O Brother, Where Art Thou? We Need Your Help

Working Paper 2014-8

Steven Stern

June 9, 2014

Stony Brook University

O Brother, Where Art Thou? We Need Your Help

Steven Stern

University of Virginia

June 9, 2014

Abstract

I generalize a model of children strategizing where to live by Konrad et al. (2002) (KKLR) to include more variation in family size and to allow for uncertainty in how each family member behaves. The inclusion of uncertainty has large effects on the predictions of a KKLR-type model. I find that the KKLR model does not explain American data well, but an alternative model with the same reduced form does explain the data. Finally, the estimated amount of uncertainty has larger welfare implications than imperfections associated with the sequential structure of the KKLR game.

1 Introduction

Geographic proximity is a powerful predictor of provision of informal care by children for elderly parents (Ikels, 1983; Litwak and Kulis, 1987; Stern, 1995; McGarry, 1998; Hiedemann and Stern, 1999; Engers and Stern, 2002; Bonsang, 2009; Compton and Pollak, 2009; Hiedemann, Sovinsky, and Stern, 2014). While the majority of adults in the United States live close to their mothers (Compton and Pollak, 2009), geographic dispersion within families is still an important policy issue. Most of the literature (Ikels, 1983; Litwak and Kulis, 1987; McGarry, 1998; Hiedemann and Stern, 1999; Engers and Stern, 2002; Bonsang, 2009) treats geographic distance between elderly parents and their children as exogenous. Stern (1995) and Hiedemann, Sovinsky, and Stern (2014) consider the possibility of endogeneity and find evidence supporting the exogeneity assumption. However, Litwak and Longino (1987), Silverstein (1995), Rogerson, Burr, and Lin (1997), and Compton and Pollak (2009) all provide evidence suggesting the importance of understanding mobility better.

Konrad et al. (2002) (KKLR) presents a model of a family with two children and parents. Each of the children want a child to live near the parents so that, when the parents grow old and need assistance, at least one of the children will be available to provide such help. However, both prefer that it is the other child who provides that help. KKLR suggests a sequential game between the children where the oldest child decides where to live (nearby or far), and then the youngest decides.[1] The oldest child has the advantage of being able to commit to not being available, thus providing the younger child with strong incentive to live nearby. KKLR uses German data to show

[1] Antmann (2012) suggests that living far may not get a child "off the hook;" they still may be expected to provide financial assistance.

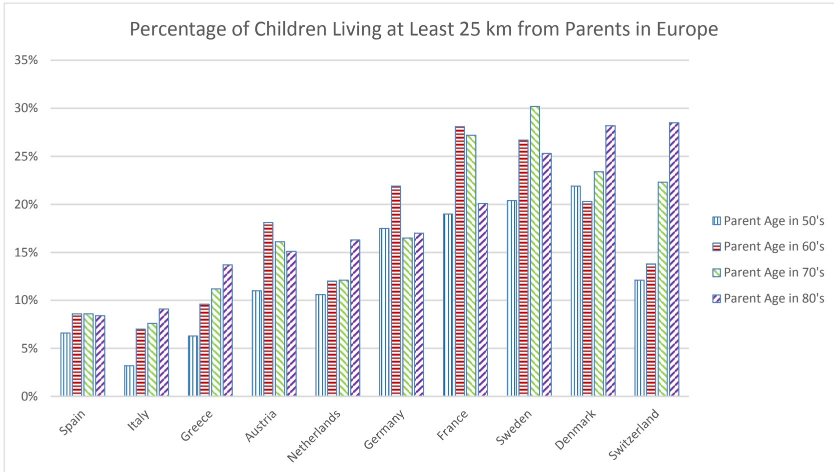

Figure 1: Percentage of Children Living at Least 25 km from Parents in Europe

that older children are more likely to live far than younger children and are more likely to live far than only children. Using a similar model, Rainer and Siedler (2005) (RS) provides some empirical results based on the same German data used in KKLR and some American data showing that children with siblings are more likely to live far than only children, but it finds no birth order effect; it is not clear what generates the differences in empirical results between the two papers.[2] Rogerson, Weng, and Lin (1993) finds results consistent with KKLR.

It is somewhat surprising that a theory predicting that the oldest child moves far would gain traction. Besides the mixed empirical results discussed above, there is a long-held belief (Treas, Gronvold, and Bergtson, 1980; Horowitz, 1985; Brody, 2004) that the oldest daughter is most likely to become the primary caregiver. Also, in many other countries, there is a strong tradition for the oldest son to be the primary caregiver.[3] Hank (2007, table 2) provides information for Europe on the proportion of children living far from parents, disaggregated by age and country, displayed in Figure 1. Unfortunately, no information is provided on the effects of birth order. However, the displayed variation both over age and across countries suggests a) the impact of some cultural differences (Hank, 2007), and b) steady movement away from parents as they age.

This paper presents a model with multiple children and some uncertainty for each child about how younger siblings will behave. Because of the implied randomness, it is possible for a child to make a decision expecting a younger sibling to stay near the parent and then be surprised; thus the title of the paper. The paper generalizes from KKLR and RS in that it includes families with more than two children and it allows for uncertainty. It generalizes from Johar and Maruyama (2012) (JM) with the inclusion of uncertainty. On the other hand, JM construct a model

[2] RS provides some imprecise arguments about why there might be no birth order effect, but the model appears to imply important birth order effects. Holmlund, Rainer, and Siedler (2009b) suggests a model similar to KKLR with the claim that any birth order effect is small. But, the model seems somewhat rigid and, more importantly, it implies much more sharing of caregiving activities than is observed in the data (Checkovich and Stern, 2002).

[3] See, for example, Arnold, Choe, and Roy (1998) for India, Bian, Logan, and Bian (1998) for China and Kureishia and Wakabayashi (2010) for Japan. Martin (1990) presents similar patterns for Japan, Korea, China but suggests they are all changing to look more Western.

Table 1: Selection Criteria

Cause	# Rejected	# Remaining
Initial Size		2813
No Child	564	2249
Missing Child	53	2196

which facilitates the discussion of welfare questions cleaner than mine.

Section 2 discusses the construction and characteristics of the data including an analysis of the distribution of child distance from parent. Section 3.1 presents a model similar to KKLR and RS to set notation and discuss problems. Randomness is added to the model in Section 3.2. Included in this section is the construction of the likelihood function used to estimate the parameters of the model. Section 4 presents parameter estimates and discusses how they compare to the literature and their implications. It turns out that the model presented in Section 3.2 does not explain the data, and a slightly different alternative model, not separately identified from the original model, is proposed. A conclusion finishes the paper.

2 Data

2.1 Source

I start off using the first year of data, 1993, from AHEAD/HRS from Byrne et al. (2009).[4] I further exclude families with no children or where a child is missing in the first year. I exclude families with no children because they have no information relevant for our model, and I exclude families with missing children because the likelihood contribution for such families is poorly defined.[5] The number of observations lost are displayed in Table 1.

There are 201 observations where the children are not included in the sample in descending order by age. I estimate the model three different ways to measure the sensitivity of results with respect to this problem. First, I assume that the children are included in descending age order, and, in those cases where the reported ages are not monotone declining, I infer that there is a mistake in one of the reported ages. Second, I exclude the 201 observations not in descending order. Third, I reorder the observations not in descending order so that they are in descending order. As discussed in Section 4, the parameter estimates are not sensitive to how one deals with this problem. The explanatory variables to be used are those that might affect the decision-making process of each child at the time that the child decides where to live. In fact, there is no one time that an individual makes such a decision. For example, Kennan and Walker (2011) shows that a large fraction of people move more than once and many return to where they grew up. However, Figure 2, based on US Census (2012), shows that migration rates fall precipitously as people age. I ignore all of the dynamics of mobility and pretend that the observed location in 1993 was a one-time decision made at some common age. RS also ignores the dynamics of mobility while KKLR considers the possibility

[4] JM uses a later wave of the same data. In theory, this could suffer from mobility among children and parents once parents need help. Hiedemann, Sovinsky, and Stern (2014) shows that such mobility is not very common, while Bianchi, McGarry, and Seltzer (2010) finds some evidence for noticeable mobility tied to informal care.

[5] For such families, one would need to integrate over the joint density of the missing child's explanatory variables and location decision.

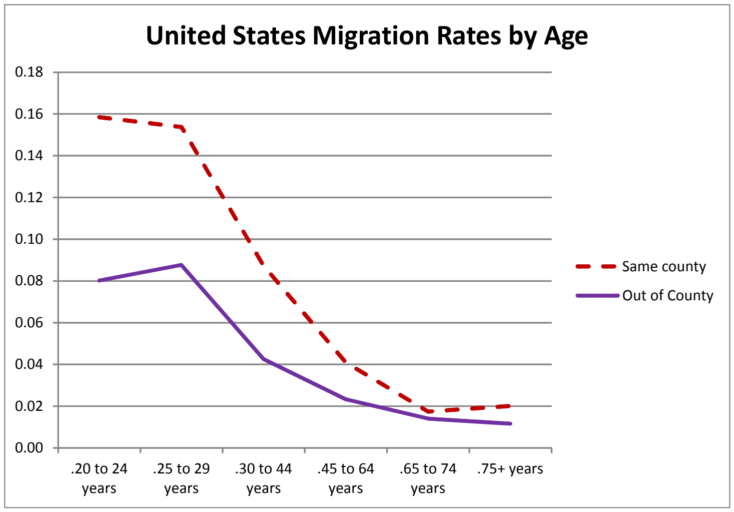

Figure 2: United States Migration Rates by Age

that the parent moves late in life if beneficial. Even if one wanted to model dynamics, it would not be obvious how to identify any of the parameters associated with dynamics using available data.

The parent explanatory variables include race (black or white), Hispanic, Protestant, and Catholic. All of these do not change over the course of time within a family. I exclude variables such as measures of the parent's health (ADL's), age of the parent, or gender of the parent because none of these were known at the decision-making time. The child explanatory variables include gender, age, education, and marital status. Age should be interpreted as a cohort effect. Marital status should be interpreted as a proxy for marriage preferences at the decision-making time. As with location, marital status is dynamic (e.g., divorces are common), and I ignore those dynamics as well. KKLR uses some variables such as parents' health and marital status and existence of grandchildren to explain geographic distance despite recognizing the point that they were not known at the decision-making time. RS does not use as many problematic explanatory variables but still uses marital status of the parent. Hank (2007), [6]Holmlund, Rainer, and Siedler (2009a) and JM use many problematic variables.[7] Both KKLR and RS use a measure of economic well-being of the place where the children grew up, and such a variable plays an important role in both of their discussions. I have no such variable available.

2.2 Suggestive Statistics

The models in KKLR and RS imply that, other things equal, the youngest child should live near the parents and all of the other children should live far away. KKLR and RS model families with at most two children, but they easily generalize to families with more than two children. Table 3 displays the density of family geography disaggregated

[6]Obviously, the parents know their own gender. However, the meaning of gender in the first year of the data is a function of the survival status of the two parents in 1993. The survival status of the parents in 1993 was not known at the decision-making time.

[7]Silverstein (1995) and Rogerson, Burr, and Lin (1997) show changes in parent health and changes in parent marital status affect distance between parent and child. These changes obviously were not anticipated when the adult children were initially leaving home.

Table 2: Explanatory Variable Moments

Parent Characteristics			Child Characteristics		
Variable	Mean	Std Dev	Variable	Mean	Std Dev
# Obs	2196		# Obs	4965	
Black	0.085	0.275	Female	0.516	0.500
Hispanic	0.038	0.191	Age	49.290	8.220
Protestant	0.652	0.475	Education	13.887	2.390
Catholic	0.243	0.427	Married	0.740	0.439

Table 3: Family Geography Density

Families with 1 Child

Child 1	# Families
1	272
2	237
3	44

Families with 2 Children

Child 1	Child 2	# Families
1	1	345
1	2	160
2	1	125
2	2	124
3	1	31
3	2	24
3	3	4
Proportion Consistent w/ Rainer & Siedler:		0.197

Families with 3 Children

Child 1	Child 2	Child 3	# Families
1	1	1	161
1	1	2	54
1	2	1	60
1	2	2	39
2	1	1	68
2	1	2	39
2	2	1	31
2	2	2	34
3	1	1	27
3	1	2	8
3	2	1	6
3	2	2	6
3	3	1	1
3	3	2	2
3	3	3	1
Proportion Consistent w/ Rainer & Seidler:			0.101

Families with 4 Children

Child 1	Child 2	Child 3	Child 4	# Families
1	1	1	1	58
1	1	1	2	31
1	1	2	1	18
1	1	2	2	15
1	2	1	1	14
1	2	1	2	15
1	2	2	1	11
1	2	2	2	6
2	1	1	1	25
2	1	1	2	10
2	1	2	1	8
2	1	2	2	11
2	2	1	1	7
2	2	1	2	8
2	2	2	1	9
2	2	2	2	9
3	1	1	1	9
3	1	1	2	1
3	1	2	1	2
3	1	2	2	4
3	2	1	1	4
3	2	1	2	3
3	2	2	1	5
3	2	2	2	6
3	3	2	1	2
3	3	2	2	1
3	3	3	1	2
Proportion Consistent w/ Rainer & Seidler:				0.105

Notes:
1) 1 = > 10 miles; 2 = within 10 miles; 3 = living with parent
2) Rows in olive are consistent w/ Rainer & Seidler

by family size in my sample with those elements consistent with KKLR and RS highlighted. For example, among families with 2 children, 160 out of 813 have child 1 (the older child) living far away and child 2 living nearby. These are the families being modeled in KKLR and RS, and they represent 19.7% of the 2-child families. The proportion is approximately half as large for families with 3 or 4 children.

Figure 3 and Table 4 provide more information about the joint density of geography. Figure 3 shows that the oldest child is as likely to live nearby as the youngest child for each family size. As in RS, Holmlund, Rainer, and Siedler (2009a), and JM, this is not consistent with KKLR and, in fact, contradicts the empirical results KKLR presents for Germany. Also, among those families where at least one child lives nearby, a significant proportion, increasing with family size has more than one child close by (in each # Children group, compare the third bar to the fourth bar).

Table 4 tests H_0 : children are randomly distributed geographically against H_A : children distribute themselves according to the KKLR and RS models. In particular, define $Z_{ij} = 1$ iff child j from family i lives nearby. Under

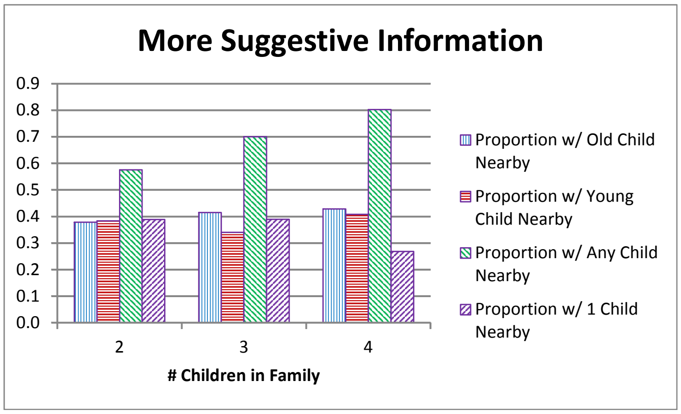

Figure 3: More Suggestive Information

Table 4: Test of KKLR & RS

	# Children		
	2	3	4
Proportion Consistent w/ KKLR & RS	0.197	0.101	0.105
Probability Lives Far Away	0.619	0.636	0.611
Probability Lives Nearby	0.381	0.364	0.389
Probability of KKLR & RS Under H0	0.236	0.231	0.238

H_0, assume that

$$Z_{ij} \sim indBernoulli\left(p_{jJ_i}\right)$$

where p_{jJ_i} is allowed to vary over both j and family size J_i. The first row displays the proportion of families (from Table 3) with the youngest child living nearby and all other children living far away. The second and third rows show the proportions living far away and nearby, respectively. The last row shows the proportion of families that would have followed the KKLR/RS pattern under H_0,

$$q = p_{J_iJ_i} \prod_{j=1}^{J_i-1} \left(1 - p_{jJ_i}\right).$$

A formal test would involve finding a critical value greater than q such that one would reject H_0 iff the statistics in the first row are larger than q. In fact, for every family size, the proportion that actually followed the KKLR/RS pattern is less than would have predicted under H_0. Thus, it is clear that H_0 should not be rejected in favor of H_A.

Both KKLR and RS provide evidence that children with siblings are more likely to live far than those without siblings. Table 5 sheds some evidence on this hypothesis using my data. Among families with 1 child, 49.2% have children living far. Among families with 2 children, 38.9% have the pattern proposed by the KKLR/RS model (ignoring birth order effects implied by the model) while the null hypothesis that family size has no effect (with the alternative being the KKLR/RS model ignoring birth order effects) would imply that 50.0% should follow the pattern. Since the theoretical probability is greater than the sample proportion, one should not reject the null

Table 5: Test of KKLR & RS

	Sample Proportion	Theoretical Proportion	Std Error
Prop. Of Families w/ 1 Child Living Far	0.492		
Prop. Of Families w/2 Children Living Far:			
1 of 2 Living Far	0.389	0.500	
At Least 1 Living Far	0.813	0.742	0.015
Prop. Of Families w/3 Children Living Far:			
2 of 3 Living Far	0.389	0.369	0.021
At Least 1 Living Far	0.920	0.869	0.015
Prop. Of Families w/4 Children Living Far:			
3 of 4 Living Far	0.330	0.242	0.025
At Least 1 Living Far	0.946	0.933	0.015

Note: elements in olive are statistically significant at the 5% level.

hypothesis in favor of the KKLR/RS models. If one considers a much weaker implication of the KKLR/RS models, families with multiple children should be more likely to have some living far, then the sample proportion is 81.3%, while the theoretical probability is 74.2%. The standard error under H_0 (no effect) is 1.5%; thus, one would reject H_0 in favor of the KKLR/RS models. The same pattern of results occurs for families with 3 children. For families with 4 children, one rejects H_0 (no effect) against H_A : all but one of the children live far but accepts H_0 (no effect) against H_A : at least one of the children live far. On the whole, this is not very supportive of the KKLR/RS models.

3 Model

3.1 Deterministic (KKLR & RS) Model

Consider a family with J children indexed (in order of age, oldest first) by $j = 1, 2, .., J$. Each child must make a decision whether to live near the parent or move far away. Let $y_j = 1$ iff child j lives near the parent. The children make decisions sequentially by age. At the time of j's decision, denote the probability that at least one of the other children will live nearby as P_j.[8] At the time of j's decision, children $k = 1, 2, .., j-1$ have already made decisions y_k and children $k = j+1, j+2, .., J$ have not yet made decisions. Let $\widetilde{p}_k$ be the probability that $k > j$ chooses to live nearby as a function of j's choice. Then

$$P_j = 1 - \left[\prod_{k=1}^{j-1}(1 - y_k)\right]\left[\prod_{k=j+1}^{J}(1 - \widetilde{p}_k)\right];$$

i.e., if $\exists k < j : y_k = 1$, then $P_j = 1$, and, if $y_k = 0\ \forall k < j$, then

$$P_j = 1 - \left[\prod_{k=j+1}^{J}(1 - \widetilde{p}_k)\right].$$

[8] In this subsection, P_j turns out to be either 0 or 1, thus not really a probability. However, later in the next subsection the probability notion is useful.

From j's point of view, there are three possible outcomes:

1. No child lives nearby. This happens with probability $(1 - y_j)(1 - P_j)$, and it delivers utility u_{00}.

2. Child j lives nearby. This happens iff $y_j = 1$, and it delivers utility $u_1(x_j)$ where x_j is a vector of characteristics of j affecting the net benefit (or loss) associated with living nearby.

3. Child j lives far away, but one of the other children live nearby. This happens with probability $(1 - y_j) P_j$, and it delivers utility u_{01}.

Thus, the expected value of j's utility at the decision-making time is

$$U_j(x_j, p_j, y_j, Y_j) = (1 - y_j)((1 - P_j) u_{00} + P_j u_{01}) + y_j u_1(x_j).$$

In general,

$$y_j = 1 \text{ iff } u_1(x_j) > (1 - P_j) u_{00} + P_j u_{01} \tag{1}$$

The problems in KKLR and RS are deterministic with

$$u_1 < u_{01}, u_{00} < u_{01}. \tag{2}$$

Note that equation (1) implies that

$$P_j = 1 \Rightarrow y_j = 1 \text{ iff } u_1(x_j) > u_{01}.$$

However, by equation (2), this is false. Thus,

$$P_j = 1 \Rightarrow y_j = 0,$$

$$y_j = 1 \Rightarrow y_k = 0 \ \forall k > j.$$

JM, besides allowing for more than 2 children, adds randomness to the model in the sense of part of x_j not being observed in the data. It assumes that all relevant information is known by everyone in the family; thus the family plays a game with perfect information. This is similar to many of the models with multiple equilibria in the empirical IO literature (e.g. Tamer, 2003; or Fontaine, Gramain, and Wittwer, 2009 in a model of child caregiving). However, the assumption of sequential decision-making in KKLR, RS, and JM avoids the multiple equilibrium problem. On the other hand, the perfect information assumption may make it difficult to explain the behavior in some observed families. In particular, with perfect information, if one observes a family where the oldest child behaves in a low probability way, then the model must explain such behavior in terms of a low-density realization of the unobserved component of x_j. In a model with some private information, besides the same source of randomness, there is also the possibility that the older child incorrectly predicted how the younger child would behave. I find this to be empirically important as is seen in Section 4.3.

3.2 Model with Randomness

3.2.1 Family Behavior

Now consider generalizing the models in KKLR and RS so that a) there is some unobserved heterogeneity in family preferences and some randomness. The unobserved heterogeneity is the component of preferences common knowledge to all family members but not observed in the data. On the other hand, each family member has some component of preferences known only by him/her; other family members know only the distribution of these "random" components.

As in Section 3.1, there are three states of the world for each child:

1. No child lives near the parent. In this case, child j receives utility

$$u_{00j} = \varepsilon_{0j} \tag{3}$$

   where ε_{0j} is a random component known only by j. There is no need to consider a more complex structure for this choice because all that matters is the utility of each choice relative to a base choice. I have chosen this choice as the base choice.

2. One of the other children lives near the parent and child j lives far away. In this case, child j receives utility

$$u_{01j} = \alpha + e_{0j} + \varepsilon_{1j} \tag{4}$$

   where α is a common value across children and families of having a child live near the parent, e_{0j} is a common-knowledge, child-specific component of the same value, and ε_{1j} is a random component known only by j.

3. Child j lives near the parent. In this case, child j receives utility

$$\begin{aligned} u_1(x_j) &= u_{01j} + g(x_j) + e_{1j} + \varepsilon_{2j}; \\ g(x_j) &= x_j\beta \end{aligned} \tag{5}$$

   where e_{1j} is a common-knowledge, child-specific component associated with child j living nearby and ε_{2j} is a random component known only by j.

There are two possible interpretations of the inclusion of the random components $\varepsilon_j = (\varepsilon_{0j}, \varepsilon_{1j}, \varepsilon_{2j})'$. One possibility is that there are some random events that affect the preferences of each child and are revealed only at the time that each child must make a location decision. Thus, even though child j may observe ε_{j+s} for $s > 0$ at the time that child $j+s$ makes a decision, child j does not observe ε_{j+s} at the time child j makes a decision. A second possibility is that each child knows his/her own ε at an early age and either chooses not to share it with siblings or can not do so in a credible way. Better understanding of this issue is left for later work.

In one way, JM is more general than my model. In particular, while I treat α as a constant, JM treats it as a function of the number of children living near. Also, $g(x_j)$ is allowed to depend upon the number of children living near. I perform a Lagrange Multiplier test to see if the extra generality is worthwhile and report the results in Section 4.

For the joint distribution of the errors, define $e_j = (e_{0j}, e_{1j})'$, and assume that

$$\begin{aligned}
\varepsilon_j &\sim iidN\left(0, \sigma_\varepsilon^2 I\right); \\
e_j &= v_0 + v_{1j}; \\
v_0 &\sim iidN\left(0, \Omega_0\right); \\
v_{1j} &\sim iidN\left(0, \Omega_1\right).
\end{aligned} \tag{6}$$

Without loss of generality, one can restrict $\sigma_\varepsilon^2 = 1$ because, as is true in any discrete choice problem, there is an unidentified scaling parameter.

Next, one can solve each child's problem backward recursively. Child J's problem is

$$y_J = 1 \text{ iff } u_1(x_J) > (1 - P_J) u_{00J} + P_J u_{01J}$$

$$\begin{aligned}
&\text{iff } \alpha + e_{0J} + \varepsilon_{1J} + g(x_J) + e_{1J} + \varepsilon_{2J} \\
> \; &(1 - P_J)\varepsilon_{0J} + P_J\alpha + P_J e_{0J} + P_J \varepsilon_{1J} \\
&\text{iff } (1 - P_J)\alpha + g(x_J) + (1 - P_J) e_{0J} + e_{1J} \\
> \; &(1 - P_J)(\varepsilon_{0J} - \varepsilon_{1J}) - \varepsilon_{2J}.
\end{aligned}$$

If $y_k = 0 \; \forall k < J$, then $P_J = 0$ which implies that

$$y_J = 1 \text{ iff } \alpha + g(x_J) + e_{0J} + e_{1J} > \varepsilon_{0J} - \varepsilon_{1J} - \varepsilon_{2J}.$$

Alternatively, if $\exists k < J : y_k = 1$, then $P_J = 1$ which implies that

$$y_J = 1 \text{ iff } g(x_J) + e_{1J} > -\varepsilon_{2J}.$$

Thus, the probability that J lives nearby if $y_k = 0 \; \forall k < J$ is

$$\widetilde{p}_J = \Pr\left[\alpha + g(x_J) + e_{0J} + e_{1J} > \varepsilon_{0J} - \varepsilon_{1J} - \varepsilon_{2J}\right],$$

and, if $\exists k < J : y_k = 1$, it is

$$\widetilde{p}_J = \Pr\left[g(x_J) + e_{1J} > -\varepsilon_{2J}\right].$$

Using equation (6) results in

$$\widetilde{p}_J = \begin{cases} \Phi\left[\frac{\alpha + g(x_J) + e_{0J} + e_{1J}}{\sqrt{3}}\right] & \text{if } y_k = 0 \; \forall k < J \\ 1 - \Phi\left[g(x_J) + e_{1J}\right] & \text{if } \exists k < J : y_k = 1 \end{cases}.$$

Define

$$\widetilde{P}_j = \Pr\left[\exists k > j : y_k = 1 \mid y_k = 0 \; \forall k < j\right];$$

i.e., $\widetilde{P}_j$ is the probability that one of the kids younger than j will live near condtional on all of the kids older than j living far away. Then $\widetilde{P}_{J-1} = \widetilde{p}_J$.

Next, child $(J-1)$'s problem is

$$y_{J-1} = 1 \text{ iff } u_1(x_{J-1}) > (1-P_{J-1})\, u_{00J-1} + P_{J-1} u_{01J-1}$$

$$\begin{aligned} & \text{iff } (1-P_{J-1})\,\alpha + g(x_{J-1}) + (1-P_{J-1})\, e_{0J-1} + e_{1J-1} \\ > \; & (1-P_{J-1})\,(\varepsilon_{0J-1} - \varepsilon_{1J-1}) - \varepsilon_{2J-1}. \end{aligned}$$

If $y_k = 0\ \forall k < J-1$, then

$$\begin{aligned} y_{J-1} &= 1 \text{ iff } \left(1-\widetilde{P}_{J-1}\right)\alpha + g(x_{J-1}) + \left(1-\widetilde{P}_{J-1}\right) e_{0J-1} + e_{1J-1} \\ &> \left(1-\widetilde{P}_{J-1}\right)(\varepsilon_{0J-1} - \varepsilon_{1J-1}) - \varepsilon_{2J-1}, \end{aligned}$$

and, if $\exists k < J : y_k = 1$, then $P_J = 1$ which implies that

$$y_{J-1} = 1 \text{ iff } g(x_{J-1}) + e_{1J-1} > -\varepsilon_{2J-1}.$$

Thus, if $y_k = 0\ \forall k < J-1$,

$$\begin{aligned} \widetilde{p}_{J-1} &= \Pr\Big[\left(1-\widetilde{P}_{J-1}\right)\alpha + g(x_{J-1}) + \left(1-\widetilde{P}_{J-1}\right) e_{0J-1} + e_{1J-1} \\ &\quad > \left(1-\widetilde{P}_{J-1}\right)(\varepsilon_{0J-1} - \varepsilon_{1J-1}) - \varepsilon_{2J-1}\Big], \end{aligned}$$

and, if $\exists k < J-1 : y_k = 1$,

$$\widetilde{p}_{J-1} = \Pr\left[g(x_{J-1}) + e_{1J-1} > -\varepsilon_{2J-1}\right].$$

Again, using equation (6) results in

$$\widetilde{p}_{J-1} = \begin{cases} \Phi\left[\frac{\left(1-\widetilde{P}_{J-1}\right)\alpha + g(x_{J-1}) + \left(1-\widetilde{P}_{J-1}\right) e_{0J-1} + e_{1J-1}}{\sqrt{2\left(1-\widetilde{P}_{J-1}\right)^2 + 1}}\right] & \text{if } y_k = 0\ \forall k < J-1 \\ 1 - \Phi\left[g(x_{J-1}) + e_{1J-1}\right] & \text{if } \exists k < J-1 : y_k = 1 \end{cases}$$

and

$$\widetilde{P}_{J-2} = 1 - (1-\widetilde{p}_{J-1})\left(1-\widetilde{P}_{J-1}\right).$$

In general, child j's problem is

$$\begin{aligned} y_j = 1 \text{ iff } & u_1(x_j) > (1-P_j)\, u_{00j} + P_j u_{01j} \\ \text{iff } & (1-P_j)\,\alpha + g(x_j) + (1-P_j)\, e_{0j} + e_{1j} > (1-P_j)(\varepsilon_{0j} - \varepsilon_{1j}) - \varepsilon_{2j}. \end{aligned} \tag{7}$$

If $y_k = 0\ \forall k < j$, then

$$y_j = 1 \text{ iff } \left(1-\widetilde{P}_j\right)\alpha + g(x_j) + e_{1j} > \left(1-\widetilde{P}_j\right)\varepsilon_{0j} - \varepsilon_{2j},$$

and, if $\exists k < J : y_k = 1$, then $P_J = 1$ which implies

$$y_j = 1 \text{ iff } g(x_j) + e_{1j} > -\varepsilon_{2j}.$$

Thus,

$$\widetilde{p}_j = \begin{cases} \Phi\left[\frac{(1-P_j)\alpha+g(x_j)+(1-P_j)e_{0j}+e_{1j}}{\sqrt{2\left(1-\widetilde{P}_j\right)^2+1}}\right] & \text{if } y_k = 0 \ \forall k < j \\ 1-\Phi\left[g\left(x_j\right)+e_{1j}\right] & \text{if } \exists k < j : y_k = 1 \end{cases}, \tag{8}$$

and

$$\widetilde{P}_{j-1} = 1-(1-\widetilde{p}_j)\left(1-\widetilde{P}_j\right). \tag{9}$$

3.2.2 Likelihood Function

Consider data for family i, $\{y_i, x_i\}$ where $y_i = (y_{i1}, y_{i2}, .., y_{iJ_i})$ and $x_i = (x_{i1}, x_{i2}, .., x_{iJ_i})$. Then

$$\begin{aligned} \Pr\left[y_i \mid x_i\right] &= \int \Pr\left[y_i \mid x_i, e\right] f(e)\, de; \\ \Pr\left[y_i \mid x_i, e\right] &= \prod_{j=1}^{J_i} \widetilde{p}_{ij}^{y_{ij}} \left(1-\widetilde{p}_{ij}\right)^{1-y_{ij}}. \end{aligned}$$

This is the likelihood contribution for observation i, and it can be simulated easily as

$$\Pr\left[y_i \mid x_i\right] \approx R^{-1}\sum_{r=1}^{R} \Pr\left[y_i \mid x_i, e^r\right]$$

where e^r is a draw from the distribution implied by equation (6).[9]

4 Results

4.1 Parameter Estimates

Estimation results are reported in Tables 6 and 7. The estimate of α, the preference of children to have someone live nearby, is negative (-0.246). The whole premise of KKLR and RS is that children want someone to live nearby. The negative estimate suggests that each child wants to avoid a situation where he/she is the only child living nearby (which happens with probability $1-P_j$). For example, when all of the older children have made a decision, the youngest child is more likely to live nearby if one of the other, older children also lives nearby.

The model consistent with this interpretation is somewhat different than the one presented in Section 3.2. Let $g(x_j)$ be a deterministic component of living nearby and α be a deterministic component of being the only child living nearby. Replace the utilities in equations (3), (4), and (5) with:

1. No child lives near the parent with utility u_{00j} having the same structure as in equation (3),

$$u_{00j} = \eta_{0j} \tag{10}$$

2. Only child j lives near the parent with utility

$$u_{01j} = \alpha + g(x_j) + \xi_{0j} + \eta_{1j} \tag{11}$$

[9] I use antithetic acceleration to reduce simulation error.

3. Child j lives near the parent along with at least one other child with utility

$$u_{1j} = u_{01j} - \alpha + \xi_{1j} + \eta_{2j}. \tag{12}$$

Then,

$$\begin{aligned} y_j = 1 \text{ iff } & P_j u_{1j} + (1 - P_j) u_{01j} > u_{00j} \\ \text{iff } & P_j \left(g(x_j) + \xi_{0j} + \xi_{1j} + \eta_{1j} + \eta_{2j} \right) + (1 - P_j) \left(\alpha + g(x_j) + \xi_{0j} + \eta_{1j} \right) > \eta_{0j} \\ \text{iff } & (1 - P_j) \alpha + g(x_j) + \xi_{0j} + P_j \xi_{1j} > \eta_{0j} - \eta_{1j} - P_j \eta_{2j}. \end{aligned} \tag{13}$$

If one sets

$$\begin{aligned} \xi_{0j} &= e_{0j} + e_{1j}, \\ \xi_{1j} &= -e_{0j}, \\ \eta_{0j} - \eta_{1j} &= \varepsilon_{0j} - \varepsilon_{1j} - \varepsilon_{2j}, \text{ and} \\ \eta_{2j} &= \varepsilon_{0j} - \varepsilon_{1j}, \end{aligned}$$

then the first order condition for this problem in equation (13) is identical to the first order condition for the model in Section 3.2, equation (7). Thus, the two models are not separately identified other than by the sign of α; if $\alpha > 0$, then the model in Section 3.2 makes sense and the alternative model does not, and if $\alpha < 0$, then the opposite is true.

The deterministic component of the net benefit of living nearby, $g(x_j) = x_j \beta$, evaluated at the mean child, is -0.015. Children of black families prefer to live nearby (0.306), and Protestant (0.348) and Catholic (0.486) children prefer to live nearby. Consistent with KKLR, daughters prefer to live nearby relative to sons (0.121), older cohorts prefer living nearby (0.005) but only by a small amount, and married children prefer living far (-0.142).[10] There is no evidence that education affects child location choice; this was an important variable in the analysis in KKLR and RS in that both models predicted that those with better economic opportunities would be more likely to live far. My estimate is inconsistent with Lin and Rogerson (1995), Kalmijn (2006), and Vidal (2008), all finding that children with more education are more likely to live far from their parents.

For Protestant children who are otherwise average, $g(\cdot) = -0.012$; for otherwise average Catholic children, $g(\cdot) = 0.126$; and for otherwise average black children, $g(\cdot) = 0.265$. Except for black and Catholic children, the deterministic component of the net benefit of living nearby is negative. It is not obvious that this should have occurred. While this includes the cost of providing care to the parent (multiplied by the conditional probability of providing care), it also includes the net benefit of having grandparents nearby for any potential children.

[10] Løken, Lommerud, and Lundberg (2011) and Huang (2012) suggest that marital effects may be influenced by tradeoffs in distance between parents and parents-in-law and provides some Norwegian and Taiwanese evidence respectively in support of the idea. Both papers also suggests that marital status should be interacted with gender.

Table 6: Family and Child Parameter Estimates

Family Characteristics			Child Characteristics		
Variable	Estimate	Std Err	Variable	Estimate	Std Err
Alpha	-0.351 **	0.124	Female	0.121 **	0.054
Constant	-0.393 *	0.247	Age	0.005 **	0.003
Black	0.306 **	0.128	Education	-0.015	0.012
Hispanic	-0.170	0.174	Married	-0.142 **	0.063
Protestant	0.348 **	0.105			
Catholic	0.486 **	0.116			

Note: Double-starred items are statistically significant at the 5% level, and single-starred items are significant at the 10% level.

Table 7: Choleski Decomposition Terms

Term	Estimate	Std Err
Family (1,1)	1.495 **	0.260
Family (2,1)	-0.081	0.127
Family (2,2)	-0.636 **	0.079
Child (1,1)	-0.803 **	0.367
Child (2,1)	0.292 **	0.143
Child (2,2)	-0.211 *	0.131

Note: Double-starred items are statistically significant at the 5% level, and single-starred items are significant at the 10% level.

Hank (2007) has many more statistically significant parameter estimates associated with parent and child variables. Differences between mine and Hank (2007) are the source of data, the difference between cross-section and panel data, and the lack of emphasis in Hank (2007) on family size and birth order.

Table 7 provides the estimates of the family-specific and person-specific Choleski decomposition terms. They result in a covariance matrix of 2×2 blocks with each block corresponding to covariances between the unobserved heterogeneity e errors from equation (6) of two children. The diagonal blocks include both family- and person-specific effects, and the off-diagonal blocks include only family-specific effects. Given the estimates in Table 7, the covariance matrix of the unobserved heterogeneity for a 3-child family is displayed in Table 8. These should be compared to the covariance matrix of the true randomness I. These are of the same order of magnitude implying that it is important both to allow for unobserved heterogeneity and for some uncertainty on the part of family members when making location decisions.

The net utilities of the different choices also include unobserved heterogeneity terms who distribution is described in Table 8. Figure 4 displays the distribution of $\Pr(u_{01} > u_{00})$, $\Pr(u_1 > u_{00})$, and $\Pr(u_1 > u_{01})$ with respect to the unobserved heterogeneity. The median probabilities are all very close to 0.43. However, extreme values of $\Pr(u_{01} > u_{00})$ are less likely than low values of either $\Pr(u_1 > u_{00})$ and $\Pr(u_1 > u_{01})$, while moderate values of $\Pr(u_{01} > u_{00})$ are more likely than low values of either $\Pr(u_1 > u_{00})$ and $\Pr(u_1 > u_{01})$. In other words, $\Pr(u_1 > u_{00})$ and $\Pr(u_1 > u_{01})$ second-order dominate $\Pr(u_{01} > u_{00})$. Also, note that $\Pr(u_1 > u_{00})$ and $\Pr(u_1 > u_{01})$ have almost

Table 8: 3-Child Family Covariance Matrix

		Child1		Child 2		Child 3	
		e(0)	e(1)	e(0)	e(1)	e(0)	e(1)
Child 1	e(0)	2.879					
	e(1)	-0.355	0.541				
Child 2	e(0)	2.235	-0.121	2.879			
	e(1)	-0.121	0.411	-0.355	0.541		
Child 3	e(0)	2.235	-0.121	2.235	-0.121	2.879	
	e(1)	-0.121	0.411	-0.121	0.411	-0.355	0.541

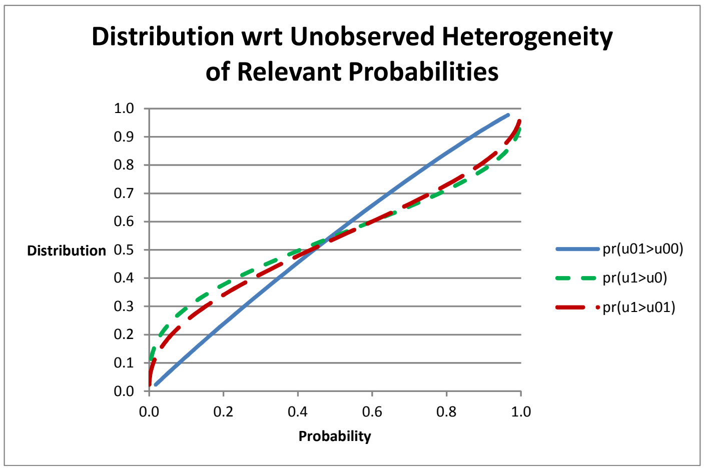

Figure 4: Distribution wrt Unobserved Heterogeneity of Relevant Probabilities

the same distribution. Therefore, the remainder of this analysis, only $\Pr(u_{01} > u_{00})$ and $\Pr(u_1 > u_{00})$ are compared.

Figures 5 and 6 show how race and religion affect the distribution of probabilities. Figure 5 shows that the distributions for blacks first-order stochastically dominate the distribution for whites. This occurs because, as seen in Table 6, blacks get more net benefit from living nearby. Figure 6 shows that the distributions for Catholics first-order stochastically dominate the distributions for Protestants, and the distributions for Protestants first-order stochastically dominate the distributions for non-Christians.

4.2 Specification Tests

As discussed in Section 2.1, there are 201 observations with children not in age-descending order. This is a critical issue in that the model requires that the children make decisions sequentially by age.[11] I estimate the model under three different assumptions concerning such cases (see Section 2.1) and, with one exception,[12] find no significant differences in parameter estimates across the three methods.

JM allows for externalities to depend on the number of other siblings living nearby, while my specification considers only the existence of a nearby child (α in equations (4) and (11)). Given results in Checkovich and Stern

[11] Chan and Ermisch (2013) provide evidence that, in England, the order of children leaving the childhood home depends on more variables than age, for example gender. I chose not to pursue a more sophisticated ordering of children in this paper.

[12] The constant in the model with excluded observations is statistically significant at the 10% level than the constant in the model with reordered observations.

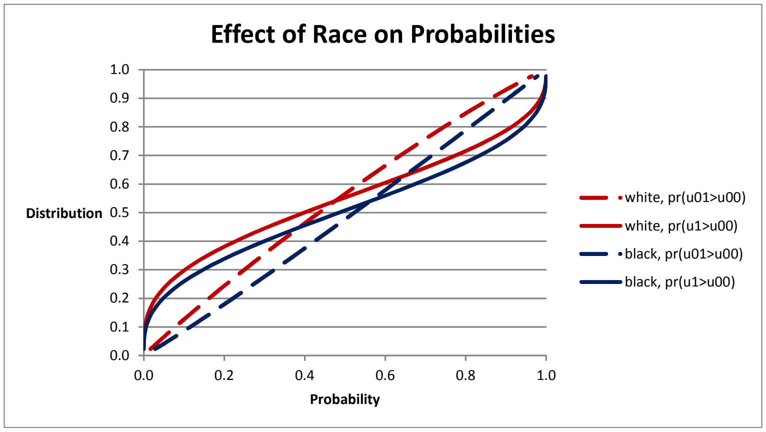

Figure 5: Effect of Race on Probabilities

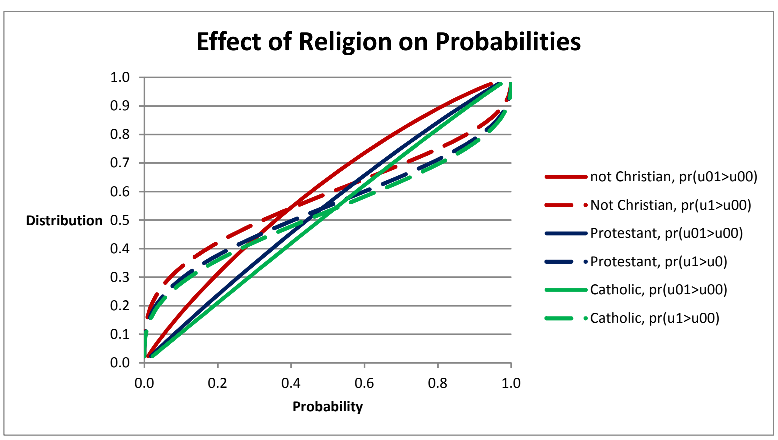

Figure 6: Effect of Religion on Probabilities

Table 9: Residual Characteristics Disaggregated by Family Size and Birth Order

	Birth Order				
# Children	1	2	3	4	Overall
1	0.041 **				0.041 **
	(0.331)				(0.331)
2	-0.016	0.004			-0.006
	(0.373)	(0.340)			(0.357)
3	-0.016	-0.013	-0.035 **		-0.021 **
	(0.373)	(0.368)	(0.349)		(0.363)
4	-0.055 **	-0.041 *	-0.024	0.002	-0.030 **
	(0.369)	(0.362)	(0.376)	(0.368)	(0.369)
Overall	-0.005	-0.009	-0.031 **	0.002	-0.010 *
	(0.363)	(0.353)	(0.358)	(0.368)	(0.359)

Notes:
1) Numbers in parentheses are standard deviations.
2) Double-starred items are statistically significant at the 5% level, and single-starred items are statistically signigicant at the 10% level.

(2002) detailing the lack of frequency of many children within a family providing significant care, one might conclude that only the existence of a nearby child is important. However, not all nearby children provide care, even when no other nearby child exists, so the number of nearby children might matter. I could allow for such effects in my model but at significant computation cost because it would not be enough to keep track of just the probability of any other younger child living near; the probability terms analogous to those in equations (8) and (9) would be much more complicated. If, as in JM, it is important to allow the number of children nearby to affect utility, then, in my results, I should find different behavior among residuals across families of different sizes. Table 9 presents the mean residual and standard deviation of residuals disaggregated by family size and birth order. Most of the mean residuals are statistically insignificant. If one were to use a Bonferroni correction to account for the large number of test statistics, one would not reject the null hypothesis that there are no family size or birth-order effects in the residuals. However, aggregating over birth order within family size, one finds statistically significant effects for family size. This suggests that it might be fruitful to allow for a more general specification including a linear effect in number of near children (as in JM) along with a discrete step at one near child (as in this model).

4.3 Variation in Siblings' Private Information

There are two types of errors in the model, common knowledge errors ξ and private information errors η. It may be that the information that siblings have about each other's private information errors depends upon characteristics of the siblings.[13] For example, consider decomposing each private information error η_{kj} into two parts,

$$\eta_{kj} = \lambda_{jm}\eta^o_{kj} + (1 - \lambda_{jm})\,\eta^u_{kj}$$

where η^u_{kj} is unobserved by sibling $m < j$, η^o_{kj} is observed by sibling m,

$$\begin{aligned} \eta^o &\sim iidN\left(0, \Omega_\eta\right), \\ \eta^u &\sim iidN\left(0, \Omega_\eta\right), \end{aligned}$$

[13] This idea was suggested to me by Alicia Baek.

$$\lambda_{jm} = \lambda e^{-(a_m - a_j)},$$

(a_m, a_j) are the ages of m and j with $a_m - a_j \geq 0$, and λ is a parameter determining how much difference in age affects information siblings know about each other. Then m's belief about the probability that j will live nearby is

$$\widetilde{p}_{jm} = \begin{cases} \Phi\left[\frac{(1-P_{jm})\alpha + g(x_j) + \xi_{0j} + P_{jm}\xi_{1j} - \lambda_{jm}\left(\eta^o_{0j} - \eta^o_{1j} - P_{jm}\eta^o_{2j}\right)}{(1-\lambda_{jk})\sqrt{\left(2\left(1-\widetilde{P}_{jm}\right)^2+1\right)}}\right] & \text{if } y_k = 0 \ \forall k < j \\ 1 - \Phi\left[g(x_j) + \xi_{0j} + \xi_{1j}\right] & \text{if } \exists k < j : y_k = 1 \end{cases},$$

and m's prior that $j-1$ thinks that one of the younger siblings will live near is

$$\widetilde{P}_{j-1,m} = 1 - (1 - \widetilde{p}_{jm})\left(1 - \widetilde{P}_{jm}\right).$$

This implies that the probability that j lives nearby is

$$\widetilde{p}_j = \begin{cases} \Phi\left[\frac{(1-P_{jj})\alpha + g(x_j) + \xi_{0j} + P_{jj}\xi_{1j}}{\sqrt{\left(2\left(1-\widetilde{P}_{jm}\right)^2+1\right)}}\right] & \text{if } y_k = 0 \ \forall k < j \\ 1 - \Phi\left[g(x_j) + \xi_{0j} + \xi_{1j}\right] & \text{if } \exists k < j : y_k = 1 \end{cases}.$$

This can be plugged into the log likelihood function and then used to perform a Lagrange Multiplier test for $H_0 : \lambda = 0$ vs $H_A : \lambda > 0$.[14] The t-test associated with this test has a value of 0.001. Thus, it appears that age differences in siblings do not help predict the amount of information siblings have about each other, at least in terms of where each chooses to live. The lack of significance suggests that it may be inappropriate to assume that children within a family know everything about each other. For example, this may be a problem for the perfect information assumption in JM.

4.4 Welfare Analysis

There are two sources of welfare loss in the model. The first is that older children do not know how younger children will behave and so can only maximize expected utility over the joint distribution of younger children behavior. If older children knew with certainty how younger children were going to behave, then they could do no worse and, at least in some cases, do better. The second is that, even in a world with certainty about behavior by one's siblings, the lack of availability of sidepayments preclude some preferable equilibria. For example, in the KKLR model, in some circumstances, it would be beneficial for the older child to commit to living near in return for a sidepayment from the younger child.

To gain some understanding of the magnitude of the welfare losses in this model, I simulate draws of errors multiple times for each multiple-child family in the data[15] and compare the difference between utility in the model described above and the utility that would be achieved by an omniscient social planner maximizing the sum of child utilities. The results are reported in Table 10. By definition, the social planner never does worse than the equilibrium, and losses can be large. The mean loss varies by family size and increases almost proportionately with family size. Relative to the variation in utility due to randomness, welfare losses are in the range of 50%.

[14] To construct the Lagrange Multiplier statistic, one needs to use the derivatives included in the appendix.

[15] Note that there can be no welfare loss for families with one child.

Table 10: Aggregate Level Moments of Welfare Losses Disaggregated by Family Size

Family Size	# Obs	Level Mean	Relative Mean	Level Std Dev	Level Minimum	Level Maximum
2	1026	-0.151	-0.446	0.082	-0.542	0.000
3	665	-0.229	-0.506	0.112	-0.640	-0.002
4	318	-0.261	-0.510	0.106	-0.560	-0.052

Note: "Relative Mean" is relative to the standard deviation of losses

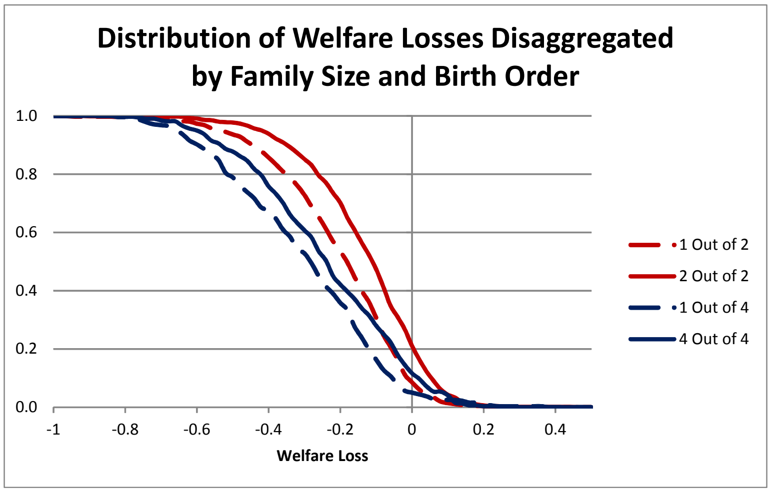

Figure 7: Distribution of Welfare Losses Disaggregated by Family Size and Birth Order

Figure 7 shows the distribution of welfare loss disaggregated by family size and by birth order. The horizontal component of a point in the curve is a particular utility loss between equilibrium utility and utility from the social planner, and the vertical component is the proportion of children with a particular family size and birth order with losses at most equal to the horizontal component. For example, using the curve for 2nd children in 2-child families (2 Out of 2), 70% have losses of -0.2 or less. Note that some children do better in equilibrium than they would with the social planner. For example, 20% of 2nd children in 2-child families do better. Also note that younger children do better than older children.[16] Younger children can condition their behavior on the behavior of their older siblings which is more important than their choices being limiting by the decisions of their older siblings; resolution of randomness is more important than the value of first mover. Finally, note that welfare loss is greater in larger families than in smaller families.[17]

Instead of disaggregating by birth order, one could disaggregate by the $g(x_j)$ terms in equations (5) and (11). This results in curves stochastically dominated by decreases in $g(x_j)$. The children with large values of $g(x_j)$ are more likely to want to live nearby independent of the choices of other children, so they are not as adversely affected by the randomness in the behavior of the other children. Overall, the results strongly suggest that uncertainty with respect to the behavior of siblings is a more important feature of the environment than the effect of birth order on

[16] The same pattern exists if curves for 3-child families and the curves for middle children in 4-child families are included.

[17] The large-family curves stochastically dominate the small-family curves.

decision-making.

5 Conclusion

My results suggest that the model in Konrad et al. (2002) may not fit the data well. In particular, they suggest that uncertainty about the strategy of other siblings plays a large part in the behavior of each sibling and that the existence of such uncertainty has significant welfare effects associated with the geographic distribution of siblings.

Our results suggest that the most important goal of each child is to avoid being the only caregiver available for a needy parent. This is interesting, especially in light of the empirical result that, conditional on provision of some informal care to the parent, it is usually just one child who provides informal care for the parent (Checkovich and Stern, 2002). In other words, sharing care responsibilities among children is somewhat rare.

There is also some desire among the siblings to avoid leaving the parent with no care, and much of the loss in utility associated with living nearby is mitigated by having another sibling nearby. Thus, it appears that there is much room for coordination of actions among siblings. An important open question is under what conditions the siblings have incentive to coordinate or to reveal personal characteristics unobserved by other family members. This is left for future work.

References

[1] Antman, Francisca (2012). "Elderly Care and Intrafamily Resource Allocation when Children Migrate." *Journal of Human Resources.* 47(2): 331-363.

[2] Arnold, F., M. Choe, and T. Roy (1998). "Son Preference, the Family-Building Process and Child Mortality in India." *Population Studies.* 52: 301–315.

[3] Bian, F., J. Logan, and Y. Bian (1998). "Intergenerational Relations in Urban China: Proximity, Contact and Help to Parents." *Demography.* 35(1): 115–124.

[4] Bianchi, Suzanne, Kathleen McGarry, and Judith Seltzer (2010). "Geographic Dispersion and the Well-Being of the Elderly." Unpublished manuscript.

[5] Bonsang, E. (2009). "Does Informal Care from Children to Their Elderly Substitutes from Formal Care in Europe?" *Journal of Health Economics.* 28: 143–154.

[6] Brody, Elaine (2004). *Women in the Middle: Their Parent-Care Years.* New York City: Springer Publishing Co.

[7] Byrne, David, Michelle Goeree, Bridget Hiedemann, and Steven Stern (2009). "Formal Home Health Care, Informal Care, and Family Decision-Making." *International Economic Review.* 50(4): 1205-1242.

[8] Chan, Tak Wing and John Ermisch (2013). "Family Geography and Family Demography in the UK: Cross-Sectional Perspective." Unpublished manuscript.

[9] Checkovich, Tennille and Steven Stern (2002). "Shared Caregiving Responsibilities of Adult Siblings with Elderly Parents." *Journal of Human Resources*. 37(3): 441-478.

[10] Compton, Janice and Robert Pollak (2009). "Proximity and Co-residence of Adult Children: Description and Correlates." Unpublished manuscript.

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

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

[13] Hank, K. (2007). "Proximity and Contacts between Older Parents and Their Children: A European Comparison." *Journal of Marriage and Family*. 69(1): 157–173.

[14] Hiedemann, Bridget, Michelle Sovinsky, and Steven Stern (2014). "Will You Still Want Me Tomorrow? The Dynamics of Families' Long-Term Care Arrangements." Unpublished manuscript.

[15] Heidemann, Bridget and Steven Stern (1999). "Strategic Play Among Family Members when Making Long-Term Care Decisions." *Journal of Economic Behavior and Organization*. 40(1): 29-57.

[16] Holmlund, Helena, Helmut Rainer, and Thomas Siedler (2009a). "Meet the Parents? The Causal Effect of Family Size on the Geographic Distance between Adult Children and Older Parents." *IZA DP No. 4398*.

[17] Holmlund, Helena, Helmut Rainer, and Thomas Siedler (2009b). "Geography of the Family: Comment." Unpublished manuscript.

[18] Horowitz, Amy (1985). "Family Caregiving to the Frail Elderly." *Annual Review of Gerontology and Geriatrics*. 5: 195-246.

[19] Huang, Ziyan (2012). "The Residential Choice of Adult Children: Competition between Parents and Parents-in-Law." Unpublished manuscript.

[20] Ikels, Charlotte (1983). "The Process of Caregiver Selection." *Research on Aging*. 5(4): 491-509.

[21] Johar, Meliyanni and Shiko Maruyama (2012). "Externality and Strategic Interaction in the Location Choice of Siblings under Altruism toward Parents." Unpublished manuscript.

[22] Kalmijn, Matthijs (2006). "Educational Inequality and Family Relationships: Influences on Contact and Proximity." *European Sociological Review*. 22(1): 1-16.

[23] Kennan, John and James Walker (2011). "The Effect of Expected Income on Individual Migration Decisions." *Econometrica*. 79(1): 211-251.

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

[25] Kureishia, Wataru and Midori Wakabayashi (2010). "Why do First-Born Children Live Together with Parents?" *Japan and the World Economy*. 22(3): 159–172.

[26] Lin, Ge and Peter Rogerson (1995). "Elderly Parents and the Geographic Availability of their Adult Children." Research on Aging. 17(3): 303-331.

[27] Litwak, E. and S. Kulis (1987). "Technology, Proximity, and Measures of Kin Support." *Journal of Marriage and the Family*. 49(3): 649-661.

[28] Litwak, E. and C. Longino (1987). "Migration Patterns among the Elderly: A Developmental Perspective." *Gerontologist*. 27(3): 266-272.

[29] Løken, Katrine, Kjell Lommerud, and Shelly Lundberg (2011). "Your Place or Mine? On the Residence Choice of Young Couples in Norway." Unpublished manuscript.

[30] Martin, L. (1990). "Changing Intergenerational Family Relations in East Asia." *Annals of the American Association of Political and Social Science*. 510: 102-114.

[31] McGarry, Kathleen (1998). "Caring for the Elderly: The Role of Adult Children." *Inquiries in the Economics of Aging*. David Wise (ed.). Chicago: University of Chicago Press.

[32] Rainer, Helmut and Thomas Siedler (2009). "O Brother, Where Art Thou? The Effects of Having a Sibling on Geographic Mobility and Labour Market Outcomes." *Economica*. 76(303): 528-556.

[33] Rogerson, Peter, Jeffrey Burr, and Ge Lin (1997). "Changes in Geographic Proximity between Parents and their Adult Children." *International Journal of Population Geography*. 3: 121-136.

[34] Rogerson, P., R. Weng, and G. Lin (1993). "The Spatial Separation of Parents and their Adult Children." *Annals of the Association of American Geographers*. 83(4): 656-671.

[35] Silverstein, Merril (1995). "Stability and Change in Temporal Distance between the Elderly and Their Children." *Demography*. 32(1): 29-45.

[36] Stern, Steven (1995). "Estimating Family Long-Term Care Decisions in the Presence of Endogenous Child Characteristics." *Journal of Human Resources*. 30(3): 551-580.

[37] Tamer, E. (2003). "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria." *Review of Economic Studies*. 70(1): 147-167.

[38] Treas, Judith, R. Gronvold, and V. Bergtson (1980). "Filial Destiny, the Effect of Birth Order on Relations with Aging Parents." Paper presented at the Annual Scientific Meeting of the Gerontological Society, San Diego.

[39] United States, Bureau of Census (2012). *Geographical Mobility: 2010-2011*. https://www.census.gov/hhes/migration/data/cps/cps2011.html.

[40] Vidal, Sergi (2008). "Internal Migration, the Family Context and Transitions to Adulthood in West-Germany." Unpublished manuscript.