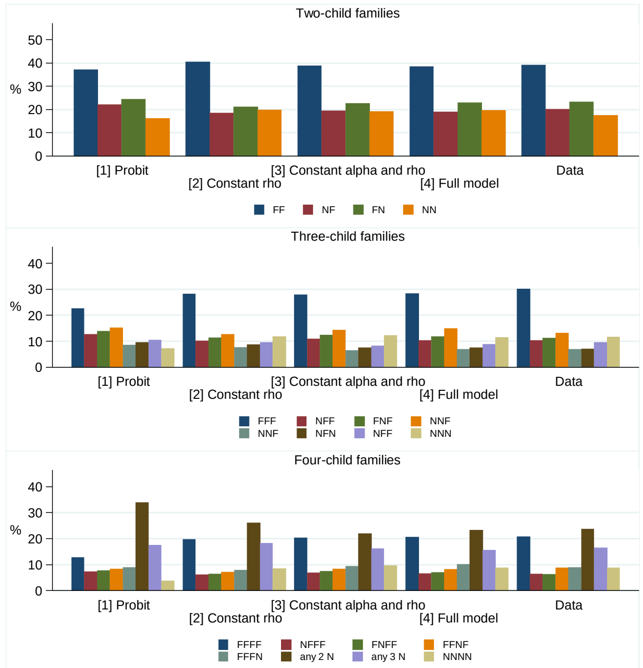

Figure 3: Predicted and Observed Location Configurations

to children is whether at least one child lives near the parent. Based on the distribution of X_i^α, the value of u_i^α ranges $[0.120, 1.370]$ with its mean 0.377.[18] The cooperation term, u_i^γ, also exhibits heterogeneity, ranging $[-0.046, 0.361]$ with its mean 0.199. The absolute size of u_i^γ is overall smaller than that of u_i^α. The negative and significant coefficient on C_age in u_i^γ indicates greater cooperation between younger children. One interpretation of this heterogeneity is that younger siblings have less experience of care provision, hence mutual assistance reduces the cost of providing care and attention. Alternatively, younger siblings may enjoy living close to each other. This interpretation has little to do with caregiving. Similarly to α_1 and α_2, the estimate of γ_1 indicates that having the third sibling nearby has no significant effect on the degree of cooperation. Thus,

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

externality does not distort the behavior of families in which more than two siblings choose to live near parents.

Heterogeneity in altruism and cooperation determines the extent of inefficiency and strategic interaction in each family. Inefficiency is larger in families with larger u_i^α and u_i^γ, and our results reveal that these are families with a single mother with limited education, poor health, and relatively younger children. Prisoners' dilemma is more likely in these families. First-mover advantage, on the other hand, is larger when u_i^α and is larger and u_i^γ is smaller. We find that relatively older children do not value cooperation greatly, and if their parent is a non-widowed single mother with limited education and poor health, the incentive to free-ride is large. The first child also has a large first-mover advantage.

The ranges of u_i^α and u_i^γ indicate that the vast majority of families show a certain altruism and cooperation. On the other hand, the range of $X_i^\beta \beta$ is $[-2.193, 0.861]$ with its mean -0.545. Given that the variance of ε_i is unity, the range of the three preference components suggests that although the two externalities are not negligible, the private cost component, $u_i^\beta + \varepsilon_i$, is the primary determinant of location decisions.

Examining the estimated coefficients on covariates that appear in both u_i^β and u_i^α offers additional insight. Whereas Specifications [1]-[3] find that parents with poor health are more likely to have their children nearby, this effect in u_i^β in Specification [4] 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 children are more concerned about the well-being of those parents and (2) because poor parental health increases children's net utility of living near the parent. The latter holds despite the expected large cost of care provision, probably because children value sharing time with parents who have shorter life expectancy. The education levels of children provide another contrast. Specifications [1]-[3] reveal a significant negative relationship between the child's education level and that child's propensity to live near their parents. Specification [4] confirms that this negative effect arises completely through the private utility component, u_i^β, probably reflecting the high opportunity cost for educated children of staying near the parent. The estimated coefficients in u_i^α show no evidence that well-educated children are less concerned about the well-being of their parents than children with limited education.

Lastly, our coefficient estimates offer a partial explanation for the birth-order asymmetry. We find a negative age effect in both 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 older siblings less. Both of these effects contribute to the lower tendency of older siblings to live near their parents, and these effects have nothing to do with first-mover advantage. At the same time, the significant estimates of altruism, u_i^α, and cooperation, u_i^γ, indicate the existence of sequential strategic interaction. In the next section, we quantify how much of the birth-order asymmetry in our data can be attributed to the first-mover advantage.

7 Counterfactual Simulations

7.1 Method

Counterfactual simulations allow us to quantitatively illustrate how the game structure and game outcomes vary across families under different settings. 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 equilibrium and optimal location configurations would be trivial, but we do not observe $\varepsilon_{i,h}$ in the data. We thus rely on Monte Carlo simulations, in which we generate simulated values of $\varepsilon_{i,h}$ 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 family h's 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 the solution function. Second, because this multidimensional integral does not have an analytical solution, a simulation method is necessary to numerically approximate the integral. Third, this simulation-based integration is complicated by strategic interaction among siblings. We evaluate this integral and the probability in the denominator by the Monte Carlo integration method explained in Section 5.5.

7.2 Normal-Form Game Structure

We first examine the simulated normal-form representation of corresponding simultaneous games, which provides useful information to understand the nature of the games played by American siblings. Table 4 characterizes the payoff matrices of two-child families by observed SPNE location configuration. The top panel reports whether siblings have dominant strategies in their 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 private cost, u_i^β, is so large that altruism, u_i^α, and cooperation, u_i^γ, has no influence on their decisions. It is trivial to show that, when every child has a dominant strategy, the equilibrium outcome of the simultaneous game is always achieved as an SPNE. Table 4 thus highlights limited strategic behavior in two-child families. The table also shows that when we observe (Far, Far) or (Near, Near) in the data, it almost always implies that both children in those families have a dominant strategy. The last column of Table 4 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 4: Characteristics of Simultaneous Normal-Form Games in Two-Child Families]

The bottom panel of Table 4 characterizes the Nash equilibrium of the simultaneous game, showing limited strategic behavior even more clearly. More than 99% of the 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 first-mover advantage.

7.3 Joint-Utility Optimal Location Configuration

We now turn to the joint-utility inefficiency of SPNE location configurations. We simulate the location configuration that maximizes each family's utility sum, which is compared in Table 5 with the actual location configuration by family size. There are many families in which the optimal number of children living near the parent is one or more but no child lives 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 in multi-child families, 18.3% more parents (= 32.5% − 14.2%) would have had at least one child living nearby had location decisions been made cooperatively.[19] On the other hand, the over-provision of proximate living exists among three- and four-child families but much less frequently.

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

The observed SPNE location configurations 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 6 presents the shares of these three groups by family size across different externality parameter values. Panel [1], which is based on the estimated parameters, shows that prisoners' dilemma is observed only for 2.0% of multi-child families, but that its presence increases with family size. More importantly, although 98.0% of multi-child families achieve Pareto efficiency, more than a quarter 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 those families achieve the joint-utility optimum. The simulation results reported in Panels [2]-[4] confirm the theoretical predictions: larger altruism, u_i^α, and cooperation, u_i^γ, lead to larger joint-utility inefficiency and Pareto inefficiency, and u_i^α explains a larger part of joint-utility inefficiency than u_i^γ, whereas a large u_i^γ is necessary for prisoners' dilemma to occur.

[Insert Table 6: Efficiency Type by Family Size]

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

Table 7 compares the shares of the efficiency types in two-child families by observed location configuration (the top panel) and by joint-utility optimal location configuration (the bottom panel). The first number in each cell represents its column share and the second number its row share. The table illustrates how prisoners' dilemma occurs. In families in the prisoners' dilemma situation, 70.5% 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. When (Far, Far) is joint-utility optimal, a family can always achieve it as an SPNE outcome. In this type of family, positive externality is very small compared to large private costs. When we observe (Near, Near) in the data, it is always the joint-utility optimal, whereas when we observe (Far, Far), it is joint-utility efficient only for 61.5% of those families, showing the importance of the under-provision of proximate living, rather than over-provision.

[Insert Table 7: Location Configurations and Efficiency Type in Two-Child Families]

7.4 First-Mover Advantage

To quantify the first-mover advantage, an ideal benchmark is the equilibrium outcome that arises in the simultaneous setup, but because simulating simultaneous games is not straightforward due to the multiplicity of equilibria, we instead employ a sequential game with reversed order (i.e. the youngest child makes a decision first and the oldest last). If order reversion does not affect the game outcome, it implies negligible first-mover advantage. The top panel in Table 8 compares the simulated location configurations of two-child families in the observed and reverse-order SPNE. The bottom panel investigates how reversing the order alters each child's utility. Overall, the sequential interaction is negligible. Reversing the order affects only 1.9% of two-child families. When it does affect a family, it is almost always the case that the SPNE outcome changes from (Far, Near) to (Near, Far), decreasing the first child's utility and increasing the second child's utility. The joint utility may or may not increase. If we double the degree of altruism, the share of families with a first-mover advantage increases from 1.9% to 9.3%.

Konrad, Künemund, Lommerud, and Robledo (2002) argue that observed birth-order asymmetry in location supports the first-mover advantage hypothesis. In our data, the number of two-child families that result in (Far, Near) and (Near, Far) is 658 and 564, respectively. The difference between these two numbers, 94 families, is the birth-order asymmetry in our data. An interesting question is how much of this difference is attributable to the first-mover advantage. As shown in Table 8, 7.8% of the 658 families with (Far, Near), or 51 families, change their location configuration from (Far, Near) to (Near, Far) after order reversion. If we assume that imposing reversed order affects twice as many families as imposing simultaneous move, removing the first-mover advantage should affect 26 (half of 51) families and result in 632 and 590 families with (Far, Near) and (Near, Far) configurations, respectively. The resulting difference of 42 families is the remaining birth-order asymmetry that is unexplained by the first-mover advantage. Hence, even though the first-mover advantage implied by our estimates is small, birth-order asymmetry in the US data is also small, thus the first-mover advantage explains roughly half of the asymmetry (42/94=45%).

[Insert Table 8: Reverse-Order SPNE in Two-Child Families]

8 Robustness and Validity of Results

8.1 Sensitivity Check

We have attempted various sample 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 (i.e. the birth order recorded in the data) contains measurement error, the estimated strategic effect may be biased toward zero. Although we expect little measurement error in the recorded birth order, birth order may not necessarily coincide with the actual order of location decisions. There may be a number of temporary moves when siblings are in their twenties, and some of those moves may become permanent; for example, younger siblings may make a permanent move before their older siblings complete post-graduate education. Conceptually, decision order in our model is a broader notion than the mere timing of migration, involving any credible commitment related to a permanent move, such as the choice of occupation and spouse. Hence, although younger siblings may make a permanent location choice before their older siblings, this does not necessarily contradict the use of birth order. Nevertheless, an important question is how well birth order approximates the true decision order, because the degree of measurement error in the decision order determines the size of the bias. Maruyama (2014) conducts a Monte Carlo experiment by applying the same estimation method for a sequential entry game, and reports that such bias tends to be marginal if the decision order is correctly specified in more than 90% of game observations.

One way to investigate potential bias resulting from misspecified 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 re-estimate the same model. Our main results are not affected by this additional restriction,[20] nor when we increase the minimum age difference to three years.

Are Only Children Special? We include one-child families in our sample because they aid identification; however, the results could be biased if only children differ considerably from children with siblings (after controlling for observable characteristics). To address this concern, we estimate our model without one-child families. We find that excluding one-child families makes the parameter estimates less robust. Standard errors tend to be larger with slightly worse goodness-of-fit. Although these findings suggest that one-child families play an important role in estimation, the results are consistent with our main results overall, indicating that our results are unlikely to be an 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 a cross-sectional approach, abstracting from the dynamic aspects of siblings' location decisions with the exception of birth order. For our estimates to be meaningful and credible, our empirical framework must be approximately consistent with the underlying data generating process. In particular, the explanatory variables used in estimation are taken from information recorded many years after children have made their lo-

[20]We find slightly larger estimates of externality as well as a smaller proportion of families experiencing inefficient family location. These two findings arise at the same time because although 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.

cation decisions. The results can be interpreted consistently with our behavioral model if all our explanatory variables were either observed or accurately predicted at the time children made those decisions. For this reason, we have carefully selected our independent variables such that they can be argued as being time-invariant or reasonably stable and predictable in the long run. Nevertheless, a number of factors may undermine the validity of our cross-sectional approach. 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 first left the parent (misspecification of the time frame), and current variables might have accumulated stochastic errors since the child makes the decision; 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 model that excludes parental health and marital variables, which may be endogenous events in later life. We find that the results of the simplified model are consistent with the full model overall, despite its poorer model fit. Counterfactual simulation results also remain similar. This finding provides some assurance that our main findings are not driven by the time inconsistency due to the time-variant variables.[21]

8.2 Alternative Behavioral Assumptions

We have so far centered our analysis on perfect-information sequential games. To investigate the appropriateness of this behavioral assumption, we discuss three alternative models.

[21] A more conservative view 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 for the empirical game-theoretic analysis of an inherently dynamic subject to start with a cross-sectional framework. The econometric literature on firms' market entry, for example, started with the analysis of a cross-sectional snapshot of market structures.

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

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

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 the validity of the perfect-information assumption, we estimate an incomplete-information model, maintaining the same functional-form specification 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. The conjectures underlie utility maximization because they affect one's expected utility. Child i's strategy, or decision rule, is denoted as $a_i\left(\varepsilon_i\right)$, 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, $\left\{a_i\left(\varepsilon_i\right)\right\}_{i=1,\ldots,I^h}$, constitutes a Nash equilibrium if:

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

We estimate this multivariate probit model using 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 that the distribution of the error component is independent across players, but our random terms are correlated between 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.[22] 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).[23] We find that this numerical iteration procedure is well-behaved as long as parameter values are not far from reasonable ranges. Because the mapping defined by (8), $f : \mathbf{a}^t \rightarrow \mathbf{a}^{t+1}$, 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, it is trivial to show the uniqueness as long as f is decreasing or moderately increasing (derivatives less than one) at any point of $\mathbb{R}^I$. 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

[22] Because the value of ε_i affects the net utility gain not only as the additive random 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 solution. Thus, the optimal strategy is solved by numerical iteration using the fact that expected net utility gain is given by a continuous increasing function of ε_i within the region of parameter values of our interest.

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

perfect-information model indicate that this condition is very likely to hold. 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 the information structure and the timing of decisions. A direct way to disentangle these two effects would be to estimate a perfect-information simultaneous game, but its estimation is not trivial due to the multiplicity of equilibrium. We instead estimate a perfect-information sequential game with reversed order, that is, we estimate our preferred model under the assumption that the youngest child makes the decision first and the oldest last. This experiment allows us to examine the relevance of our decision order assumption.

Model Fit Comparison Table 9 compares the goodness-of-fit of six alternative models with different behavioral assumptions: independent maximization under no externalities (Specification [1]), the non-cooperative perfect-information sequential model (Specifications [3] and [4]), joint maximization, the non-cooperative private-information model, and the non-cooperative perfect-information sequential model with reversed order. The last four columns compare different behavioral assumptions based on the same functional form assumption as Specification [4]. The table reports three comparison measures: the log likelihood values, the Akaike information criterion, and the percentage of correct prediction.

[Insert Table 9: Comparison of Alternative Behavioral Assumptions]

Overall, the comparison supports the use of a 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.[24] The private-information model fits the data better than the joint-decision model, but not as well as the perfect-information sequential model. Between these two lies the model with reversed order, supporting the use of both the perfect-information framework and birth order.[25] We also conduct the same comparison using simpler specifications and find that our conclusion is not affected.

9 Conclusion

We study externality and strategic interaction among adult siblings regarding their location decisions relative to their elderly parents, by estimating a sequential participation game that exceeds the scope of previous studies. We find a positive externality and strategic interaction. Siblings make location decisions non-cooperatively and their free-riding behavior results in the under-provision of proximate living to their elderly parents. Whereas the size of the strategic behavior is limited, the impact of the public good problem is striking; in multi-child families, 18.3% more parents would have had at least one child living nearby had location decisions been made cooperatively.

The complex nature of the subject requires us to employ a tractable framework: we rely on a 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

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

[25] Unlike our finding, Stern's (2014) study on the location choice game of siblings finds the empirical importance of private information. This may be because his game-theoretic framework is not as comprehensive as ours or because his model of private information is more general, in the sense that it incorporates private information in addition to unobservable common-knowledge heterogeneity.

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 the child with the smallest opportunity cost to assume caregiving obligations regardless of his/her willingness so that all other siblings can free-ride on this child. Historically, social norms and traditions in many countries have forced daughters, who supposedly have a smaller opportunity cost than sons, to fulfill caregiving obligations (see, e.g., Holroyd (2001) and Silverstein, Gans, and Yang (2006)). These social norms and traditions have served 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 undermined this mechanism and thus reduced the joint utility of families. The maximum joint utility can also be achieved by a transfer scheme from those who free-ride to those who provide care, but this option may be difficult in practice. Parents can utilize inheritance to enforce such a transfer, but this option is not available for socioeconomically disadvantaged parents, who face a particularly severe free-rider 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/her next-of-kin. Free-riding is thus likely to be more severe in jurisdictions that have legitim, such as Scotland, Japan, and, until recently, the US state of Louisiana. In general, policies that reduce the private cost of caring for elderly parents, such as tax benefits for carers, increase proximate living, but if the costs of such policies are financed by taxing other children equally, the overall welfare effect is ambiguous. The welfare effect of public support for parents is similarly ambiguous, depending on families' preferences and how such policies are financed.

Misleading conclusions may be drawn from future research if the free-rider problem identified in this study is not taken into consideration. Future research should direct its attention toward externality, the free-rider problem, and the under-provision of care and attention rather than to strategic interactions such as the first-mover advantage.

A Results of Selected Robustness Tests

[Insert Table A1: Robustness of Results]

References

[1] Antman, F. (2012), "Elderly care and intrafamily resource allocation when children migrate." *Journal of Human Resources*, 47(2), 331-363.

[2] Bernheim, B. D., A. Schleifer, and L. H. Summers (1985), "The strategic bequest motive." *Journal of Political Economy*, 93, 1045–1076.

[3] Bolin, K., B. Lindgren, and P. Lundborg (2008), "Informal and formal care among single-living elderly in Europe." *Health Economics*, 17(3), 393-409.

[4] Bonsang, E. (2009), "Does informal care from children to their elderly parents substitute for formal care in Europe?" *Journal of Health Economics*, 28(1), 143-154.

[5] Booth, A. L. and H. J. Kee (2009), "Birth order matters: the effect of family size and birth order on educational attainment." *Journal of Population Economics*, 22(2), 367–397.

[6] Börsch-Supan, A., L. J. Kotlikoff, and J. Morris (1988), "The dynamics of living arrangements of the elderly." Working Paper 2787, NBER.

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

[8] Checkovich, T. and S. Stern (2002), "Shared caregiving responsibilities of adult siblings with elderly parents." *Journal of Human Resources*, 37, 441-478.

[9] Compton, J. and R. Pollak (2013), "Proximity and coresidence of adult children and their parents in the United States: description and correlates." *IZA Discussion Paper*, No. 7431.

[10] Davis, J. N. (1997), "Birth order, sibship size, and status in modern Canada." *Human Nature*, 8(3), 205–230.

[11] Dostie, B. and P. T. Léger (2005), "The living arrangement dynamics of sick, elderly individuals." *Journal of Human Resources*, 40(4), 989-1014.

[12] Engers, M. and S. Stern (2002), "Long-term care and family bargaining." *International Economic Review*, 43(1), 73-114.

[13] Ettner, S. L. (1996), "The opportunity costs of elder care." *Journal of Human Resources*, 31(1), 189-205.

[14] Fontaine, R., A. Gramain, and J. Wittwer (2009), "Providing care for an elderly parent: interactions among siblings?" *Health Economics*, 18(9), 1011-1029.

[15] Gal-Or, E. (1985), "First mover and second mover advantages." *International Economic Review*, 26(3), 649-653.

[16] Hank, K. (2007), "Proximity and contacts between older parents and their children: a European comparison." *Journal of Marriage and Family*, 69(1), 157-173.

[17] Hiedemann, B., M. Sovinsky, and S. Stern (2014), "Will you still want me tomorrow? The dynamics of families' long-term care arrangements." mimeo.

[18] 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.

[19] Hoerger, T. J., G. A. Picone, and 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.

[20] Holroyd, E. (2001), "Hong Kong Chinese daughters' intergenerational caregiving obligations: a cultural model approach." *Social Science and Medicine*, 53, 1125-1134.

[21] Hotz, V. J., K. McGarry, and E. Wiemers (2010), "Living arrangements of mothers and their adult children over the life course." mimeo.

[22] Johar, M. and S. Maruyama (2011), "Intergenerational cohabitation in modern Indonesia: filial support and dependence." *Health Economics*, 20(S1), 87-94.

[23] 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.

[24] Johar, M. and S. Maruyama (2014), "Does coresidence improve an elderly parent's health?" *Journal of Applied Econometrics*, 26(9), 965-983.

[25] 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.

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

[27] Lilly, M. B., A. Laporte, and P. C. Coyte (2010), "Do they care too much to work? The influence of caregiving intensity on the labour force participation of unpaid caregivers in Canada." *Journal of Health Economics*, 29(6), 895-903.

[28] Maruyama, S. (2014), "Estimation of finite sequential games." *Journal of Econometrics*, 178(2), 716-726.

[29] Maruyama, S. (2015), "The effect of coresidence on parental health in Japan." *Journal of the Japanese and International Economies*, 35, 1-22.

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

[31] Matthews, S. H. and T. T. Rosner (1988), "Shared filial responsibility: the family as the primary caregiver." *Journal of Marriage and the Family*, 50, 185-195.

[32] 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.

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

[34] 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.

[35] Sakudo, M. (2008), "Strategic interactions between parents and daughters: co-residence, marriage, and intergenerational transfers in Japan." mimeo, University of Chicago.

[36] Schmidt-Dengler, P. (2006), "The timing of new technology adoption: the case of MRI." mimeo.

[37] 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), 1068-1084.

[38] Stern, S. (2014), "O brother, where art thou? We need your help." mimeo.

[39] Sulloway, F. J. (2007), "Birth order and inteliggence." *Science*, 316, 1711-1712.

[40] van den Berg, B., W. B. Brouwer, and M. A. Koopmanschap (2004), "Economic valuation of informal care." *European Journal of Health Economics*, formerly HEPAC, 5(1), 36-45.

Table 1: Sibling Location Configurations by Birth Order with Implied Shares under Independence

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

N of children living near	Total share	Detailed location configurations with birth order	Implied share under independence (p=0.404)
0	48.7%	Far: 48.7%	59.6%
1	51.3%	Near: 51.3%	40.4%

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

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

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

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

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

N of children living near	Total share	Detailed location configurations with birth order	Implied share u/ independence (p=0.404)
0	20.5%	FFFF: 20.5%	12.6%
1	30.4%	NFFF: 6.7% ; FNFF: 6.3% ; FFNF: 8.5% ; FFFN: 8.9%	34.2%
2	24.2%	NNFF: 3.1% ; NFNF: 3.4% ; NFFN: 4.7% ; FNNF: 3.3% ; FNFN: 4.9% ; FFNN: 4.8%	34.8%
3	16.3%	NNNF: 3.7% ; NNFN: 4.4% ; NFNN: 3.8% ; FNNN: 4.4%	15.7%
4	8.7%	NNNN: 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 oldest child, e.g., “FFN” indicates the location configuration of a three-child family in which only the youngest child lives near the parent. “N” includes coresidence. As a benchmark, the last column shows the shares computed under the assumption that each child makes a location decision independently and chooses “N” with probability 0.404 (=overall average).

Table 2: Definition and Summary Statistics of Variables

Variable	Definition	Mean	Std. Dev.
Outcome			
Near	=1 if the child lives with or within 10 miles of the parent	0.404	0.491
Parent			
P_cohab	=1 if the respondent lives with a partner, regardless of marital status (reference group)	0.447	0.497
P_father_widow	=1 if the respondent is a father living with no partner, widowed	0.062	0.242
P_father_nonwidow	=1 if the respondent is a father living with no partner, not a widow (e.g. separated/divorced)	0.057	0.232
P_mother_widow	=1 if the respondent is a mother living with no partner, widowed	0.302	0.459
P_mother_nonwidow	=1 if the respondent is a mother living with no partner, not a widow (e.g. separated/divorced)	0.132	0.339
*P_age**	Parent's age	71.939	7.576
*P_white^*	=1 if race is white	0.838	0.369
*P_healthy**	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. The larger the healthier.	-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
P_Geo_HighPop	=1 if lives in a metro area of 1 million population /more (reference group)	0.441	0.476
P_Geo_MedPop	=1 if lives in a metro area of 250,000 to 1 million population	0.250	0.433
P_Geo_LowPop	=1 if lives in a metro area of fewer than 250,000 population or non-metro area	0.283	0.450
P_Geo_missing	=1 if geographical information is missing	0.026	0.160
P_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
C_age_difference	Age difference between child *i* and child *j* (absolute value)		
C_sex_difference	=1 if child *i* and child *j* are of different sex; 0 otherwise		
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

	[1] Probit		[2] Constant ρ; u^{α}=0 (no altruism, constant correlation)		[3] Constant u^{α} and ρ (constant altruism and correlation)		[4] Full model	
	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_healthy	-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
P_Geo_MedPop	-0.009	0.022	-0.009	0.027	-0.007	0.028	-0.007	0.029
P_Geo_LowPop	-0.091***	0.021	-0.091***	0.026	-0.095***	0.027	-0.095***	0.028
P_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} (altruism) 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 u_i^{γ} (cooperation))							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 (correlation))			0.238***	0.014	0.361***	0.021	0.476***	0.035
C_age_difference							-0.008**	0.003
C_sex_difference							-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 section 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, *P_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: Characteristics of Simultaneous Normal-Form Games in Two-Child Families

Who has dominant strategy:	Observed location configuration (SPNE) (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%

Equilibrium patterns in simultaneous normal-form games:	Observed location configuration (SPNE) (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 5: Observed and Family-Optimal Location Configurations by Family Size

	Number of children living near the parent in SPNE (observed location configuration)					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: 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 Panel [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 7: Location Configurations and Efficiency Type in Two-Child Families

	Efficiency type: Joint-utility optimal	Joint-utility suboptimal but Pareto efficient	Prisoners' dilemma	Total
Observed location configuration (SPNE)				
(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 the location configuration that maximizes the sum of siblings' utility. Results are based on empirical distribution with Monte Carlo simulation for the error terms with 1,000 random draws.

Table 8: Reverse-Order SPNE in Two-Child Families

Location configuration of reverse-order SPNE:	Observed location configuration (SPNE): (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:	Observed location configuration (SPNE): (Far, Far) N=1,125	(Far, Near) N=658	(Near, Far) N=564	(Near, Near) N=493	Total	Total when u_i^{α}×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. The second child's utility may also decrease, but these two events never occur at the same time (i.e. there is no second-mover advantage). Results are based on empirical distribution with Monte Carlo simulation for the error terms with 1,000 random draws.

Table 9: 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:	$u^{\alpha}=u^{\gamma}=\rho=0$ (Model [1])	u^{α}, ρ constant; $u^{\gamma}=0$ (Model [3])	heterogeneous externality (Model [4])	heterogeneous externality	heterogeneous externality	heterogeneous externality
Log L	-11,951.0	-11,759.6	-11,693.9	-11,957.7	-11,727.8	-11,711.4
# 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 $u^{\alpha}=u^{\gamma}=\rho=0$, there is no 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 largest predicted probability among all possible location configurations.

Table A1: Robustness of Results

	[4] Full model	[5] Without siblings of age difference < 2 years	[6] Multi-child families	[7] Simplified model without potentially endogenous variables	
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_healthy	-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
P_Geo_MedPop	-0.007*	-0.024	-0.004	P_Geo_MedPop	-0.002
P_Geo_LowPop	-0.095***	-0.097***	-0.097***	P_Geo_LowPop	-0.100***
P_House	0.096***	0.100***	0.105***	P_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_healthy	-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***
C_age_difference	-0.008**	-0.003	-0.009***	C_age_difference	-0.008
C_sex_difference	-0.114***	-0.108***	-0.122***	C_sex_difference	-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, *P_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*.