Australian School of Business

Working Paper

Never Stand Still Australian School of Business

Australian School of Business Research Paper No. 2012 ECON 20

Inter Vivos Health Transfers: Final Days of Japanese Elderly Parents

Shiko Maruyama

Inter Vivos Health Transfers: Final Days of Japanese Elderly Parents

Shiko Maruyama[1]

University of New South Wales, Australia

September 15, 2012

Abstract

Empirical evidence of the effect of intergenerational coresidence by elderly parents and their adult children on parental health remains inconclusive. This study provides a new estimate of the coresidence effect by addressing non-random selection and heterogeneity in the treatment effect, and investigates why coresidence can be detrimental. Studying Japanese data reveals a negative coresidence effect on the treated. I argue that coresidence may worsen parental health because care burdens on children create disincentives for parents to invest in their longevity. The results support this theory: ceteris paribus, the coresidence effect is negatively associated with coresidence burdens such as disability.

[1] The author gratefully acknowledges financial support from the Australian Research Council's *Discovery Projects* funding scheme (project number DP110100773). Helpful comments from Susumu Imai, Hajime Katayama, and Kevin Lang are gratefully acknowledged. I would like to thank the Nihon University Center for Information Networking for the use of the Nihon University Japanese Longitudinal Study of Aging data.
E-mail: *s.maruyama@unsw.edu.au*.

1 Introduction

In the theory of health capital, individuals make health-related decisions taking their future consequences into consideration. Expected marginal benefits and costs of health investments determine the optimal trajectory of health over one's remaining life and thereby determine one's demand for longevity (Grossman, 1972; Ehrlich and Chuma, 1990). The literature has discussed various sources of benefits and costs of health investments, but the presence of family members has rarely received attention in this context. If elderly parents are altruistic toward their children, how does the presence of beloved children affect their health related decisions and demand for longevity? What if the parents have high care needs?

Intergenerational coresidence by elderly parents and their adult children are often formed to meet parental care needs and to provide family support for security in old age. However, despite the wide recognition of the importance of informal care, despite the general notion that coresiding children are caregivers who provide comprehensive, essential assistance to elderly parents, and despite numerous studies on the coresidence effect conducted in sociology, demography, public health, and gerontology, the empirical literature has yet to reach any general consensus on the effect of coresidence with an adult child on parental health, let alone an understanding of the underlying mechanism.

Two questions are asked in this study. First, is the coresidence effect indeed positive? Second, how do we explain the sign and size of the coresidence effect? If coresidence is something that is formed to meet the care needs of a parent, the existing literature should have found broad unambiguous evidence of positive coresidence effects. In this study, I propose a simple theory as to why intergenerational coresidence may affect parental health negatively, and derive testable implications. I then provide a new estimate of the treatment effect of coresidence and test the implications of the theory.

The burden of informal caregiving borne by carers is widely documented. Caregiving may significantly affect the labor supply decisions of carers (Ettner, 1995) and their quality of life (Kenny et al., 2010). I argue that in this setup, coresidence may cause parents with high care needs to compromise remaining life years, because prolonged life means additional caregiving burdens on their children, which create disincentives for parents to invest in their health and live long. Parents may reduce their health investments through two different channels. On the one hand, altruistic parents may voluntarily reward their children's dedication. On the other hand, non-altruistic parents may passively respond to pressures from grudging children, or in other words, pay the price necessary to maintain coresidence as a quid pro quo. At the same time, for parents with no care needs, coresidence will create additional incentives to invest in their health in order to stay healthy so that parents can delay the onset of potential caregiving burdens on coresiding children. This theory, hence, predicts a negative relationship between parental care needs and the coresidence effect.

The existing literature on the coresidence effect has so far paid limited attention to the non-experimental nature of data, and this may be the reason for the mixed results. I advance the literature by relying on the program evaluation literature. In my setup, coresidence by an elderly parent and an adult child is referred to as the "treatment". The treatment group consists of elderly parents who live with a child and the control group consists of elderly parents who have at least one child but live with no child.

Two econometric issues might have biased the results of previous studies. The first source of bias is non-random selection into coresidence, especially selection by unobservable factors, which refers to the possibility that the treated group and the control group are systematically different even after controlling for observable characteristics. For instance, suppose sociable parents are more likely to coreside with children. If an econometrician does not observe their sociability and if sociability contributes to future health, the coresidence effect estimate will be overestimated. The existing literature has overlooked this potential bias with the exception of Do and Malhotra (2012) and Johar and Maruyama (2011). Second, given the great heterogeneity in the motives and behavior of families, it is reasonable to suspect that the coresidence effect would vary by various observable and unobservable family characteristics. Suppose different families experience different coresidence effects but the distribution of the coresidence effect centers around zero. Then the resulting overall effect averages zero, and in a homogeneous treatment effect framework, an econometrician will incorrectly conclude that coresidence has no sizable effect. To address these two issues, I employ the factor structure model of Aakvik et al. (2005). This model provides a flexible yet parsimonious and tractable specification that accounts for non-random selection into coresidence and heterogeneity in the coresidence effect.

I use the two recent waves of the Nihon University Japanese Longitudinal Study of Aging (NUJLSOA), a nationally representative panel survey of elderly Japanese. Japan features East Asian close family ties and still has one of the highest intergenerational coresidence rates among developed countries, thus providing an ideal setting for this study. The population in the analysis consists of elderly Japanese who have at least one adult child in the base year, 2003. The dependent variable is the health outcome three years later, which is a combination of the mortality status and five-categorical self-assessed health scores if the parent survives. I build on the framework of Aakvik et al. (2005) by adding this ordered dependent variable structure so that all the available information is fully exploited to facilitate the identification of the model. This combination of mortality and self-reported health also allows me to avoid truncation and attrition bias due to death. To address non-random selection into coresidence, I use land prices and rurality as instruments.

The main results are as follows. First, the point estimate of the average treatment effect of coresidence on the three-year survival is -0.0158 with a standard error of 0.0180. If coresidence is randomly assigned, the three-year survival rate of elderly parents *decreases* by 1.58 percentage points on average. Although the estimate is not statistically significant, the effect is not larger than $+2.0$ percentage points at the 95% confidence level; thus, there is no support for a positive coresidence effect. Second, the estimated average treatment effect *on the treated* is statistically significant and strongly negative. The three year survival rate of elderly parents in coresidence would be 5.2 percentage points higher (1.7 percentage point per annum) if they lived independently. Third, the results are consistent with the theory: ceteris paribus, the coresidence effect is negatively associated with coresidence burdens such as disability. In particular, vulnerable parents with high care needs who have limited potential for making financial compensation, typically older widowed mothers with disability, are most likely to be in coresidence and experience a large negative coresidence effect.

2 Mixed Findings on the Effect of Coresidence

Researchers in various fields have studied the effect of living arrangements on the elderly's health. These studies cover both Western and Asian countries and look at various population and various outcomes. The studies before 2000 mostly rely on the cross-sectional approach, and hence are likely to suffer from reverse causation — the possibility that coresidence occurs in response to declining parental health. Recently, the use of panel data has become standard. Table 1 summarizes recent studies as well as some early panel studies. Their findings are indeed mixed and it is not easy to draw a general conclusion. For example, Li et al. (2009) study the oldest old in China and find that whereas having a spouse in the household provides the best health protection, the effect of living with children is mixed. Similar results are found for mental health of parents in Singapore (Chan et al., 2011). Using data from Israel, Walter-Ginzburg et al. (2002) find that elderly parents who live with a child have higher mortality risk than parents living alone. In Japan, elderly mothers are found to have more than double the risk of heart disease when they coreside with children (Ikeda et al., 2009). Elderly Japanese are also found to have higher mortality rates when they are cared for by daughters compared to living with a spouse (Nishi et al., 2010).

To the best of my knowledge, Do and Malhotra (2012) and Johar and Maruyama (2011) are the only studies that attempt to correct for non-random selection into coresidence. Do and Malhotra (2012) use the number of sons as an instrument in their two-stage least squares framework and argue that it is an appropriate instrument in the Korean setting because it is related to traditional rules of coresidence and should not directly affect parental health. They find that living with a child reduces depressive symptoms among South Korean widowed elderly mothers.[2] Johar and Maruyama (2011) apply the same econometric model as the present study to Indonesian data. Using community level traditions and customs as instruments, we find a significant, negative coresidence effect on the treated.

[2] Informal care provided by children living elsewhere may affect the exogeneity of this instrument.

[Insert Table 1]

3 The Determinants of Coresidence Effect

It is difficult to cast doubt on the positive, vital role of coresidence with children. In ageing societies, care and attention provided by adult children remains an important source of support in old age. While the demand for aged care is growing at an unprecedented rate, around 80% of the hours of care are provided informally, with children providing 41% of all informal care in the U.S., 43% in the U.K., and 60% in Japan (OECD, 2005). In Japan, informal care provided by children quite often coincides with parent-child coresidence. Though such coresidence is declining in Japan (Johar et al., 2010), for elderly Japanese receiving any nursing care, the most common primary caregiver is still a coresident child or a coresident child's spouse (32%) (The Ministry of Health, Labour and Welfare, 2008).

Nevertheless, the literature has not found robust support for the beneficial effect of coresidence on parental health. In this section, I propose a possible explanation based on the theory of the demand for longevity as to why intergenerational coresidence may negatively affect parental health and derive testable implications.

3.1 Theory of the Demand for Longevity

Coresidence with a child may or may not benefit parents. Coresidence benefits come not only from informal health care but also domestic assistance, risk sharing, economy of scale, and probably most important, companionship, comfort, and other mental and emotional benefits. Coresidence may also create costs such as reduced privacy, loss of self-esteem, and conflicts in social relations. Similarly, coresidence may or may not benefit children.

Children may benefit from housing, economies of scale, and additional time spent with parents. If parents have care needs, however, caregiving burdens reduce the utility of children in coresidence, as widely documented (e.g., Ettner, 1995; Kenny et al., 2010). Despite the reduced utility due to physical and mental exertion and opportunity costs, children live with parents with care needs for various reasons, such as the consumption value of shared time, altruism, filial piety, social and cultural norms, and inheritance anticipation, but children will avoid coresidence if their expected lifetime utility loss exceeds a certain threshold.

A parent makes health-related decisions taking their marginal benefits and costs into consideration. Health-related decisions cover various aspects of one's life: regular exercise, healthy diet, engaging in social activities, frequent medical checkups, search for good doctors, hospitals, and insurance plans, etc. Health-related decisions affect the future trajectory of health over one's remaining life and thereby determine one's longevity.

The parent's problem is to optimally allocate her endowed resources, such as wealth, time, and effort, to maximize her expected lifetime utility. The parent's lifetime utility is an increasing function of her remaining life years and the child's utility, if she has altruism toward her child. The child's expected lifetime net utility gain due to coresidence is the per period net utility gain from coresidence multiplied by the expected remaining life years of the parent.

This structure creates a trade-off for the parent regarding longevity, which is the central idea of the theory. Although additional life years increase the lifetime utility of the parent, when she has care needs and lives with her child, the child incurs a coresidence burden and this disutility of the child provides the parent with a disincentive to make health investments through the following two channels: altruism and exchange.

First, parental altruism is likely to reduce the marginal benefit of health investments, because the prolonged coresidence expected from health investments further burdens her child. Second, even if the parent is not altruistic, she must compensate the child to maintain coresidence, or the parent needs to *buy* coresidence, by providing financial transfers, by offering time-related services, or by shortening the remaining life years. This "exchange channel" also lowers the marginal benefit of health investments because a prolonged life increases the *price* of coresidence.

These disincentives to make health investments may generate a negative coresidence effect. If the expected lifetime utility gain from coresidence is sufficiently large, the parent may prefer to live with her child by sacrificing her life expectancy to lessen the child's burden. I call such reduction of longevity for the sake of their children an *inter vivos health transfer*.

The idea of the rational decision on longevity and trade-off between the length and quality of life dates back to the health capital model by Grossman (1972), in which an individual chooses the utility-maximizing allocation of time and wealth, considering the marginal benefits and costs of health investments over the life-cycle, ultimately choosing the optimal date of death.[3]

The literature on resource allocation within families also dates back to the 1970s. The studies by Becker (1974) and Barro (1974) show that altruistically linked families pool their resources, and public transfers across generations are neutralized by within-family

[3] Another classical example of the rational longevity decision is the study of rational suicide by Hamermesh and Soss (1974). In rational suicide models, death occurs when the remaining lifetime utility falls below a certain threshold.

transfers (the Ricardian equivalence).[4] The literature has since discussed various parental motives behind the intergenerational financial transfers that parents make to their children in the form of bequests and inter vivos transfers. The exchange hypothesis claims that non-altruistic parents may "buy" time-related services, such as informal care, by compensating their children's disutility with additional resource transfers (Cox, 1987; Bernheim et al., 1985). I contribute to this literature by emphasizing the role of parental health as a medium of intergenerational transfers.

3.2 Testable Implications

In the discussion so far, I have provided an explanation as to why elderly parents who live with their children may "choose" shorter lives, pointing out the trade-off in the demand for longevity. Whether the decision is governed by altruism or an exchange motive, compared with parents who live with no children, parents in coresidence have greater disincentives to live long. This theory serves as a possible explanation if we empirically observe a negative coresidence effect.

In the rest of this section, using an informal *comparative statics* argument, I derive testable implications that allow me to test the validity of the theory.

Coresidence Burdens The parent's trade-off hinges on the child's utility cost, which is the per period coresidence burden times the expected length of coresidence. This implies that a larger per period burden will lead to a worse coresidence effect because of larger marginal costs of additional life years.

[4] Note that the term "altruism" here has no moral connotation. It implies that an altruistic parent maximizes her utility "selfishly" by "consuming" her child's utility. The real motive may stem from enthusiastic love or moral responsibility associated with the role in society and cultural tradition.

Prediction 1 *All else equal, a larger per period coresidence burden leads to a worse coresidence effect on parental health.*

In reality, the per period caregiving burden is not a constant; disability and care burdens worsen over time, often together with health decline. Thus, reduced health investments may enlarge care burdens, and parents may have a larger incentive to maintain their health to reduce future burdens on children. Although this creates ambiguity in the theoretical prediction of the coresidence effect, under fairly mild assumptions,[5] elderly parents with no care needs have a larger incentive than those with disabilities to remain healthy so that they can delay the outset of caregiving burdens, whereas a severe disability creates a large disincentive to make health investments. Therefore, even if the per period burden changes over time, a larger per period burden is likely to lead to a worse coresidence effect. In this setup, the coresidence effect is likely to be positive for parents without care needs.

To test this prediction, I look at a measure of disability based on activities of daily living and widowhood. After controlling for other health status variables, these variables should capture how coresidence is burdensome to children. Marital status matters because a spouse provides care quite often. If the estimated coresidence effect worsens with these coresidence burden measures, we can conclude that data is consistent with the theory.

Opportunity Costs of Children By the same token, the health reduction is expected to be larger when the child's opportunity cost is larger. If the parent is altruistic, a larger disutility put on the child also reduces the utility of the parent. Alternatively, a higher opportunity cost implies that the coresidence burden is more likely to exceed the child's

[5] Theoretical predictions depend on the dynamic transition processes of health, disability, and caregiving utility costs.

tolerance threshold. In either case, when the child who provides care has something else more important, such as work opportunities in distant cities,[6] it dampens the parent's incentive to make health investments.

Prediction 2 *All else equal, a larger per period opportunity cost for the child leads to a larger negative coresidence effect on parental health.*

To test this prediction, I use the age of the child, assuming that the opportunity cost is larger for younger children because compromise in their career development has larger consequences on their lifetime earnings when they are younger. The data do not provide any income or wage information for children.

Wealth The theoretical prediction regarding parental wealth is ambiguous. On one hand, instead of health transfers, wealthy parents can reward their children by financial transfers (a *King Lear effect* in Bernheim et al., 1985). This is plausible because for elderly individuals, the marginal utility of consumption is small relative to the marginal utility of life expectancy (Hall and Jones, 2007; Ehrlich and Chuma, 1990). Consequently, parents with larger wealth will not need to compromise their longevity and will enjoy a better coresidence effect.

On the other hand, a longer life eats up parental wealth and reduces the amount of bequest transfers. Children of wealthy parents can receive larger transfers if their parents die earlier (the *Prince Hal situation* in Goody, 1987). Hence the possible bequest transfers amplify the trade-off of altruistic parents, and larger wealth may worsen the coresidence effect.

[6] In the data used in this study, I do not observe many parents moving to the children's location.

Furthermore, as I discuss later, a negative relationship between wealth and the coresidence effect is also consistent with elderly abuse and neglect, an alternative explanation of a negative coresidence effect. When coresiding children have a means to negatively affect parental health, their incentive to do so is larger when their parents have larger wealth.[7]

Because the data provide no direct measure of wealth, to investigate the wealth effect, I examine how the coresidence effect varies with parents' education, housing asset, and income.

Bargaining If a family is not (strongly) altruistically linked, the number of children in the family affects the bargaining power of the parent. Specifically, parents with fewer children will have smaller bargaining power and will need to reduce their longevity to a greater extent. Children with no siblings have the greatest bargaining power and the parent faces the highest *price* for coresidence. This hypothesis can be tested by the number of children.

Prediction 3 *All else equal, when intergenerational family bargaining dominates family altruism, a parent with fewer children experiences a worse coresidence effect.*

Note that if a family is predominantly altruistically linked, Predictin 3 may not hold, but the theory can still explain a negative coresidence effect and Predictions 1 and 2 still hold.

[7] In Japanese civil law, the decedent's next-of-kin has a legal right share, or the *legitim* — a statutory fraction of the decedent's gross estate from which the decedent cannot disinherit his next-of-kin.

4 Empirical Model

4.1 Latent Health Model with Heterogeneous Treatment Effects

To estimate the coresidence effect and address non-random selection and heterogeneity in the coresidence effect, I extend the factor structure framework of Aakvik et al. (2005), which provides a flexible yet parsimonious and tractable specification that yields easily interpretable expressions for treatment parameters. The original model by Aakvik et al. (2005) simultaneously estimates three binary equations for (1) selection, (2) treated outcomes, and (3) untreated outcomes. Because the identification of such a model is not necessarily easy, I extend the original binary outcome model by allowing for an ordered categorical outcome, combining categorical health status with the survival outcome. The treatment group comprises elderly parents who live with a child, and the control group is elderly parents who have at least one child but live with no child. For each elderly parent i, let D_i be the treatment status: 1 for coresidence with a child and 0 otherwise. For each parent i, define two potential outcomes (Y_{0i}, Y_{1i}) corresponding to the health outcomes after three years in the non-coresidence and coresidence states, respectively. Y_{0i} and Y_{1i} measure subjective health and take six possible values: 5 for very healthy, 4 for healthy, 3 for average, 2 for poor, 1 for very poor, and 0 for death. Parent i may be married, but because only one parent from each household is included in the analysis, Y_{0i} and Y_{1i} are assumed to be independent across observations.

Because all dependent variables are discrete, a latent index framework is used. The coresidence equation is specified as

$$
\begin{aligned}
D_i^* &= Z_i\beta_D - U_{Di}, \qquad (1)\\
D_i &= 1 \text{ if } D_i^* \geq 0,\ D_i = 0 \text{ otherwise,}
\end{aligned}
$$

where Z_i is a vector of observed characteristics that influence the family's coresidence decision, β_D is a vector of associated parameters, and U_{Di} captures unobserved costs of coresidence.[8] The latent variable, D_i^*, measures the net utility of coresidence.

The health outcomes, $Y_{ji}, j = 0, 1$, are specified as follows:

$$
Y_{ji}^* = X_i\beta_j - U_{ji}, \qquad Y_{ji} = \begin{cases} 5 \text{ if } Y_{ji}^* \geq c_4, \\ 4 \text{ if } c_4 > Y_{ji}^* \geq c_3, \\ 3 \text{ if } c_3 > Y_{ji}^* \geq c_2, \\ 2 \text{ if } c_2 > Y_{ji}^* \geq c_1, \\ 1 \text{ if } c_1 > Y_{ji}^* \geq 0, \\ 0 \text{ if } 0 > Y_{ji}^*, \end{cases} \qquad (2)
$$

where X_i is a vector of observed characteristics, β_j denotes its associated parameters, U_{ji} is an error term that captures unobserved health shocks, and c_1, c_2, c_3, and c_4 are cutoff parameters. The exclusion restriction is satisfied when $X_i \subset Z_i$. The latent health variable, Y_{ji}^*, has a structural interpretation: if it is positive, parent i survives, and a larger value indicates better health. The cutoff points determine which category parent i's health falls into. This extension of the original Aakvik et al.'s (2005) model requires only a limited

[8]Following Aakvik et al. (2005), the error terms in the three equations are defined as costs instead of benefits, without loss of generality.

number of fairly reasonable additional assumptions: (1) death is worse than the worst health status;[9] and (2) the cutoff parameters are the same across individuals and across treated and untreated states. In the following estimation, X_i contains a constant term, implying that setting the lowest cutoff to zero does not impose any restriction.

Following Aakvik et al. (2005), the error terms in equations (1) and (2) are assumed to be governed by the following normal factor structure:

$$U_{Di} = -\theta_i + \varepsilon_{Di}, \ U_{0i} = -\alpha_0\theta_i + \varepsilon_{0i}, \text{ and } U_{1i} = -\alpha_1\theta_i + \varepsilon_{1i},$$

where each of $(\theta, \varepsilon_D, \varepsilon_0, \varepsilon_1)$ follows the i.i.d. standard normal.[10] This specification implies

$$\text{Cov}(U_D, U_0) = \alpha_0, \ \text{Cov}(U_D, U_1) = \alpha_1, \text{ and } \text{Cov}(U_0, U_1) = \alpha_0\alpha_1.$$

4.2 Estimation

Estimation relies on the maximum likelihood method. The likelihood function has the form

$$L = \prod_{i=1}^{N} \int \Pr(D_i, Y_i | X_i, Z_i, \theta) \, \phi(\theta) \, d\theta,$$

where $Y_i \equiv (1 - D_i) Y_{0i} + D_i Y_{1i}$ and ϕ denotes the standard normal probability density function. Let Φ denote the standard normal cumulative distribution function. The joint

[9]This is a standard assumption in the health economics literature (e.g. health capital models and the QALY weights literature). Also, besides mortality, subjective health is the most commonly used measure of individual health in the literature (Banks and Smith, 2011). Subjective health summarizes various aspects of individual health and has been found to be highly correlated with life expectancy and the prevalence of chronic diseases.

[10]The i.i.d. normality assumption guarantees the following standard assumptions: (i) (U_D, U_0) and (U_D, U_1) are independent of (Z, X), (ii) Y_0 and Y_1 have finite first moments, and (iii) $1 > \Pr(D = 1|X) > 0$.

probability is given by

$$\Pr\left(D_i, Y_i | X_i, Z_i, \theta_i\right) = \Pr\left(D_i | Z_i, \theta_i\right) \Pr\left(Y_i | X_i, D_i, \theta_i\right),$$

where $\Pr\left(D_i = 1 | Z_i, \theta_i\right) = \Phi\left(Z_i \beta_D + \theta_i\right)$. For $j = 0, 1$,

$$\begin{aligned}
\Pr\left(Y_{ji} = 5 | D_i = j, X_i, \theta_i\right) &= \Phi\left(X_i \beta_j + \alpha_j \theta_i - c_4\right), \\
\Pr\left(Y_{ji} = 4 | D_i = j, X_i, \theta_i\right) &= \Phi\left(X_i \beta_j + \alpha_j \theta_i - c_3\right) - \Phi\left(X_i \beta_j + \alpha_j \theta_i - c_4\right), \\
\Pr\left(Y_{ji} = 3 | D_i = j, X_i, \theta_i\right) &= \Phi\left(X_i \beta_j + \alpha_j \theta_i - c_2\right) - \Phi\left(X_i \beta_j + \alpha_j \theta_i - c_3\right), \\
\Pr\left(Y_{ji} = 2 | D_i = j, X_i, \theta_i\right) &= \Phi\left(X_i \beta_j + \alpha_j \theta_i - c_1\right) - \Phi\left(X_i \beta_j + \alpha_j \theta_i - c_2\right), \\
\Pr\left(Y_{ji} = 1 | D_i = j, X_i, \theta_i\right) &= \Phi\left(X_i \beta_j + \alpha_j \theta_i\right) - \Phi\left(X_i \beta_j + \alpha_j \theta_i - c_1\right), \text{ and} \\
\Pr\left(Y_{ji} = 0 | D_i = j, X_i, \theta_i\right) &= 1 - \Phi\left(X_i \beta_j + \alpha_j \theta_i\right).
\end{aligned}$$

Identification of the model parameters follows from Heckman (1981) if $\varepsilon_D, \varepsilon_0, \varepsilon_1$, and θ are joint normal. To integrate θ_i, numerical approximation by Gauss-Hermite quadrature is used.

4.3 Treatment Parameters

An advantage of the factor structure model is that the average treatment parameters and the distributions of the treatment parameters can be obtained from the estimated structural parameters. Let Δ denote the treatment effect with regard to survival for a given parent: $\Delta = I\left[Y_1 \geq 1\right] - I\left[Y_{0.} \geq 1\right]$. Note that Δ includes a counterfactual and is not observed for each individual. Three parameters of interests are: (i) the average treatment effect (ATE);

(ii) the average treatment effect on the treated (ATT) and on the untreated (ATU); (iii) the marginal effect of covariate X_k on the ATE. Parameters, ATT, ATT, and ATU, measure the average value of Δ on different conditioning sets. The ATE is the average effect of coresidence for a parent chosen at random from the population. The ATE given a value of X is defined as

$$\Delta^{ATE}(x) \equiv E(\Delta | X = x).$$

The ATT and ATU measure the average effect for a parent who is in coresidence and for a parent who is not in coresidence, respectively, and are defined as

$$\begin{aligned}
\Delta^{ATT}(x, z, D = 1) &\equiv E(\Delta | X = x, Z = z, D = 1) = E(\Delta | X = x, U_D \leq z\beta_D), \\
\Delta^{ATU}(x, z, D = 0) &\equiv E(\Delta | X = x, Z = z, D = 0) = E(\Delta | X = x, U_D > z\beta_D).
\end{aligned}$$

The ATT is the parameter most commonly estimated in literature. The marginal effect of observed characteristics on the ATE, (iii), tells us how the coresidence effect varies across observed characteristics. It is informative in inferring the mechanism underlying the causal effect.

To obtain the estimates of these (unconditional) treatment parameters, I integrate estimated treatment parameters against the empirical distribution of X and Z. Standard errors are computed using the delta method. Table 2 summarizes the definitions of these treatment parameters and provides expressions to compute them.

[Insert Table 2]

5 Data

The data is derived from the Nihon University Japanese Longitudinal Study of Aging (NUJLSOA), a nationally representative survey of the population aged 65 and over.[11] The analyses in this study use the latest two waves: the third wave conducted in 2003 and the fourth wave in 2006.[12] The population of interest consists of elderly parents: individuals aged 65 years and older who have at least one child in the base year, 2003.[13] The unit of observation consists of an elderly parent who has completed the third wave survey. In the NUJLSOA, if the subject is incapable of being interviewed due to illness, dementia, or another cause, the interview may be conducted with a proxy. The sample does not include the spouses of interviewees. I exclude elderly parents who are in hospitals or institutions in the base year, who live with persons other than their spouse and child's family, or who live with multiple children.[14] After dropping observations with data problems, such as missing values and inconsistent answers, the final sample consists of 3,023 elderly parents. Of these parents, 1,572 (52.0%) lived with a child in the base year.

[Insert Table 3]

Table 3 defines the variables used in this study. The dependent variable, $Y_i \equiv (1 - D_i)\, Y_{0i} + D_i Y_{1i}$, is a six-categorical health outcome variable that takes the value zero if an elderly parent dies within three years (between 2003 and 2006). If the parent survives, it takes a value between 1 and 5 according to the five self-assessed health levels. Table 4 shows the

[11] For the details of the NUJLSOA, see http://www.usc.edu/dept/gero/CBPH/nujlsoa/.

[12] The earlier waves have shorter time intervals between one another. Survey intervals that are too short are not suitable for analyzing effects on the mortality rate.

[13] The definition of a child in this study includes biological, step, and adopted children but not children-in-law.

[14] These exclusions are possibly endogenous, but the number of observations dropped is quite small.

distribution of the health outcome variable. The distribution centers around "3: Average" with a significant mass in the death category. The three year mortality rate is 10.0 percent in the sample, or 3.23 percent per annum.[15] Comparison of the distributions of the non-coresidence and coresidence groups shows that coresidence is associated with a lower survival rate. Such negative association can be observed if parents with worse health conditions are more likely to be in coresidence. The bottom half of Table 4 shows a moderate negative association between the baseline health and coresidence, indicating the importance of controlling for the baseline health.

[Insert Table 4]

Table 5 provides descriptive statistics for the other variables used in this study. Variables from *Spouse* to *Community activity* are individual level characteristics of each survey respondent. The other variables in Table 5 are defined for each elderly couple instead of each individual respondent, if the respondent is married. The variable, *ADL Disability*, is constructed as the first factor from a factor analysis of seven basic activities-of-daily-living (ADL) items based on the 3,023 parent observations. Its mean value and standard deviation are roughly zero and one, and it takes a larger value when a parent has severer disability.[16]

The last two variables, *Rural* and *Land price*, are the instruments used to address non-random selection into coresidence. The land price variable is constructed from a government source as an average price of residential properties sampled in each municipal area. The logic behind these instruments is that ruralness and average land prices are not likely to

[15]This number is consistent with the Japanese vital statistics: the national mortality rate of the elderly (65 and over) in 2005, inclusive of hospitalized and institutionalized individuals, was 3.46 percent per annum.

[16]The use of simple average instead of factor analysis makes no significant differences in the results.

directly affect the health outcome of elderly parents when other covariates are held constant, but they do affect the coresidence decision by changing the setup and maintenance costs of coresidence. As can be seen below, these instruments are indeed strongly correlated with coresidence; in cities with high land prices, intergenerational coresidence, which requires a larger house, appears quite expensive to set up and maintain.[17]

The exogeneity of instruments might be a concern. Those who live in urban areas and rural areas may be different, e.g., in terms of income, education, and life-style, and land prices may reflect how livable those areas are. These instruments are thus likely to be correlated with health outcomes unconditionally in the long run. However, the exogeneity condition necessary in this study is that when controlling for the base year health and other observable characteristics, the instruments are unrelated to the three-year health outcome. All else equal, it is unlikely that the location choice makes a significant difference to the three year mortality. To further support this argument, universal access to health care services is reasonably guaranteed in Japan and virtually every elderly individual is covered under the public health insurance system and the public long-term care insurance system, which require fairly limited out-of-pocket expenses for the elderly. Although the accessibility issue exists in some remote areas, such remote and less populated areas, e.g. remote tiny islands, are not included in the data. One may also argue that the land price might reflect the lifetime earnings and other unobserved attributes of the elderly who live in that area. However, this relationship is weak for this generation in Japan. The variance in land prices rose dramatically in the 1980s, which provided some people with huge windfall capital

[17] If cultural tendency for coresidence affects the real-estate market at the municipality level, the reverse causation raises a concern about the validity of the land price instrument. I believe this is not a major concern in Japan because as discussed below, the major part of the land price dispersion emerged in the last several decades of the 20th century, whereas social norms and traditions evolve much more slowly.

gains. Hence, although current land prices indicate the cost of starting and maintaining coresidence, they are much less related to the prices at which the current elderly individuals purchased their land.[18] In addition, I conduct several statistical tests and, as shown in the next section, the results support the appropriateness of my instruments.

[Insert Table 5]

6 Results

6.1 Linear Regressions and Instrument Validity

Before reporting results of the full structural model, I discuss results from simple linear models, which illuminate the nature of the data. Table 6 reports three simple estimates of the coresidence effect: (1) the unconditional average effect; (2) the average effect from a linear probability model (LPM); and (3) the average effect from an LPM with the instruments. The unconditional average effect shows that the three year mortality rate for parents in coresidence is 3.6 percentage points higher than for those who have a child but live independently. This difference is statistically significant at the 5 percent level. This estimate, however, might capture the systematic difference between the two groups. The LPM model in the next column utilizes baseline controls available in the data to control for the differences between the two groups. The IV-LPM model in the last column controls for selection on both observables and unobservables. In neither model, the estimate of the coresidence effect is statistically significant, suggesting the importance of baseline controls.

[18] Similarly, one may argue that correlation between time trend in health and land prices may exist even after controlling for baseline health and affect the validity of the instrument, but in Japan, this will not be a major concern for the same reason.

[Insert Table 6]

Before proceeding to the full model, I discuss the quality of my instruments. First, in the first stage regression of the IV-LPM model in Table 6, the instruments are jointly significant with t-statistics of 4.45 and -2.65, respectively. Hence, they are strong predictors of coresidence. Second, the overidentification test supports the validity of the IVs: the P-value of the Hansen J-statistics is 0.429.[19] The validity of the instruments can also be tested by a "placebo test" in which I regress the three-year survival rate on the instruments and the same covariates used in the main analysis using *childless elderly individuals*. If intergenerational coresidence is the only channel through which the land price and rural dummy potentially have an effect on health outcomes, those two instruments must have no significance in the same health-outcome regression for those who do not have children. Table 7 reports the results of this placebo test based on an LPM on three year survival and an ordered probit model for the six-categorical health outcome. The rural area dummy has no power to explain the two dependent variables and the effect of land price is fairly weak and statistically not significant, supporting the validity of the instruments. Lastly, in the above IV-LPM model, my instruments pass a standard weak identification test: the Kleibergen-Paap Wald F-statistic is 18.28, and the Stock-Yogo critical value of the weak IV test for 10% maximal LIML size is 8.68 (Stock and Yogo, 2005). These tests indicate the relevance, validity, and power of the instruments.[20]

[19] Since the dependent variable is binary, statistics robust to heteroskedasticity are employed.

[20] As another test of weak instruments, I study how my covariates vary along the instruments. If the covariates in the health outcome equation and the instruments move together, this raises concerns about weak instruments. I run a simple probit regression of coresidence on the instruments, generate a propensity score, and divide it into quartiles. At the 5% significance level, 9 of the 18 covariates have no difference across quartiles. 4 covariates show differences across quartiles, but they vary non-monotonically along the quartiles. The remaining 5 covariates either increase or decrease along the quartiles, but the correlation is fairly small: the largest correlation is -0.23 between the propensity score and years of education. These results suggest a reasonable identification power of the instruments.

[Insert Table 7]

6.2 Estimated Coefficients

The full model estimates the coresidence equation, (1), and the two health outcome equations, (2), simultaneously. Table 8 reports the coefficient estimates of the coresidence equation, and Table 9 reports the coefficient estimates of the two health outcome equations for the non-coresidence and coresidence states.

The results of the coresidence equation confirm the presence of non-random selection into coresidence. An elderly parent is more likely to be in coresidence with a child if the parent has characteristics that require family support, such as being widowed, having limitations in daily activities, and having dementia. Parents are also more likely to coreside with a child when they are working, less educated, and a parent of more children. Parents who own a house are less likely to be in coresidence. Interpretation of the coefficient estimates in the coresidence equation, however, requires caution because this equation does not address unobserved heterogeneity, and hence, captures only statistical association rather than causal effects. The results also confirm the significance of the instruments. Living in an urban area with a higher land price is associated with higher costs of coresidence, which reduces the probability of coresidence significantly. The significance of the instruments in the selection equation contrasts with their weak explanatory power in the "placebo" regression for childless elderly individuals (Table 7).

[Insert Table 8]

The results of the two health outcome equations are reported in Table 9. Most of the estimated coefficients have expected signs. Parents who have better health outcomes after three years tend to be younger parents who have better base period health status, a job, and involvement in community activities. The table also reports the estimates of four cutoff points for the five discrete subjective health outcomes (c_1, c_2, c_3, c_4). These thresholds are relative to the threshold for death that is normalized at 0. They are precisely estimated and have reasonable values.

The factor structure parameter, θ, enters the health outcome equations with the coefficient parameters, α_0 and α_1. Note that a model with no selection on unobservables implies $\alpha_0 = \alpha_1 = 0$. The estimates of α_0 and α_1 are shown in the first row in Table 9 as coefficient estimates on θ. Although neither α_0 nor α_1 is statistically significant, they are jointly significant: the model with no selection on unobservables (i.e. $\alpha_0 = \alpha_1 = 0$) is rejected by the likelihood ratio test at the 99% confidence level. Thus, the null hypothesis of no selection bias due to unobservables is rejected and the reported one-factor model is the preferred specification.

[Insert Table 9]

6.3 Treatment Parameters and Selection on Unobservables

The estimated treatment parameters are reported in Table 10. The ATE is estimated to be -0.0158, implying that if coresidence with a child is randomly assigned to elderly parents, the three-year mortality rate increases by 1.58 percentage points on average. The standard error of 0.018 indicates that although the ATE estimate is not statistically significant, the effect is not larger than $+2.0$ percent at the 95% confidence level, so intergenerational coresidence is very unlikely to have a substantial positive effect. The next three rows in Table 10 show distributional parameters of the ATE. I compute the distributional parameter, $\Pr(\Delta = 1|X) \equiv \Pr(Y_1 \geq 1, Y_0 = 0|x)$, as $\Pr(Y_1 \geq 1, Y_0 = 0|x) = \int_\theta \Phi(x\beta_1 + \alpha_1\theta)[1 - \Phi(x\beta_0 + \alpha_0\theta)]\phi(\theta)\,d\theta$. $\Pr(\Delta = 0|X)$ and $\Pr(\Delta = -1|X)$ are also defined in the same manner. The results reveal that if coresidence were randomly assigned to elderly parents, 89.1% would not be affected by coresidence, 4.7% would avoid death due to coresidence, and 6.2% would die because of coresidence.

This picture drastically changes when we look at the treatment effect on the treated. The ATT is estimated to be -0.0521 and is significant at the 1 percent level. This estimate implies that if parents in coresidence had not lived with children, their three year survival rate would have been 5.2 percentage points higher (or 1.7 percentage points per annum). Families with a potentially negative coresidence effect tend to self-select into coresidence. Reflecting the heterogeneity in the coresidence effect, the ATU is estimated to be 0.0363.

[Insert Table 10]

Heterogeneity in the coresidence effect can be attributed to observed and unobserved factors. To illustrate the working of the factor structure model, Table 11 reports the correlation estimates among observables, unobservables, and the treatment effect. The first three rows show the correlations among the three equations in terms of observables and unobservables. The two health outcome equations, (Y_0^*, Y_1^*), are strongly and positively correlated due to the strong positive correlation of observables, $(X\beta_0, X\beta_1)$, though the unobservable terms, (U_0, U_1), are negatively correlated. The health outcomes in the non-coresidence state and selection into coresidence show no significant correlation, whereas selection is negatively correlated with health outcomes in the coresidence state through both observables and unobservables. Taking these relationships altogether, the last three rows show the negative relationship between coresidence and the coresidence effect, which both observables and unobservables contribute to. Unobserved factors underlying the negative selection on unobservables might capture the dependence of children, parental altruism, social norms, or unobserved mental or physical health disposition. In my empirical framework, however, I am agnostic about the underlying factors of the negative selection on unobservables.

[Insert Table 11]

6.4 Marginal Effects on *ATE* and Testing Theory Predictions

Observable characteristics also generate heterogeneity in the coresidence effect, and the relationship between observables and the heterogeneity allows us to test the theory predictions and infer the mechanism underlying the causal effect. Table 12 reports the computed marginal effects of observed characteristics on the ATE. Parents who experience a worse coresidence effect tend to be older mothers with ADL limitations, dementia, and no involvement in household financial management. They are less likely to own a house and tend to have a younger child, and though less statistically significant, they tend to be widowed and less educated. Note that the marginal effects of observed characteristics can also be estimated for a model with no selection on unobservables (i.e. $\alpha_0 = \alpha_1 = 0$). The marginal effects estimated with this restricted model show largely consistent magnitudes and significance, implying that the discussion in this subsection does not depend on the choice of my instruments.[21]

[Insert Table 12]

[21] The results are available upon request.

Are these observations consistent with the theory predictions? The first prediction on the care burden is clearly supported by the significant negative relationship between daily activity limitations and the ATE. Although disability is generally highly related to the future health outcome, the negative relationship observed here is after controlling for other health characteristics. By contrast, the subjective health dummies and the cancer dummy show no strong impact on the ATE. These variables are more closely related to health and remaining life years. Therefore, what matters to the consequences of coresidence is the burden of caregiving to coresiding children, not health conditions.[22] The second prediction about the lifetime opportunity costs of the child is also supported by the positive coefficient on the child's age. Compared with those in their fifties and sixties, children in their thirties and forties have longer expected lifetimes and tend to have a greater need of time. Facing children's higher marginal utility of time, parents who have care needs may have to accept a shorter life.

Other findings also confirm the validity of the theory. If both parents are alive, they can support each other, whereas a widowed parent's reliance on her coresiding child may be substantial. Reliance on coresiding children may also become greater as parents age. The estimated coefficient on *Spouse* and *Age* confirm these predictions: older widowed parents tend to experience a worse coresidence effect. Parents may not have to compromise their health if they can compensate their children by another means — typically, by financial transfers. Although income and working status are not a strong proxy for the wealth of the elderly, the house asset constitutes a large portion of their wealth, and this is especially the

[22] Similarly to the ADL disability variable, the presence of dementia, another contributor to the coresidence burden, has a negative effect on the ATE. The interpretation of this finding, however, is not straightforward, because it is unclear how much control parents with dementia has over their health-related decisions.

case in Japan. The results in Table 12 confirm the role of such financial transfers. Those who own a house do not need to sacrifice their health. If education is a good proxy for lifetime income, the positive coefficient on education is also consistent with this argument. The variable *income missing* exhibits a significant negative estimate, suggesting that parents who did not answer their approximate income level in the survey may have very limited transferable wealth. Lastly, the number of children is negative and insignificant, suggesting that the data does not support the family bargaining hypothesis.

Summarizing these findings, it is reasonable to conclude that elderly parents who tend to experience a larger negative impact from coresidence are those who expect to rely significantly on the child and have limited potential to make financial compensation as a quid pro quo.

7 Discussion

7.1 Comparison of Results from Different Specifications

In Table 13, I compare the estimated ATEs on three-year survival produced by different specifications: (1) the unconditional average coresidence effect; (2) an LPM, in which I impose a common treatment effect assuming no selection on unobservables; (3) an IV-LPM that allows selection on unobservables; (4) a binary probit model that assumes heterogeneity in the treatment effect and no selection on unobservables; (5) a binary factor structure model that assumes heterogeneity in the treatment effect and selection on unobservables; (6) an ordered probit model that assumes heterogeneity in the treatment effect and no selection on unobservables; and (7) an ordered factor structure model that assumes heterogeneity in the treatment effect and selection on unobservables (the preferred full model). The results show that addressing non-random selection on unobservables leads to stronger negative treatment effects, indicating positive selection on unobservables. Adding the ordered categorical structure produces more precise estimates. Both facts support the use of the full model (7). The comparison also indicates that the coresidence effect is unlikely to be substantially positive.

[Insert Table 13]

I also conduct a simple goodness-of-fit test. Based on 5,000 Monte Carlo simulation draws of $(\theta, \varepsilon_D, \varepsilon_0, \varepsilon_1)$, I check to what extent the preferred model can correctly predict the coresidence status and health outcomes. The baseline specification is an independent combination of a binary probit for coresidence and a standard ordered probit for health outcomes. This baseline model correctly predicts 15.5% of the observed data, whereas the full model (7) correctly predicts 15.7%. Thus the factor structure model is able to reproduce the observed data fairly well.

7.2 Is the Health Transfer Hypothesis Plausible?

Before I conclude, I examine alternative explanations for the negative coresidence effect.

Abuse and Neglect Nishi et al. (2010) find that Japanese women cared for by their daughters-in-law experience the lowest survival propensity, and propose abuse and neglect as a potential explanation. Lachs et al. (1998) report that elderly Americans who experienced abuse were 3.1 times more likely to die during a 3-year period than those who were not mistreated, and in Japan, major perpetrators are daughters-in-law (Soeda and

Araki, 1999). Although abuse and neglect are not uncommon in Japan and thus appear to be responsible for a part of the negative coresidence effect, they are unlikely to be the primary driving force of the negative coresidence effect for the following reasons. First, if children have a means to shorten the remaining life of their parents directly or indirectly and commit themselves to actually do so, the expected return of such conduct is higher when the parent has larger wealth. However, the signs of the calculated marginal effects of wealth-related variables (Table 13) are not consistent with this theoretical prediction. Second, such a bad relationship cannot explain the formation of coresidence. It is natural to assume that families make the coresidence decisions based on the expected costs and benefits, and a good relationship should contribute to the benefit of coresidence. Certainly, families sometimes start coresidence based on biased expectations or because they simply do not have other options, and daughters-in-law, particularly under Japanese virilocal traditions, may be forced to be the caregiver irrespective of their preferences. However, during my sample period, Japanese families are as mobile as other modern societies and exhibit a wide range of living arrangements, so the living arrangements observed in my data are more likely to reflect the preferences of the families rather than traditions and social norms. Alternatively, we can regard abuse and neglect as a special case of the longevity trade-off, in the sense that coresiding parents have no strong intention to protect their health and maximize their remaining life years, and they passively accept the situation, probably without much bargaining power.

Exposure to Microbes and Viruses Coresidence may negatively affect parental health through a greater exposure to microbes and viruses. To check whether this is the major driving force of the negative coresidence effect, I estimate a model to which I add another covariate, the proportion of children in bad health among all the children the parent has as a proxy of the coresiding child's health.[23] This model yields an ATE and ATT almost identical to the preferred model, and shows no relationship between the ATE and the health status of children.[24] Another rebuttal to this hypothesis is the fact that the presence of a spouse has no significant effect on the respondent's health (Tables 5, 6, and 8), whereas the hypothesis implies a greater exposure to viruses of married individuals compared with individuals living alone. Furthermore, the marginal effect of the presence of a spouse on the ATE is positive and almost significant at the 10% level (Table 12), which also contradicts the hypothesis.[25]

Access to Care As shown in Table 5, parents living with children are poorer than parents living without children. Then one might think that elderly parents in coresidence are generally poorer and hence cannot afford to access health and nursing care, which leads to the spurious negative coresidence effect on health. This explanation is unlikely in Japan, however. The Japanese health care system guarantees universal access, and during the sample years of this study, the out-of-pocket payments for the elderly were quite small and the affordability of health and nursing care was not a major issue.[26] In addition, Japan introduced public long-term care insurance in April 2000, and the affordability issue of basic nursing assistance was not a major issue either.

[23] In the NUJLSOA data, a child is defined to be unhealthy if the child is described by the parent as "not healthy to take care of someone".

[24] The t-value of its marginal effect is -0.25. The results are available upon request.

[25] One might argue that the reason for the non-negative effect of a spouse is that the protective effect of a spouse dominates the effect of exposure to viruses. If that is the case, however, this exposure-to-viruses story is not a convincing explanation of the negative coresidence effect.

[26] For example, from September 1997 to December 2000, an outpatient visit cost elderly Japanese 500 Japanese yen (around 5 US dollars in early 2000), only up to the fourth visit, and from January 2001 to March 2008, elderly Japanese were charged a 10% copayment, only up to a certain maximum amount. Out-of-pocket payments have significantly increased since 2008.

Note that, even if there were a significant income gradient regarding access to care, the empirical framework used above takes care of such baseline differentials. The negative coresidence effect comes from comparison between those who live with children and those who live without children, *controlling for selection on observables and unobservables.* Income, wealth, and education are included as control variables.

Data Issues: Institutional Care and Interview Proxy An alternative to coresidence and informal care is institutional care, such as a nursing home. Formal care provided in such institutions comes with professional medical assistance, and hence, receiving institutional care may be associated with better health outcomes. This might induce a spurious negative coresidence effect. In the population of this study, however, the group of elderly parents living without children does not include those living in institutions.

Another possible explanation is that elderly parents who live without children may eventually move to a nursing home once their health deteriorates, which results in positive selection in this group compared to those who live with children. However, such selection is not likely to be the major driving force of the negative coresidence effect. In Japan, institutional care, such as long-term care in hospitals and group homes, has traditionally been for those who have no alternatives. Nursing homes as a satisfactory alternative to living with a child are still under development in Japan and tend to be very expensive. In addition, few elderly parents move to a nursing home in the NUJLSOA.

In the NUJLSOA, if the subject is incapable of being interviewed due to illness or another cause, the interview may be conducted with a proxy, typically a family member in the household. This is another potential source of bias because parents living alone with severe illness are less likely to be observed in the data due to the lack of a proxy than similarly ill parents living with a child. I can show that this possibility is not a major concern by estimating the coresidence effect separately for married and unmarried parents, because this selection bias due to the lack of a proxy is only relevant for parents living alone. Estimating the LPM and IV-LPM models in Table 6 separately for married and unmarried parents reveals that the coresidence effect of married couples are even slightly worse than that of unmarried couples, though the difference in the coresidence effect is not statistically significant.

How Elderly Japanese View Death Support for the health transfer hypothesis is found in many Japanese survey studies. A survey of over 2,000 Japanese conducted in 2004 reveals that around 40% of respondents did *not* hope to live long and one of their main concerns was to impose burdens of caregiving on families (Arai and Arai, 2008; Arai et al., 2005). As Long (2006) discusses, it is a widely accepted notion that an ideal death is a death that is peaceful, painless, and sudden so that it does not impose caregiving burdens on others. Another survey based on Japanese over 40 finds that more than 60% of respondents prefer such sudden death and discusses that one of the major reasons for such preferences is to avoid imposing care burdens on families (Kotani, 2004).

8 Conclusions

Although intergenerational coresidence is declining, it is still the most common living arrangement in the world (UN, 2005), and many countries expect to rely on informal care from family members in the foreseeable future. In less developed countries, around three quarters of elderly persons live with a child. In European countries the average is about 25%, but the trend toward solitary living has slowed or halted in many countries.

In this study, I find a statistically significant, negative coresidence effect on the treated. Estimated heterogeneity in the treatment effect suggests that vulnerable parents with high care needs and limited potential to make financial compensation are most likely to be in coresidence and to experience the largest negative impact. These findings are consistent with the hypothesis that coresidence may negatively affect parental health by creating disincentives for parents to invest in their health when there is no other means of compensation.

What policy implications can be derived from these results? One might argue that as long as reduced health is a consequence of rational choice, it is not a social concern. This argument shares the same logical structure with the suicide literature. If individuals commit suicide when the net utility value of their life becomes negative, suicide is their optimal choice, and hence there is no need for society to care. Others might argue that it is a social concern based on externality to others, equity, or social justice.

If the strong negative coresidence effect on the most vulnerable group of the elderly is a social concern, it is worth mentioning three policy implications from this study. First, policies that aim to directly affect living arrangements need to be carefully designed. The Singaporean government, for example, has introduced housing tax incentives to promote intergenerational coresidence and to maintain the traditional social structure (Mehta et al., 1995). My results raise a concern that such policies may result in an unintended adverse effect on parental health. Second, my results highlight the importance of policies that reduce the burden of caregiving to coresiding children and policies that support financial independence of the elderly so that they do not have to sacrifice their health. Third, for such policies to be effective, the intergenerational link needs to be taken into consideration because it may cause a version of the Ricardian equivalence. For example, when governments provide informal care support for altruistically linked families, especially those in coresidence, and finance its cost by taxing the younger generation, families may respond in such a way that their responses dilute the effect of the policy (the Ricardian neutralization).

References

[1] Aakvik, A., J.J. Heckman, and E.J. Vytlacil, "Estimating Treatment Effects for Discrete Outcomes when Responses to Treatment Vary: An Application to Norwegian Vocational Rehabilitation Programs," *Journal of Econometrics* 125 (2005), 15-51.

[2] Altonji, J.G., F. Hayashi, and L.J. Kotlikoff, "The Effects of Earnings and Wealth on Time and Money Transfers between Parents and Children," In *Sharing the Wealth: Demographic Change and Economic Transfers between Generations*, Edited by A. Mason and G. Tapinos. New York and Oxford: Oxford University Press, 2000, pp. 306-57.

[3] Arai, A. and Y. Arai, "Advance Care Planning among the General Public in Japan: Association with Awareness about Dementia," (in Japanese), *Nippon Ronen Igakkai Zasshi* 45 (2008), 640-6.

[4] Arai, Y., K. Kumamoto, S.H. Zarit, H. Dennoh, M. Kitamoto, "Angst in Shangri-La: Japanese Fear of Growing Old," *Journal of American Geriatrics Society* 53(9) (2005), 1641-2.

[5] Banks, J. and J.P. Smith, "International Comparisons in Health Economics: Evidence

from Aging Studies," *RAND Working Paper* WR-880 (2011).

[6] Barro, R.J., "Are Government Bonds Net Wealth?" *Journal of Political Economy* 82(6) (1974), 1095-1117.

[7] Becker, G.S., "A Theory of Social Interactions," *Journal of Political Economy* 82 (1974), 1063-93.

[8] Becker, G.S., *A Treatise on The Family*, (Harvard University Press, Cambridge, 1991).

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

[10] Chan, A., C. Malhotra, R. Malhotra, and T. Østbye, "Living Arrangements, Social Networks and Depressive Symptoms among Older Men and Women in Singapore," *International Journal of Geriatric Psychiatry* 26(6) (2011), 630-639.

[11] Chou, K.-L., A.H.Y. Ho, and I. Chi, "Living Alone and Depression in Chinese Older Adults," *Aging and Mental Health* 10(6) (2006), 583-591.

[12] Cox, D., "Motives for Private Income Transfers," *Journal of Political Economy* 95 (1987), 508-46.

[13] Davis, M.A., D.J. Moritz, J.M. Neuhaus, J.D. Barclay, and L. Gee, "Living Arrangements, Changes in Living Arrangements, and Survival among Community Dwelling Older Adults," *American Journal of Public Health* 87(3) (1997), 371-377.

[14] Do, Y. and C. Malhotra, "The Effect of Coresidence with an Adult Child on Depressive Symptoms among Older Widowed Women in South Korea: An Instrumental Variable Estimation," *Journal of Gerontology B: Psychological Sciences and Social Sciences* 67B(3) (2012), 384-391.

[15] Ehrlich, I. and H. Chuma, "A Model of the Demand for Longevity and the Value of Life Extension," *Journal of Political Economy* 98(4) (1990), 761-782.

[16] Ettner, S.L., "The Impact of "Parent Care" on Female Labor Supply Decisions," *Demography* 32(1) (1995), 63-80.

[17] Fujino, Y, and S. Matsuda, "Prospective Study of Living Arrangement by the Ability to Receive Informal Care and Survival among Japanese Elderly," *Preventive Medicine* 48 (2009), 79-85.

[18] Goody, J., "Inheritance," *The New Palgrave: A Dictionary of Economics.* First Edition. Eds. J. Eatwell, M. Milgate, and P. Newman. (Palgrave Macmillan, 1987).

[19] Grossman, M., "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy* 80(2) (1972), 223-255.

[20] Hall, R.E. and C.I. Jones, "The Value of Life and the Rise in Health Spending," *The Quarterly Journal of Economics*, 122(1) (2007), 39-72.

[21] Hamermesh, D.S. and N.M. Soss, "An Economic Theory of Suicide," *Journal of Political Economy* 82(1) (1974), 83-98.

[22] Heckman, J., "Statistical Models for Discrete Panel Data," in: Manski, C., McFadden, D. (Eds.), *Structural Analysis of Discrete Data with Econometric Applications.* (MIT Press, Cambridge, MA, 1981).

[23] Heckman, J., "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations," *Journal of Human Resources* 32 (1997), 441-462.

[24] Hughes, M.E., and L.J. Waite, "Health in Household Context: Living Arrangements

and Health in Late Middle Age," *Journal of Health and Social Behavior* 43 (2002), 1-21.

[25] Ikeda, A., H. Iso, I. Kawachi, K. Yamagishi, M. Inoue, and S. Tsugane, "Living Arrangement and Coronary Heart Disease: The JPHC Study," *Heart* 95 (2009), 577-583.

[26] Jeon, G.-S., S.-N. Jang, S.-J. Rhee, I. Kawachi, and S.-I. Cho, "Gender Differences in Correlates of Mental Health Among Elderly Koreans," *Journal of Gerontology: Social Sciences* 62B(5) (2007), S323-S329.

[27] Johar, M., and S. Maruyama, "Does Coresidence Improve an Elderly Parent's Health?" UNSW Australian School of Business Research Paper, No. 2011 ECON 08, (University of New South Wales, 2011).

[28] Johar, M., S. Maruyama, and S. Nakamura, "Transition to Prent-Child Coresidence: Parental Needs and the Strategic Bequest Motive," UNSW Australian School of Business Research Paper, No. 2010 ECON 05, (University of New South Wales, 2010).

[29] Kenny, P.M., J.P. Hall, S. Zapart, and P.R. Davis, "Informal Care and Home-Based Palliative Care: The Health-Related Quality of Life of Carers," *Journal of Pain and Symptom Management* 40(1) (2010), 35-48.

[30] Kharicha, K., S. Iliffe, D. Harari, C. Swift, G. Gillmann, and A. Stuck, "Health Risk Appraisal in Older People: Are Older People Living Alone an 'At-Risk' Group?" *British Journal of General Practice* 57(537) (2007), 271-276.

[31] Kotani, M., "Shi ni taisuru ishiki to shi no osore" (in Japanese), Life Design Report,

May, Dai-ichi Life Research Institute Inc (2004).

[32] Lachs, M.S., C.S. Williams, S. O'Brien, K.A. Pillemer, and M.E. Charlson, "The Mortality of Elder Mistreatment," *Journal of the American Medical Association* 280 (1998), 428-32.

[33] Li, L.W., J. Zhang, and J. Liang, "Health among the Oldest-Old in China: Which Living Arrangements Make a Difference?" *Social Science & Medicine* 68 (2009), 220-227.

[34] Long, S.O., *Final Days: Japanese Culture and Choice at the End of Life* (University of Hawaii Press, 2006). ISBN 978-0824829643.

[35] Lund, T., P. Due, J. Modvig, B.E. Holsetin, M.T. Damsgaard, and P.K. Andersen, "Cohabitation and Marital Status as Predictors of Mortality — An Eight Year Follow-Up Study," *Social Science & Medicine* 55 (2002), 673-679.

[36] Mehta, K., M.M. Osman, and A. Lee E.Y., "Living Arrangements of the Elderly in Singapore: Cultural Norms in Transition," *Journal of Cross-Cultural Gerontology* 10 (1995), 113-143.

[37] Michael, Y.L., L.F. Berkman, G.A. Colditz, and I. Kawachi, "Living Arrangements, Social Integration, and Change in Functional Health Status," *American Journal of Epidemiology* 153(2) (2001), 123-131.

[38] Ministry of Health, Labour and Welfare, *Report on the Comprehensive Survey of Living Conditions of the People on Health and Welfare* (In Japanese) (2008). Available at: http://www.mhlw.go.jp/toukei/list/20-19-1.html.

[39] Nishi, A., N. Tamiya, M. Kashiwagi, H. Takahashi, M. Sato, and I. Kawachi, "Moth-

ers and Daughters-In-Law: A Prospective Study of Informal Care-Giving Arrangements and Survival in Japan," *BMC Geriatrics* 10:61 (2010).

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

[41] Rogers, R.G., R.A. Hummer, and C.B. Nam, *Living and Dying in the USA; Behavioral, Health and Social Differentials of Adult Mortality.* (San Diego, CA: Academic Press, 2000).

[42] Russell, D. and J. Taylor, "Living Alone and Depressive Symptoms: The Influence of Gender, Physical Disability, and Social Support Among Hispanic and Non-Hispanic Older Adults," *Journal of Gerontology: Social Sciences* 64B(1) (2009), 95-104.

[43] Sarwari, A.R., L. Fredman, P. Langenberg, J. Magaziner, "Prospective Study on the Relation between Living Arrangement and Change in Functional Health Status of Elderly Women," *American Journal of Epidemiology* 147(4) (1998), 370-378.

[44] Silverstein, M., Z. Cong, and S. Li, "Intergenerational Transfers and Living Arrangements of Older People in Rural China: Consequences for Psychological Well-Being," *Journal of Gerontology: Social Sciences* 61B(5) (2006), S256-S266.

[45] Soeda, A. and C. Araki, "Elder Abuse by Daughters-in-Law in Japan," *Journal of Elder Abuse and Neglect* 11 (1999), 47-58.

[46] Stock, J. H. and M. Yogo, "Testing for weak instruments in linear IV regression." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. D. W. K. Andrews and J. H. Stock, 80–108 (Cambridge: Cambridge University Press, 2005).

[47] United Nations *Living Arrangements of Older Persons around the World*. Department of Economic and Social Affairs, Population Division, United Nations, 2005.

[48] Walter-Ginzburg, A., T. Blumstein, A. Chetrit, and B. Modan, "Social Factors and Mortality in the Old-Old in Israel: The CALAS Study," *Journal of Gerontology: Social Sciences* 57B(5) (2002), S308-318.

[49] Wang, D., J. Zheng, M. Kurosawa, Y. Inaba, and N. Kato, "Changes in Activities of Daily Living (ADL) among Elderly Chinese by Marital Status, Living Arrangement, and Availability of Healthcare over a 3-Year Period," *Environmental Health and Preventive Medicine* 14 (2009), 128-141.

Table 1: Studies of the Effect of Coresidence on Health Outcomes

	country	population	approach	treatment[a]	outcome	finding[b]
Do and Malhotra 2012	South Korea	widowed women ≥ 65	panel/IV	child	depression	+
Chan et al. 2011	Singapore	elderly ≥ 60	cross-section	child	depression	mixed
Johar and Maruyama 2011	Indonesia	elderly ≥ 60	panel/IV	child	health/survival	−
Nishi et al. 2010	Japan	elderly ≥ 65	duration	daughters	survival	−
Fujino and Matsuda 2009	Japan	elderly ≥ 60	panel	coresidence	survival	0 or −
Ikeda et al. 2009	Japan	age $40 - 69$	panel	child	cardio/survival	−
Li, Zhang, Liang 2009	China	elderly ≥ 77	panel	child	ADL/health/survival	mixed
Russell and Taylor 2009	USA	elderly ≥ 60	cross-section	coresidence	depression	mixed
Wang et al. 2009	China	elderly ≥ 65	panel	coresidence	ADL/survival	mixed
Jeon et al. 2007	South Korea	elderly ≥ 65	cross-section	coresidence	depression	mixed
Kharicha et al. 2007	UK	non-disabled ≥ 65	cross-section	coresidence	various health	+
Chou, Ho, Chi 2006	Hong Kong	elderly ≥ 60	cross-section	coresidence	depression	mixed
Silverstein, Cong, Li 2006	Rural China	parents ≥ 60	cross-section	3-generation	mental	+
Hughes and Waite 2002	USA	age $51 - 61$	panel	child	ADL/health/mental	0 or −
Lund et al. 2002	Denmark	age 50, 60, 70	duration	spouse	survival	+
Walter-Ginzburg et al. 2002	Israel	elderly ≥ 75	panel	child	survival	−
Michael et al. 2001	USA	women age 60-72	panel	spouse	ADL/mental/survival	−
Rogers et al. 2000	USA	adults	cross-section	child	survival	−
Sarwari et al. 1998	USA	women ≥ 65	panel	coresidence	functional health	mixed
Davis et al. 1997	USA	elderly ≥ 70	panel	coresidence	health/survival	mixed

Note: ADL denotes "activity of daily living".
[a] Many studies compare more than two living arrangements, in which case I report a treatment most relevant to this study. "coresidence" in this column means living with somebody.
[b] Positive indicates a favorable impact on health, regardless of how a health measure is defined.

Table 2: Treatment Parameters of Interest

Treatment Parameter	Definition	Estimation
ATE	$\Delta^{ATE}(x) \equiv E(\Delta \mid X = x)$	$E_X\left[\Delta^{ATE}(x)\right] = E_X\left[\Pr(Y_1 \geq 1 \mid X = x) - \Pr(Y_0 \geq 1 \mid X = x)\right]$
(average treatment effect)		$= \int_x \int_\theta \left[\Phi(x\beta_1 + \alpha_1\theta) - \Phi(x\beta_0 + \alpha_0\theta)\right] \phi(\theta)\, d\theta dF_X(x)$
ATT	$\Delta^{ATT}(x, z, D = 1)$	$E_{X,Z}\left[\Delta^{ATT}(x, z, D = 1)\right]$
(average treatment	$\equiv E(\Delta \mid X = x, Z = z, D = 1)$	$= E_X\left[\Pr(Y_1 \geq 1 \mid X = x, Z = z, D = 1) - \Pr(Y_0 \geq 1 \mid X = x, Z = z, D = 1)\right]$
effect on the treated)	$= E(\Delta \mid X = x, U_D \leq z\beta_D)$	$= \int_{(x,z)} \int_\theta \left[\Phi(x\beta_1 + \alpha_1\theta) - \Phi(x\beta_0 + \alpha_0\theta)\right] \frac{\Phi(z\beta_D + \theta)}{\Phi\left(z\beta_D/\sqrt{2}\right)} \phi(\theta)\, d\theta dF_{X,Z}(x, z)$
ATU	$\Delta^{ATU}(x, z, D = 0)$	$E_{X,Z}\left[\Delta^{ATU}(x, z, D = 0)\right]$
(average treatment	$\equiv E(\Delta \mid X = x, Z = z, D = 0)$	$= E_X\left[\Pr(Y_1 \geq 1 \mid X = x, Z = z, D = 0) - \Pr(Y_0 \geq 1 \mid X = x, Z = z, D = 0)\right]$
effect on the untreated)	$= E(\Delta \mid X = x, U_D > z\beta_D)$	$= \int_{(x,z)} \int_\theta \left[\Phi(x\beta_1 + \alpha_1\theta) - \Phi(x\beta_0 + \alpha_0\theta)\right] \frac{(1-\Phi(z\beta_D + \theta))}{1-\Phi\left(z\beta_D/\sqrt{2}\right)} \phi(\theta)\, d\theta dF_{X,Z}(x, z)$
Marginal effect of kth	$\frac{\partial \Delta^{ATE}(x)}{\partial X_k}$	$E_X\left(\frac{\partial E(\Delta \mid X = x)}{\partial X_k}\right) = \int_x \frac{\partial \Delta^{ATE}(x)}{\partial X_k} dF_X(x)$
element of X on ATE		$= \int_x \int_\theta \left[\phi(x\beta_1 + \alpha_1\theta)\,\beta_{1k} - \phi(x\beta_0 + \alpha_0\theta)\,\beta_{0k}\right] \phi(\theta)\, d\theta dF_X(x)$

Note: $F_X(x)$ is the distribution of X. $F_{X,Z}(x, z)$ is the joint distribution of X and Z.

Table 3: Definitions of Variables

Dependent variables:	
Y_{0i}, Y_{1i}	=0 if the elderly parent dies within three years; 1 for very unhealthy; 2 unhealthy; 3 average; 4 healthy; and 5 very healthy
Coresidence	=1 if the elderly parent lives with a child; 0 otherwise
Explanatory variables:	
Spouse	=1 if living with a spouse; 0 otherwise
Age	Age of the elderly parent
Female	=1 if female; 0 otherwise
Health: Very good	=1 if the answer to the subjective health question is "very good".
Health: Good	=1 if the answer to the subjective health question is "good".
Health: Average	=1 if the answer to the subjective health question is "average" (reference group).
Health: Poor	=1 if the answer to the subjective health question is "poor".
Health: Very poor	=1 if the answer to the subjective health question is "very poor".
ADL Disability	The first factor in factor analysis applied for 7 basic activities-of-daily-living (ADL) items based on the 3,023 observations (the larger the severer disability)
Dementia	=1 if dementia is an existing condition; 0 otherwise
Cancer	=1 if cancer is an existing condition; 0 otherwise
Work	=1 if working; 0 otherwise
Education	Years of schooling (if missing, imputed from educational attainment information)
Community activity	The number of social/community groups/activities (top-coded at 3)
Income	The parent's own and spousal income, in millions of Japanese yen[a,b]
Income missing	=1 if income data is missing; 0 otherwise
Own house	=1 if living in a house self-owned or owned by spouse; 0 otherwise
Number of children	The number of living children
Age youngest child	Age of the youngest child
Instrumental variables:	
Rural	=1 if living in a rural area; 0 otherwise
Land price	Municipal level average land price (in 10,000 Jp yen / m^2, residential area)[b,c]

[a] Imputed from 13 interval categories in the questionnaire. [b] 1,000 JPY = approx. 9 US dollars in 2003. [c] As of January 1, 2003. Source: Public notice of land prices, Ministry of Land, Infrastructure, Transport and Tourism.

Table 4: Distribution of Health Outcomes after Three Years and Baseline Health

	All parents		Non-Coresidence		Coresidence	
Total	3,023	(100%)	1,451	(100%)	1,572	(100%)
Y_i: health outcomes						
5: Very healthy	322	(10.7%)	159	(11.0%)	163	(10.4%)
4: Healthy	474	(15.7%)	220	(15.2%)	254	(16.2%)
3: Average	1,090	(36.1%)	546	(37.6%)	544	(34.6%)
2: Poor	672	(22.2%)	327	(22.5%)	345	(22.0%)
1: Very poor	163	(5.4%)	81	(5.6%)	82	(5.2%)
0: Dead	302	(10.0%)	118	(8.1%)	184	(11.7%)
Baseline health						
Very healthy	429	(14.2%)	226	(15.6%)	203	(12.9%)
Healthy	545	(18.0%)	257	(17.7%)	288	(18.3%)
Average	1,251	(41.4%)	601	(41.4%)	650	(41.4%)
Poor	675	(22.3%)	308	(21.2%)	367	(23.4%)
Very poor	123	(4.1%)	59	(4.1%)	64	(4.1%)

Table 5: Means and Standard Deviations of Variables

	3,023 elderly parents		1,451 parents living without children		1,572 parents living with a child	
	Means	Std. dev.	Means	Std. dev.	Means	Std. dev.
Spouse	0.64	0.48	0.75	0.43	0.55	0.50
Age	75.36	7.03	74.38	6.58	76.27	7.32
Female	0.54	0.50	0.49	0.50	0.59	0.49
Health: Very good	0.14	0.35	0.16	0.36	0.13	0.34
Health: Good	0.18	0.38	0.18	0.38	0.18	0.39
Health: Poor	0.22	0.42	0.21	0.41	0.23	0.42
Health: Very poor	0.04	0.20	0.04	0.20	0.04	0.20
ADL Disability	0.00	0.98	−0.11	0.59	0.10	1.22
Dementia	0.04	0.20	0.02	0.13	0.07	0.25
Cancer	0.03	0.17	0.03	0.18	0.03	0.17
Work	0.23	0.42	0.23	0.42	0.22	0.42
Education	9.53	2.60	9.96	2.71	9.15	2.43
Community activity	1.12	1.06	1.16	1.07	1.09	1.05
Income	2.57	2.02	2.88	2.08	2.27	1.92
Income missing	0.18	0.39	0.16	0.37	0.20	0.40
Own house	0.72	0.45	0.82	0.39	0.64	0.48
Number of children	2.46	1.06	2.30	0.98	2.61	1.12
Age youngest child	43.93	7.64	43.05	7.04	44.74	8.07
Rural	0.38	0.48	0.31	0.46	0.44	0.50
Land price	10.67	11.52	11.87	12.80	9.56	10.07

Table 6: Linear Regressions of 3 Year Survival

Dependent variable:	Mean difference			LPM			IV-LPM[a]		
3 year survival	Coefficient		t-value	Coefficient		t-value	Coefficient		t-value
Coresidence	−0.036	***	−3.30	−0.002		−0.17	−0.037		−0.40
Spouse				0.002		0.13	−0.003		−0.17
Age				−0.010	***	−6.41	−0.010	***	−6.33
Female				0.038	***	3.45	0.038	***	3.45
Health: Very good				0.005		0.42	0.005		0.41
Health: Good				−0.000		−0.01	0.001		0.07
Health: Poor				−0.031	**	−2.15	−0.032	**	−2.15
Health: Very poor				−0.094	**	−2.14	−0.099	**	−2.13
ADL Disability				−0.058	***	−5.55	−0.057	***	−5.21
Dementia				−0.106	**	−2.27	−0.101	**	−2.04
Cancer				−0.048		−1.33	−0.048		−1.33
Work				0.008		0.80	0.010		0.84
Education				0.002		0.96	0.001		0.65
Community activity				0.013	***	2.67	0.013	***	2.69
Income				−0.000		−0.16	−0.001		−0.31
Income missing				−0.017		−1.22	−0.015		−1.06
Own house				−0.013		−0.97	−0.017		−0.98
Number of children				−0.004		−0.63	−0.002		−0.33
Age youngest child				0.000		0.37	0.000		0.30
Constant	0.919	***	127.99	1.597	***	18.52	1.622	***	15.21

Note: $N = 3,023$. LPM stands for linear probability model. White's robust standard errors are used.
[a] Two instruments, "rural" and "landprice", are applied to Coresidence. In the first stage regression, these are jointly significant, with t-statistics of 4.45 and −2.65 respectively.

Table 7: Instrument Validity Test Using Childless Parents

Dependent variable:	3 year survival (LPM)			health outcome (ordered probit)		
	Coefficient		t-value	Coefficient		t-value
Spouse	0.004		0.09	0.023		0.17
Age	−0.008	***	−2.95	−0.024	***	−2.59
Female	0.068		1.64	0.080		0.57
Health: Very good	−0.016		−0.29	0.632	***	3.31
Health: Good	0.031		0.60	0.515	***	2.99
Health: Poor	−0.072	*	−1.72	−0.569	***	−4.04
Health: Very poor	−0.233	***	−2.98	−0.930	***	−3.41
ADL Disability	−0.058	**	−2.29	−0.247	***	−2.64
Dementia	−0.181	**	−2.09	−0.549	*	−1.85
Cancer	−0.028		−0.28	−0.454		−1.35
Work	0.019		0.41	0.146		0.91
Education	0.001		0.21	0.023		1.16
Community activity	−0.009		−0.51	0.079		1.36
Income	−0.011		−0.92	−0.031		−0.78
Income missing	−0.058		−1.24	−0.034		−0.22
Own house	0.030		0.79	0.021		0.16
Rural	−0.001		−0.02	0.048		0.37
Land price	0.001		1.11	0.003		1.04
Constant	1.464	***	6.42			

Note: $N = 379$. LPM stands for linear probability model. For the ordered probit model, five cutoff points are estimated as −2.88, −2.58, −1.69, −0.44, and 0.09.

Table 8: Coresidence Equation

	Coefficient		t-value	Marginal Effects[a]
Spouse	−0.524	***	−6.02	−0.135
Age	0.011		1.19	0.003
Female	0.039		0.51	0.010
Health: Very good	−0.008		−0.08	−0.002
Health: Good	0.104		1.10	0.027
Health: Poor	−0.047		−0.52	−0.012
Health: Very poor	−0.648	***	−3.24	−0.167
ADL Disability	0.160	***	3.06	0.041
Dementia	0.756	***	3.33	0.195
Cancer	0.046		0.23	0.012
Work	0.165	**	1.97	0.042
Education	−0.034	**	−2.28	−0.009
Community activity	−0.000		−0.01	−0.000
Income	−0.026		−1.36	−0.007
Income missing	0.268	***	3.02	0.069
Own house	−0.548	***	−6.36	−0.141
Number of children	0.118	***	3.37	0.030
Age youngest child	−0.012		−1.55	−0.003
Rural	0.241	***	3.47	0.062
Land price	−0.008	***	−2.84	−0.002
Constant	0.528		0.98	0.136

[a] Defined as the analytical derivative averaged over the entire sample: $(1/N)\sum_{i=1}^{N}\left[\int\left(\partial \Pr\left(D=1|Z=z,\theta\right)/\partial z_k\right)\phi\left(\theta\right)d\theta\right]$.

Table 9: Health Outcome Equations

	Non-coresidence outcome: Y_0				Coresidence outcome: Y_1			
	Coefficient		t-value	M.E.[a]	Coefficient		t-value	M.E.[a]
Factor (θ)	1.844		1.36	0.075	−1.982		−1.42	−0.086
Spouse	−0.227		−0.98	−0.009	0.170		0.84	0.007
Age	−0.032		−1.44	−0.001	−0.084	*	−1.76	−0.004
Female	0.469	*	1.71	0.019	0.040		0.29	0.002
Health: Very good	1.583	*	1.88	0.065	1.402	*	1.88	0.061
Health: Good	0.801	*	1.74	0.033	0.637	*	1.80	0.028
Health: Poor	−0.984	*	−1.87	−0.040	−1.007	*	−1.86	−0.044
Health: Very poor	−2.110	*	−1.71	−0.086	−1.163	*	−1.82	−0.050
ADL Disability	0.006		0.07	0.000	−0.383	*	−1.69	−0.017
Dementia	0.038		0.08	0.002	−0.940		−1.51	−0.041
Cancer	−0.287		−0.79	−0.012	−0.285		−0.79	−0.012
Work	0.523		1.59	0.021	0.256		1.40	0.011
Education	−0.001		−0.05	−0.000	0.057		1.30	0.002
Community activity	0.199	*	1.68	0.008	0.126		1.50	0.005
Income	0.035		0.98	0.001	0.015		0.39	0.001
Income missing	0.152		0.79	0.006	−0.485		−1.60	−0.021
Own house	−0.319		−1.17	−0.013	0.172		0.88	0.007
Number of children	0.025		0.38	0.001	−0.077		−0.97	−0.003
Age youngest child	−0.020		−1.16	−0.001	0.028		1.32	0.001
Constant	6.843	*	1.72	0.280	8.452	*	1.82	0.366
Cutoff parameters:								
c_1	0.631	*	1.86					
c_2	2.295	*	1.89					
c_3	4.430	*	1.91					
c_4	5.762	*	1.93					

[a] M.E. denotes marginal effects, defined as the analytical derivative averaged over the entire sample. For $Y_j, j = 0, 1$: $(1/N)\sum_{i=1}^{N}\left[\int\left(\partial \Pr\left(Y_j \geq 1|X = x, \theta\right)/\partial x_k\right)\phi\left(\theta\right)d\theta\right]$.

Table 10: Mean and Distributional Treatment Parameters on 3-Year Survival

ATE (average treatment effect): $(1/N)\sum_{i=1}^{N}\left[\int \Delta^{ATE}\left(x,\theta\right)\phi\left(\theta\right)d\theta\right]$	−0.0158
standard error[a]	0.0180
Distributional version of ATE	
$\Pr\left(\Delta = 1\right)$	0.047
$\Pr\left(\Delta = 0\right)$	0.891
$\Pr\left(\Delta = -1\right)$	0.062
ATT (average treatment effect on the treated): $(1/N)\sum_{i=1}^{N}\left[\int \Delta^{ATT}\left(D = 1, x, z, \theta\right)\phi\left(\theta\right)d\theta\right]$	−0.0521
standard error[a]	0.0198
ATU (average treatment effect on the untreated): $(1/N)\sum_{i=1}^{N}\left[\int \Delta^{ATU}\left(D = 0, x, z, \theta\right)\phi\left(\theta\right)d\theta\right]$	0.0363
standard error[a]	0.0095

[a] Standard errors are computed using the delta method for given (x, θ) and averaged over the sample.

Table 11: Correlations between Observables, Unobservables, and Treatment Effect

$Corr\left(X\beta_0, X\beta_1\right) = 0.823$	$Corr\left(Z\beta_D, X\beta_0\right) = -0.035$	$Corr\left(Z\beta_D, X\beta_1\right) = -0.527$
$Corr\left(U_0, U_1\right) = -0.785$	$Corr\left(U_D, U_0\right) = 0.622$	$Corr\left(U_D, U_1\right) = -0.631$
$Corr\left(Y_0^*, Y_1^*\right) = 0.821$	$Corr\left(D^*, Y_0^*\right) = -0.034$	$Corr\left(D^*, Y_1^*\right) = -0.527$
$Corr\left(Z\beta_D, X\beta_1 - X\beta_0\right) = -0.877$		
$Corr\left(U_D, U_1 - U_0\right) = -0.663$		
$Corr\left(D^*, Y_1^* - Y_0^*\right) = -0.875$		

Table 12: Marginal Effects of Covariates on Average Treatment Effect

	Marginal effects on ATE [a]		t-value
Spouse	0.0167		1.61
Age	−0.0024	**	−2.08
Female	−0.0174	*	−1.80
Health: Very good	−0.0040		−0.20
Health: Good	−0.0052		−0.40
Health: Poor	−0.0034		−0.22
Health: Very poor	0.0360		1.21
ADL Disability	−0.0168	***	−3.06
Dementia	−0.0423	*	−1.74
Cancer	−0.0006		−0.28
Work	−0.0103		−0.99
Education	0.0025		1.46
Community activity	−0.0027		−0.68
Income	−0.0008		−0.33
Income missing	−0.0273	***	−2.68
Own house	0.0205	**	2.04
Number of children	−0.0044		−1.13
Age youngest child	0.0020	**	2.27

[a] The marginal effects, $E_{X,\theta}\left[\left(\partial\Delta^{ATE}\left(X=x,\theta\right)/\partial x_k\right)\right]$, are obtained by the analytical derivative averaged over the sample: $(1/N)\sum_{i=1}^{N}\left[\int\left(\partial\Delta^{ATE}\left(X=x,\theta\right)/\partial x_k\right)\phi\left(\theta\right)d\theta\right]$.

Table 13: Average Treatment Parameters on 3-Year Survival from Different Methods

	ATE	Standard errors
(1) Unconditional mean differences[a]	−0.0357	0.0108
(2) Linear probability model with common treatment effect[b]	−0.0017	0.0103
(3) Linear IV common treatment effect[c]	−0.0372	0.0934
(4) Binary model without selection on unobservables[d,e]	−0.0009	0.0463
(5) Binary model with selection on unobservables[d]	−0.0241	0.0952
(6) Ordered model without selection on unobservables[e]	0.0063	0.0268
(7) Ordered model with selection on unobservables	−0.0158	0.0180

[a] $\widehat{E}\left(Y_1|D=1\right)-\widehat{E}\left(Y_0|D=0\right)$

[b] Based on an LPM form and a common treatment assumption $\left(Y=X\beta+\gamma D+U\right)$.

[c] Linear IV estimation for the above LPM using the instruments.

[d] The dependent variable $Y=1$ if a parent survives and 0 otherwise.

[e] No selection on unobservables imposed, i.e., $\alpha_0=\alpha_1=0$.