COMBINATORIAL BOOTSTRAP INFERENCE IN PARTIALLY IDENTIFIED INCOMPLETE STRUCTURAL MODELS: SUPPLEMENTARY MATERIAL

MARC HENRY, ROMUALD MEANGO AND MAURICE QUEYRANNE

1. Min-Cut Max-Flow Theorem

The network flow problem refers to the optimal way to route quantities through a given exogenous network to meet supply, demand and capacity constraints.

**Definition 1** (Directed graph). A directed graph is a pair $G = (V, E)$, where $V = \{1, \ldots, n\}$ is a finite set of points, called nodes, and $E$ is a set of ordered pairs $e = (v^1, v^2)$ of elements of $V$, called arcs.

**Definition 2** (Path). Fix a graph $G = (V, E)$ and a sequence $v^1, \ldots, v^r$ of nodes in $V$. A path is a sequence of arcs $e^1, \ldots, e^{r-1}$, such that $e^i = (v^i, v^{i+1})$, for each $i = 1, \ldots, r - 1$.

**Definition 3** (Network). A network is a directed graph $G = (V, E)$ with two distinct nodes $s$ and $t$ designated as source and sink, respectively, such that there is at least one path between $s$ and $t$ and such that each arc $e \in E$ is endowed with a positive number $c(e)$ called its capacity. A network will be denoted $N = (V, E, c)$.

A flow problem is defined on a network.

**Definition 4** (Flow). A flow is a assignment of weights $f(e)$ to each arc $e \in E$ of a network $N = (V, E, c)$, which satisfies the following constraints:

1. **Capacity constraint:** $0 \leq f(e) \leq c(e)$ for any $e \in E$;
2. **Flow conservation:** for any $v \in V \setminus \{s, t\}$, the sum of flows $f(e)$ on arcs $e = (v, \cdot)$ starting at $v$ is equal to the sum of flows $f(e')$ on arcs $e' = (\cdot, v)$ ending at $v$.

Date: June 3, 2014. The authors thank David Straley for providing the data and Daniel Stubbs for computing assistance. Correspondence address: Economics Department, Penn State, University Park, PA 16802, USA. marc.henry@psu.edu.
The sum of \( f(\cdot, t) \) flowing into the sink \( t \) will be denoted \( F \) and called value of the flow.

The maximum flow problem is the problem of finding a flow that maximizes the sum of \( f(\cdot, t) \) flowing into the sink \( t \). It is dual of a problem known as minimum cut.

**Definition 5** (Cut). A cut in a network is a node partition \((S, T)\), i.e., \( S \cup T = V \) and \( S \cap T = \emptyset \), such that \( s \in S \) and \( t \in T \). The capacity \( c(S, T) \) of a cut \((S, T)\) is the sum of capacities \( c(v, w) \) for all arcs \((v, w)\) such that \( v \in S \) and \( w \in T \) (i.e., arcs leaving \( S \)).

The minimum cut problem is the problem of finding a cut in a network with minimum capacity.

Observe the following facts:

1. By definition of a cut \((S, T)\), the net flow across a cut, i.e., the sum of \( f(v, w) \) for all arcs \((v, w)\) such that \( v \in S \) and \( w \in T \) (i.e., arcs leaving \( S \)), is equal to the value of the flow, i.e., the amount \( F \) reaching the sink \( t \).
2. By the capacity constraints and the previous observation, the value \( F \) of the flow is smaller than the capacity of any cut.
3. From the previous observation, it immediately follows that if there is a flow and a cut such that value \( F \) is equal to the capacity of the cut, then the flow is maximum and the cut is minimum.

We can now state the main theorem, from Ford and Fulkerson (1957).

**Theorem 1** (Min-Cut Max-Flow). In any network, the value of the maximum flow equals the capacity of the minimum cut.

Before we prove the theorem and present the Ford-Fulkerson algorithm to find a maximum flow, two more definitions are needed.

**Definition 6** (Residual network). A residual network is obtained from a network endowed with a flow by removing all arcs \((v, w)\) such that \( f(v, w) > 0 \) and replacing them by and arc \((w, v)\) with capacity equal to \( f(v, w) \) and, when \( f(v, w) < c(v, w) \), arc \((v, w)\) with capacity equal to \( c(v, w) - f(v, w) \).
Definition 7 (Augmenting path). An augmenting path in a network endowed with a flow is a path in the residual network.

The Ford-Fulkerson algorithm to find a maximum flow in a network consists in finding augmenting paths and increasing flow along the latter. Its convergence is based on the proof below.

Proof of Theorem 1. The method of proof is to show the equivalence of the following three statements:

1. Flow $F$ is maximum.
2. There is no augmenting path,
3. There exists a cut with capacity $F$.

By definition, if the flow is maximum, there is no augmenting path. By previous observation, if there exists a cut with capacity $F$, then the cut is minial and the flow is maximum. There remains to show that if there is no augmenting path, then there is a cut with capacity $F$. Let $S$ be the set of vertices $v$ such that there is a path from $s$ to $v$ in the residual network. Then we have the following:

- $S$ contains $s$.
- Since there is no augmenting path, $t \notin S$.
- All arcs $e$ leaving $S$ in the original network have $f(e) = c(e)$.

Hence, $(S, V\setminus S)$ is a cut with capacity $F$ as required. \qed

2. Elderly Care Provision

We estimate the determinants of long term care option choices for elderly parents in American families. The model we use closely follows the one proposed by Engers and Stern (2002) who present these choices as the result of a participation game. The family members decide simultaneously whether to participate in a family reunion where the care option maximizing the participants’ utility is chosen. Profits are then split among these participants according to some benefit-sharing rule.
The data consists of a sample of 1,212 elderly Americans with two children drawn from the National Long Term Care Survey, sponsored by the National Institute of Aging and conducted by the Duke University Center for Demographic Studies under Grant number U01-AG007198, Duke (1999). Elderly people were interviewed in 1984 about their living and care arrangements. The survey questions include gender and age of the children, the distance between homes of the elderly parent and each of the children, the disability status of the elderly parent (where disability is referred to as problems with “Activities of Daily Living or Instrumental Activities of Daily Living (ADL)” and the number of days per week each of the children devotes to the care of the elderly parent. The dependent variable is the care provision for the parent. The parent is asked to list children (either at home or away from home) and how much each provides help. If only one child is listed as providing significant help, that child is designated the primary care giver. If more than one child is listed, the one providing the most time is designated the primary care giver. If the elderly parent lives in a nursing home, then the nursing home is the primary care giver. If no child is listed and the parent does not live in a nursing home, then the parent is designated as “living alone”. Table 1 presents the list of variables used in the analysis. They include parent characteristics, characteristics of the children and the care option chosen.

3. Data summary statistics

The dependent variable is the care option chosen by the family. We classify it in 4 categories: “Child 1” (resp. “Child 2”) means that the firstborn child (resp. the second born child) provides the most care in terms of days spent helping the parents. Our data exhibits a strong dominance of the choice of child 1 over the choice child 2. Child 1 is listed as primary care giver in 26.81% of the families, while child 2 is only listed in 6.75 % of our observations. On average, the older child provides 3.50 (3.12) days of care while the second child provides 0.48 (1.34) days. When Child 1 is the chosen option, the average numbers of days spent by the primary care giver climbs to 4.32 (2.92), while the other child provides on average only 0.17 (0.77) days of care. When Child 2 is the preferred option, the firstborn will spend on average only 0.29 (1.29) days of care while the second child provides 1.74 (2.15) days of care. The third care option, for the parent to enter a
nursing home, represents 19.92% of the families in our sample. The remaining option, for parents to live alone, includes all the cases where no child is listed and the parent does not live in a nursing home (it could well be the case that another individual, other than the children provides care for the parent). Parent is living alone in the remaining 46.74% of the sample, which makes it the most frequent option in our sample. We use two types of explanatory variables, those relative to the elderly parent, and those relative to children. The first variable relative to the elderly parent is their disability status based on six types of daily living activities: bed, bathing, dressing, eating, toilet, walking inside. As much as 65% of the parents in the sample population suffer at least one disability, of which 51.5% have 4 or more. In the following, we introduce a categorical variable DA with value 1 when the parent shows at least 4 disabilities, 0 otherwise. The second variable relative to the parent characteristics is a categorical variable for the presence of the parent’s spouse in the household. We denote it DM. Previous studies show a positive effect on the incentive to remain home when the parent lives with his or her spouse. In our sample, 40.36% live with a spouse. We consider three characteristics of each child: their distance to the parent, their birth order and their gender. The distance can be viewed as a cost for a child to provide days of care to the parent. The survey measures the time required to travel from each child’s residence to the parent’s residence. About one half of the children in the sample live at a travel distance of 30 minutes or more, and 12% of the children live in the same household as their parent. Among care givers, the distribution of distance is quite different. Among the firstborn children in charge of the parent, 25% live in the same household, and the percentage of those living more than 30 minutes away drops to 31%. Similarly, among primary care giver who are second born children, the percentage of those living more than 30 minutes away is only 25%. However, the number of those living in the same household as the parents is 7%. As noted by Engers and Stern (2002), the observations where the parent lives in the same household as the child induces some statistical problems. It is unclear if the child is the actual care giver for the parent or if the parent is actually the care giver for the child. To capture the effect of distance on children’s incentives, we will therefore conduct estimation conditional on the children living in a different household than the parent’s household. We include in our analysis, a dummy variable for each one of the
Variables Equal to 1 if: Percentage of sample

<table>
<thead>
<tr>
<th>Care Option</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Living with child 1</td>
<td>26.81</td>
<td></td>
</tr>
<tr>
<td>Living with child 2</td>
<td>6.75</td>
<td></td>
</tr>
<tr>
<td>Living in nursing home</td>
<td>19.92</td>
<td></td>
</tr>
<tr>
<td>Living home alone</td>
<td>46.54</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parent Variables</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$DA$</td>
<td>Highly disabled</td>
<td>33.81</td>
</tr>
<tr>
<td>$DM$</td>
<td>Living with the spouse</td>
<td>40.36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Children Variables</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$DD0$</td>
<td>Living with parent</td>
<td>11.55</td>
</tr>
<tr>
<td>$DD1$</td>
<td>Distance from parent: 31 min and more</td>
<td>49.45</td>
</tr>
<tr>
<td>$DS$</td>
<td>Female</td>
<td>49.26</td>
</tr>
</tbody>
</table>

Table 1

children: $DD1_i = 1$ if child $i$ lives 30 minutes away or more, 0 otherwise. The child’s gender is an important variable in most of the analysis. Previous work (see, for example Horowitz (1982), Treas, Gronvold, and Bergston (1980)) suggests that being female makes a child most likely to provide care. However, Stern (1995) and Engers and Stern (2002) indicate that the data set suffers from significant misclassification of gender (see Stern (1995) for a detailed discussion of the problems with the data set). We introduce in the specification a random error term to account for misclassification of gender.

4. The game

The observable choice of care option is modeled as in Engers and Stern (2002) as the outcome of a participation game. Before going further, we introduce some notation. We
will index family members as follows. Parent: 0, Firstborn child: 1 and Second born child: 2.

Define $V_{0j}$ as the value to parent 0 and $V_{ij}$ as the value to child $i$ of care option $j$, where $j \in \{1, 2\}$ means child $j$ becomes the primary care giver, $j = 0$ means the parent remains at home with no significant help from either child and $j = 3$ means the parent is moved to a nursing home. The matrix $V = (V_{ij})_{ij}$ is known to both children and the parent. We suppose it takes the form

$$V_{ij} = \gamma_{ij} + W\beta_{ij} + Z_j\psi_{ij}$$

where $W$ indicates the characteristics of the parents (DA and DM), and $Z_j$ indicates the characteristics of care option $j$ (DS, DD1 and DD2) and $X = (W, Z)$. $\theta = (\gamma_{ij}, \beta_{ij}, \psi_{ij})'$ is unknown to the analyst and the object of inference.

**Example 1.** Consider the following family, in which the matrix where given value of $X$ and $\theta$ result in $V$ that takes the form:

$$V = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 4 & -1 & 1 \\ 0 & -1 & 4 & 1 \end{bmatrix}$$

Rows indicate family member $i = 0, 1, 2$ and columns represent care giving options $j = 0, 1, 2, 3$ in that order. In this example, the parent is indifferent between all the care options, except the one where she has to move to the Nursing home. Each child prefers to be the primary care giver to any other care option, followed by living in a nursing home, the parent living at home, and being taken care of by the other child, in that order.

We assume that it is always in the interest of the parent to attend the family reunion. However, child $i$ ($i = 1, 2$) can refrain from participating in the meeting. By choosing not to participate, a member of the family agrees on whatever is decided but can neither ensure the role of primary care giver, nor can he be involved in any side payment. Both children simultaneously decide whether or not to participate in the long term care decision. Suppose $M$ is the set of children who participate. The option chosen is option $j \in M \cup \{0, 3\}$ which maximizes the participants’s total utility $\sum_{i \in M} V_{ij}$. It is assumed that participants
abide by the decision and that benefits are then shared equally\(^1\) among parent and children participating in the decision through a monetary transfer \(s_i\), which is the second term in the children’s payoff.

Finally, \(\epsilon_i\) is a random cost or benefit from participation, which is 0 for children who decide not to participate and distributed according to absolutely continuous distribution \(\nu(\cdot|\theta)\) for each child who participates. All children observe the realizations of \(\epsilon\), whereas the analyst only knows its distribution.

Therefore, the payoff matrix of the participation game can be determined in the following way.

- If both children decide to participate, denoted \(PP\), child \(i\)'s payoff (for \(i = 1, 2\)) is

  \[
  \Pi_i = \epsilon_i + w^{PP}
  \]

  \[
  \Pi_i = \epsilon_i + \frac{1}{3} \times \left( \max_{j \in \{0,1,2,3\}} \sum_{k \in \{0,1,2\}} V_{kj} \right)
  \]

  where \(w^{PP}\) is the share of the overall benefit that each child gets when both children participate. Note that we have 3 participants and an equal sharing rule, which explains the term \(\frac{1}{3}\).

- If neither child participates, denoted \(NN\), child \(i\) payoff is

  \[
  \Pi_i = V_{ij} \text{ with } j \text{ such that } j = \arg \max \{V_{00}; V_{03}\}
  \]

  Indeed, since none of the children participate, the parent (0) picks the best option among the two available (0 and 3). We will note this quantity \(w^{NN}_i\).

- If only child 1 participates, (option denoted \(PN\), child 1’s payoff is

  \[
  \Pi_1 = \epsilon_1 + w^{PN}
  \]

  \[
  \Pi_1 = \epsilon_1 + \frac{1}{2} \times \left( \max_{j \in \{0,1,2,3\}} \sum_{i \in \{0,1\}} V_{ij} \right)
  \]

\(^1\)Other benefit-sharing rule can be explored. Engers and Stern (2002) study different possible rules: two Pareto optimal rules and one based on the Shapley value
where $w_1^{PN}$ is the share of the overall benefit that player 1 gets when she the only one to participate. And child 2’s payoff is

$$\Pi_2 = V_{2j} : j = \arg\max_{j \in \{0,1,3\}} \sum_{i \in \{0,1\}} V_{ij}$$

Following previous notations, we will call this quantity $w_2^{PN}$.

- For option $NP$, $w_i^{NP}$, $i = 1,2$ are defined similarly.

The Payoff matrix can then be written as in 4.

<table>
<thead>
<tr>
<th>Child 1</th>
<th>N</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>$w_1^{NN}, w_2^{NN}$</td>
<td>$w_1^{NP}, \varepsilon_2 + w_2^{NP}$</td>
</tr>
<tr>
<td>P</td>
<td>$\varepsilon_1 + w_1^{PN}, w_2^{PN}$</td>
<td>$\varepsilon_1 + w^{PP}, \varepsilon_2 + w^{PP}$</td>
</tr>
</tbody>
</table>

We derive best responses (in pure strategies) $br_i(s_{3-i})$ for child $i$ to a strategy $s_{3-i}$ played by her sibling. Then we have:

$$br_1(P) = \begin{cases} 
P & \text{if } \varepsilon_1 \geq w_1^{NP} - w^{PP} \\
N & \text{if } \varepsilon_1 \leq w_1^{NP} - w^{PP}
\end{cases}$$

$$br_1(N) = \begin{cases} 
P & \text{if } \varepsilon_1 \geq w_1^{NN} - w_1^{PN} \\
N & \text{if } \varepsilon_1 \leq w_1^{NN} - w_1^{PN}
\end{cases}$$

$$br_2(P) = \begin{cases} 
P & \text{if } \varepsilon_2 \geq w_2^{PN} - w^{PP} \\
N & \text{if } \varepsilon_2 \leq w_2^{PN} - w^{PP}
\end{cases}$$

$$br_2(N) = \begin{cases} 
P & \text{if } \varepsilon_2 \geq w_2^{NN} - w_2^{NP} \\
N & \text{if } \varepsilon_2 \leq w_2^{NN} - w_2^{NP}
\end{cases}$$
These lead to the following Nash Equilibria in pure strategies:

- \((N, N)\) ⇔ \(\varepsilon_1 \leq w_1^{NN} - w_1^{PN}\) and \(\varepsilon_2 \leq w_2^{NN} - w_2^{NP}\)
- \((P, P)\) ⇔ \(\varepsilon_1 \geq w_1^{NP} - w_1^{PP}\) and \(\varepsilon_2 \geq w_2^{PN} - w_2^{PP}\)
- \((N, P)\) ⇔ \(\varepsilon_1 \leq w_1^{NP} - w_1^{PP}\) and \(\varepsilon_2 \geq w_2^{NN} - w_2^{NP}\)
- \((P, N)\) ⇔ \(\varepsilon_1 \geq w_1^{NN} - w_1^{PN}\) and \(\varepsilon_2 \leq w_2^{PN} - w_2^{PP}\)

Therefore, the equilibrium correspondence \(\varepsilon \mapsto G(\varepsilon|x; \theta)\) depends on the rankings of the terms \(w_i^r - w_i^s, i \in 1, 2\) and \(r, s \in \{NN, NP, PN, PP\}\).

It can also be shown that there exists a unique Nash Equilibrium in mixed strategies as follows. A mixed profile \((\alpha P + (1 - \alpha)N; \beta P + (1 - \beta)N)\) is a Nash equilibrium if and only if

\[
\alpha = \frac{\varepsilon_2 - (w_2^{NN} - w_2^{NP})}{(w_2^{PP} - w_1^{PP}) - (w_2^{NN} - w_2^{NP})} \quad \text{and} \quad \beta = \frac{\varepsilon_1 - (w_1^{NN} - w_1^{PN})}{(w_1^{PP} - w_1^{PP}) - (w_1^{NN} - w_1^{PP})},
\]

the denominators are non zero and

\[
\min \{w_1^{NN} - w_1^{PP}; w_2^{NP} - w_2^{PP}\} < \varepsilon_1 < \max \{w_1^{NN} - w_1^{PN}; w_1^{NP} - w_1^{PP}\}
\]
\[
\min \{w_2^{NN} - w_2^{PP}; w_2^{NP} - w_2^{PP}\} < \varepsilon_2 < \max \{w_2^{NN} - w_2^{PN}; w_2^{NP} - w_2^{PP}\}
\]

Each action profile results in a (almost surely) unique care option choice, hence for each participation shock \(\varepsilon\), we can derive \(G(\varepsilon|x; \theta)\) as the set of probability measures on the set of care options \(\{0, 1, 2, 3\}\) induced by mixed strategy profiles, which are probabilities on the set of participation profiles \(\{NN, NP, PN, PP\}\).

5. Specification

We provide estimates for two specifications of the utility matrix presented in this paragraph.
5.1. **Specification 1.** The first matrix of interest is the following:

\[
V (X; \theta) = \begin{pmatrix}
  \beta_{00} + \beta_m DM + \beta_{ah} DA & \alpha + \psi DS_1 & \psi DS_2 & 0 \\
  \beta_m DM + \beta_{ah} DA & \beta_{11} + \psi_1 DD_1 & 0 & 0 \\
  \beta_m DM + \beta_{ah} DA & 0 & \beta_{11} + \psi_1 DD_2 & 0 \\
  & & & + \beta_{ac} DA
\end{pmatrix}
\]

Recall that the columns indicate the options, in the following order \{0, 1, 2, 3\}, and the rows represent each member of the family, in the following order Parent, Child 1, Child 2. For example, the value the first born child (family member 1) living less than 30 minutes away from the parent’s home attaches to the fact that she takes care of a non-disabled, non-married parent is measured by \(\beta_{11}\), whereas for a disabled parent, it is \(\beta_{11} + \beta_{ac}\). Note some implications of the model:

1. When the parent is unmarried and has “no serious disability” (DA = 0), the children are indifferent between the option where they are not in charge of the parent.
2. The presence of problems with ADL affects the utility of the family in two ways. First, the utility of each member is affected by the term \(\beta_{ah}\), if the option “live alone” is preferred. Second, if a child is chosen as primary care giver, she will bear the “cost” (under the hypothesis that \(\beta_{ac} < 0\)) of taking care of the parent with a disability. This cost is transferred to the Nursing Home, if this option is the one preferred by the family.
3. The parameter \(\beta_m\) associated to the variable DM measures the additional utility for all family member to choose the option “live alone”, when a spouse is present in the same household.
(4) The effect of distance on children’s utility is introduced in the valuation of the primary care giver. \( \psi_d \) measures a cost for child \( i \) to travel (30 minutes or more) to provide care for the parent on a regular basis.

(5) As mentioned earlier, we introduce the parameter \( \alpha \) which allows for a preference of the parent for the oldest child. \( \alpha \) measures the incremental utility for the parent of being taken care of by the firstborn child, as compared with the second born child. This specification allows for the presence of favoritism, as defined by Li, Rosenzweig, and Zhang (2010). In their words: “favoritism exists if the parent derives more utility by spending the same time with [...] one child versus another. Such favoritism could be based on the child’s endowment”, the endowment here being simply birth order. The gender effect is introduced in the same manner.

5.2. **Specification 2.** We summarize the second specification of the utilities that we study in the following matrix:

\[
V(X; \theta) = \begin{bmatrix}
\beta_{00} + \beta_{m} DM + \beta_{ah} DA \\
\alpha + \psi_s DS_1 \\
\psi_s DS_2 \\
\delta V_{00} \\
\beta_{11} + \psi_1 DD_1 + \beta_a DA \\
\delta V_{03} \\
\delta V_{04} \\
\delta V_{00} + \delta V_{03} \\
\beta_{11} + \psi_1 DD_2 + \beta_a DA \\
\delta V_{04} + \delta V_{04}
\end{bmatrix}
\]

The main difference with the first model is the introduction of the term \( \delta V_{0i} \) in the evaluation of option \( j \) by child \( i \). This term measures the degree of altruism of the children. An individual shows altruism when their utility depends directly on someone else’s. The children are altruistic if \( \delta > 0 \).

We introduce the idiosyncratic part of the utility through two terms. First, random participation benefits for each child \( \varepsilon_i, i \in \{1, 2\} \). Second, error terms \( u_j, j \in \{0, 1, 2, 3\} \), which measure unobserved components of the family evaluation of the alternatives. We
Assume:
\[\varepsilon_i \sim iid N(\mu, \sigma_{\varepsilon}) ; \quad u_j \sim iid N(0, \sigma_u)\]

As mentioned above, we introduce in the specification a random error term to account for misclassification of the gender of children. Recall that we include a parameter \(\delta\), measuring the preference of the parent for a female caretaker, so that we have:

\[V_{0,i} = V'_{0,i} + \psi_s D\bar{S}_i \text{ for } i \in \{1, 2\}\]

where \(V'_{0,i}\) sums up the remaining terms entering the parent’s utility. Suppose that \((D\bar{S}_i)^*\) is the reported value of child \(i\)’s gender, when \(D\bar{S}_i\) is the true value of gender. It follows that:

\[V_{0,i} = V'_{0,i} + \psi_s (D\bar{S}_i)^* + \xi_i \text{ for } i \in \{1, 2\} \text{ with } \xi_i = \psi_s (D\bar{S}_i - (D\bar{S}_i)^*)\]

Supposing that true gender is misreported with probability \(p_\xi\) and that measurement error is independent of true gender, we have \(\xi_i = 0\) with probability \(1 - p_\xi\) and \(\psi_s (1 - 2D\bar{S}_i)\) with probability \(p_\xi\). Failing to take into account this measurement error may lead to biased estimates of the parameter \(\psi_s\) and may worsen the identification issue.

6. Results

We perform the estimation of the two specifications introduced earlier, under different values of the mean and variance of the error term. To alleviate the computational burden, we first test the significance of some of the individual parameters by checking whether the hyper planes defined by \(\theta_i = 0\) - where \(\theta_i\) is a component of \(\theta\) - intersect the 95\% confidence region. In practice, this amounts to building a constrained confidence region under the Null Hypothesis. We fail to reject the Null Hypothesis if the estimation procedure returns a non-empty set. We then obtain a constrained confidence region for the remaining parameters.

6.1. Specification 1. For each value of mean and variance of the error term, we find a non-empty intersection between the confidence region and the hyperplane defined by \(\beta_{11} = 0\). This means we fail to reject (at the 5\% level) the null hypothesis that there is no additional constant disutility for a child to take care of an elderly parent. Since, this hypothesis is not rejected, we obtain a constrained confidence region for the remaining parameters. We then
obtain confidence regions for different values of $\beta_{11}$ and discuss the latter’s effect on the regions. We note that the Null hypothesis $H_0: \beta_{00} = 0$ is always rejected. Hence, when we control for all other effects, parents are not indifferent between the first two options. They show a clear preference in favor of living in their own home (option called “living alone”) instead of living in a nursing home ($\beta_{00}$ is always positive). The results we present are then for given values of $\beta_{00}$. We provide an insight of how different values of this parameter change the results.

We report the range for each parameters in table 2. Note that the identified set is not a compact set. In particular, $\beta_{ac}$, $\beta_{ah}$, $\beta_{m}$ and $\psi$ are allowed to diverge to $-\infty$.

Results are generally consistent with expectations and previous results on the subject. Namely:

(1) The existence of several problems with the parent’s functional ability is a key determinant of the decision to enter a nursing home. $\beta_{ah}$ and $\beta_{ac}$ are both negative and can both be (very) large. The negative sign of $\beta_{ah}$ captures the fact that a parent’s

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Min</th>
<th>Max</th>
<th>Min</th>
<th>Max</th>
<th>Min</th>
<th>Max</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{00}$</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{ah}$</td>
<td>$-\infty$</td>
<td>-3.57</td>
<td>$-\infty$</td>
<td>-3.57</td>
<td>$-\infty$</td>
<td>-2.86</td>
<td>$-\infty$</td>
<td>-2.14</td>
</tr>
<tr>
<td>$\beta_{ac} = -\beta_{m}$</td>
<td>$-\infty$</td>
<td>-2.86</td>
<td>$-\infty$</td>
<td>-2.86</td>
<td>$-\infty$</td>
<td>-3.57</td>
<td>$-\infty$</td>
<td>-3.57</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.00</td>
<td>8.00</td>
<td>0.00</td>
<td>8.00</td>
<td>1.00</td>
<td>5.00</td>
<td>0.00</td>
<td>4.00</td>
</tr>
<tr>
<td>$\psi_s$</td>
<td>0.00</td>
<td>5.00</td>
<td>0.00</td>
<td>4.00</td>
<td>1.00</td>
<td>4.00</td>
<td>0.00</td>
<td>2.00</td>
</tr>
<tr>
<td>$\psi_d$</td>
<td>$-\infty$</td>
<td>-2.86</td>
<td>$-\infty$</td>
<td>-2.14</td>
<td>$-\infty$</td>
<td>-1.43</td>
<td>$-\infty$</td>
<td>-3.57</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\sigma_\varepsilon$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\sigma_\varepsilon^2$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>$p_\xi$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 2. Parameters Range for estimation of Specification 1 at $\beta_{11} = 0$, $\beta_{ac} = -\beta_{m}$ and for different values of the error terms and of $\beta_{00}$. 
disability increases the value of care provided by the family or a specialized institution. In addition, $\beta_{ac} < 0$ means that the disability entails a utility cost for the child if he is chosen as primary care giver. In other words, holding all else equal, parental disability decreases the value for family members of being primary care giver by at least 2.86 utils. Placing the elderly parent in a nursing home transfers this cost to the institution. If we think that the expansion of life expectancy observed in past years is correlated with the appearance of more functional ability problems, this may partly explain the trend away from care provision by children toward other alternatives.

(2) Parameter $\beta_m$ associated with the parent living with a spouse is positive and large. This implies that married parents are more likely to stay at home. In families where the parent is disabled, the effect of living with the spouse, somehow compensates the disutility created by the disability factor and, all other variables controlled, preserve the incentive for parents to live at home.

(3) While we cannot rule out parents being indifferent to the gender or birth order of their primary care giver, estimation shows a tilt of the confidence interval toward positive values for both parameters, with a possible positive and large magnitude of the parameter $\alpha$. In case $\mu = -1$ and $\sigma_u^2 = 0.25$, the data reveal that parents exhibit a preference for an older and for a female care giver.

(4) Children living more than 30 minutes from the parents are less likely to provide care than those living closer to the parents. Distance has a (possibly strong) disutility effect on children’s incentives to participate in the care decision. All else being equal, moving 30 minutes away from the parents reduces by at least 1.43 utils, the utility of a child when providing care.

(5) Note $H_0 : \sigma_u^2 = 0$ is rejected. The magnitude of the unobserved idiosyncratic term $u$ is linked to the value of the parameter $\beta_{00}$. Values of $\beta_{00}$ higher than 1 are rejected for $\sigma_u^2 = 0.25$, and are the only values admissible if $\sigma_u = 1$. Recall that $\beta_{00}$ measures the preference of the parent for living at their own home. Higher values of this parameter induces then higher probabilities for the choice of the option “Living alone” i.e $P(\text{Option “Living alone”}) \to 1$. Higher magnitude of the unobserved
heterogeneity is therefore needed to make sense of the difference in choices that we observe in the data. On the other hand, the unobserved heterogeneity need not have a “large” variance when $\beta_{00}$ is close to 0, which suggests that the model explains the variance in the data well.

The shape of the confidence region also conveys a considerable amount of information. We plotted in Figure 1 and 2, two dimensional and three dimensional projections and cuts of the confidence region for column 2 of table 2, i.e $\mu = 0$, $\sigma^2 = 1$, $\sigma^2_u = 1$.

Of great interest is the projection of the identified set in the plan $\beta_{ah}, \beta_m$. Figure 2(a) reveals a linear relation between the two parameters of the type : $\beta_{ah} = -\beta_m$. The estimation rejects models for which the absolute value of the two parameters are significatively different. The data suggest therefore that the disutility induced by the disability of the parent can be entirely compensated by the presence of a spouse in the same household.

Notice the triangular shape of the region plotted in figure 2(b) which entails that simultaneous large values of $\psi_s$ and $\alpha$ are rejected. This finding means that only one of the effects (gender or birth order) can be large, not both. In other words, firstborn daughter are not the only possible care givers. Note also that both effects can be very small, though not jointly insignificant.

We observe similar types of constraints for the pairs $(\alpha, \beta_{ah})$, $(\alpha, \beta_{ac})$, $(\alpha, \psi_d)$, $(\psi_s, \beta_{ac})$, $(\psi_s, \psi_d)$ as large values of parameters $\alpha$ or $\psi_s$ are only permitted when the other parameters are jointly large (see figure 2(c) to 2(f)). For example, we obtain a constrained confidence region at $\beta_{ac} = -3.5$. The ranges for the two parameters, $\alpha$ and $\psi_s$, are tighter, as $\alpha \in [1, 2]$ and $\psi \in [0, 1]$.

Figure 3 shows the effect of the variation of parameter $\beta_{11}$ on $\psi_s$ and $\alpha$. Recall that $\beta_{11}$ represents a fixed cost or benefit for the child chosen as care giver. We observe negative relations between $\beta_{11}$ and $\psi_s$, and $\beta_{11}$ and $\alpha$. Negative values of $\psi_s$ and $\alpha$ are only admissible for positive values of $\beta_{11}$. Hence a model where parents exhibit no favoritism for a daughter and/or a firstborn, or favoritism for a son and/or a second born, will be consistent with our data if and only if there exist a strictly positive constant benefit for a child to be caregiver. In the case where $\beta_{11} << 0$, the null hypothesis of indifference to gender and birth order of
the primary care giver is rejected and the minimum value of both effects (gender and birth order) increases with more negative values of the fixed cost (see figure 3(c)).

6.2. Specification 2. As discussed earlier, the second specification analyzes primarily altruistic behavior on the part of the children. The parent’s utility enters in the children’s evaluation through the parameter $\delta$. The greater $\delta$, the more the parent’s preferences influence the family’s decision. Notice that in the case where $\delta = 0$, both children value all options identically, where they are not in charge of the parent, irrespective of the parent characteristics, in particular as pertains to disability.

Not surprisingly, as in the previous specification, we reject $H_0 : \beta_{00} = 0$ (at the 5% level) as $\beta_{00}$ is always positive, and we fail to reject (at the 5% level) the null hypothesis that there is no additional constant disutility for a child to take care of the elderly parent ($H_0 : \beta_{11} = 0$). Effects of $\beta_{11}$ on remaining parameters is the same as discussed above. We obtain constrained confidence region for the remaining parameters. We report the range for each parameters in table 3. Note again, that the identified set is not a compact set as $\beta_{ac}, \beta_{ah}, \beta_m$ and $\psi$ are allowed to diverge to $-\infty$.

The data seem to comply with both types of models: altruistic and non altruistic. However, the hypothesis of non-altruism is rejected for some values of the error term, namely $\sigma_u^2 = 1$. That is, in the case where the variance of the unobserved heterogeneity is higher ($\sigma_u^2 = 1$ compared to $\sigma_u^2 = 0.25$), the behavior of children can only be rationalized in a framework where they act according to altruistic motives. On the other hand, high degrees of altruism ($\delta > 0.78$) are not permitted for a low variance of $u$. This feature can be understood by relating to the parameter $\beta_{00}$. We discussed in the previous section the fact that $\sigma_u^2$ is related to $\beta_{00}$, as large values of $\sigma_u^2$ imply large values of $\beta_{00}$ (see first line of table 3). Figure 5(a) shows a projection of the confidence region in the space $(\beta_{00}, \delta)$. We can observe a negative dependance between the two parameters. Both parameters cannot both be very large or both very small. This suggests that a model compatible with the data is either one where children are predominantly selfish and parents have a strong preference for living alone, or one where children are fairly altruistic and parents have only a weak preference for living alone, or a middle point between the two.
In this specification, we reject the joint null hypothesis of indifference of the parent to gender and birth order of the primary care giver when there exists no constant cost or benefit associated with providing care ($\beta_{11} = 0$). The estimation suggests indeed a preference of parents for an older child and for a daughter ($\alpha > 0$ and $\psi_s > 0$). Negative values of $\beta_{11}$ (i.e. existence of constant disutility for caregivers) are associated with larger minimum values of $\alpha$ and $\beta$, hence stronger preference for first-born and female caregivers.

Confidence intervals for the other parameters do not differ significantly from the previous specification. We again reject the hypotheses that distance (between parent and caretaker) and parent’s disability are insignificant, as the range for $\psi_d$ is $\psi_d \in (-\infty, -1.43]$ and the range for $\beta_a$ is $\beta_a \in (-\infty, -2.86]$. We also reject the hypothesis that living with a spouse does not increase the value for a parent to remain autonomous.

We observe very similar shapes of the confidence region as for the previous specification. Most of the analysis above still applies to the second specification. Notice that in the non altruistic case, i.e $\delta = 0$, only very large effects of the parent’s disability are consistent
with the data (see Figure 5(b)). In other words, parental disability reduces the utility of providing care for the non altruistic type of children more than for the altruistic type.

References

DUKE (1999): “National Long Term Care Survey,” Public use data set produced and distributed by the Duke University Center for Demographic Studies with funding from the National Institute on Aging under Grant No. U01-AG007198.


Figure 1. Three dimensional representations of the confidence region for Specification 1 at $\beta_0 = 3$, $\beta_1 = 0$, $\mu = 0$, $\sigma_\varepsilon = 1$, $\sigma_u = 1$, $p_\xi = 0.1$
Figure 2. Two dimensional representations of the confidence region for Specification 1 at $\beta_{00} = 3$, $\beta_{11} = 0$, $\mu = 0$, $\sigma_\varepsilon = 1$, $\sigma_u = 1$, $p_\xi = 0.1$
Figure 3. Parameter $\beta_{11}$ in relation with other parameters: Specification
1 at $\beta_{00} = 3$, $\mu = 0$, $\sigma_{\varepsilon} = 1$, $\sigma_u = 1$, $p_\xi = 0.1$
Figure 4. Three dimensional representations of the confidence region for Specification 2 at $\beta_{11} = 0$, $\mu = 0$, $\sigma_\varepsilon = 1$, $\sigma_u = 1$, $p_\xi = 0.1$
Figure 5. Parameter $\delta$ in relation with other parameters: Specification 2
at $\beta_{11} = 0$, $\mu = 0$, $\sigma_c = 1$, $\sigma_u = 1$, $p_\xi = 0.1$