Can family caregiving substitute for nursing home care?

Kerwin Kofi Charles a, Purvi Sevak b,∗

a University of Michigan and NBER, USA
b Department of Economics, Hunter College, 695 Park Avenue, New York, NY 10021, USA

Received 25 August 2004; received in revised form 19 April 2005; accepted 3 May 2005
Available online 24 June 2005

Abstract

Informal care should be a substitute for nursing homes but empirical evidence often suggests the opposite. This may be because informal care receipt is positively correlated with unobserved negative health characteristics. We exploit variation in children’s characteristics as instruments for informal care to provide Two-Stage Least Squares (TSLS) estimates of nursing home use among a sample of 6855 individuals from the 1993–2000 waves of the AHEAD survey. While OLS results suggest informal care is associated with greater future nursing home risk, TSLS estimates show that receipt of informal care statistically and substantially reduces the risk of nursing home entry. This finding has implications for Medicaid and private long-term care insurance markets.

© 2005 Elsevier B.V. All rights reserved.

JEL classification: J14; I12

Keywords: Long-term care; Nursing homes; Informal care

1. Introduction

Nursing homes are a costly and generally undesirable source of long-term care for the disabled elderly. As a result, they are often considered an option of last resort by
individuals, families, and the state. Despite various attempts by researchers to understand the
determinants of nursing home use, many aspects of the transition are not well understood.
For example, while health status is known to be a very strong determinant of nursing
home entry, many seriously disabled individuals reside in non-institutionalized settings,
while many less seriously disabled persons are institutionalized. This paper presents new
evidence on the relationship between informal home care receipt and nursing home entry
risk by following respondents in the Study of Asset and Health Dynamics among the Oldest
Old (AHEAD) from 1993 to 2000.

Informal care provided in the home by family or friends should be a substitute for nursing
homes but the available evidence in support of this proposition is quite weak; indeed there
is some evidence that the receipt of informal care and nursing home entry are positively
related. One obvious explanation for this empirical finding is that receipt of home care is
positively correlated with unobserved negative health characteristics that increase the prob-
ability of nursing home entry. The use of lagged measures of home care may reduce but not
eliminate this endogeneity problem, as health and disability are likely highly correlated over
time.

To address this problem, we estimate instrumental variables models of the effect of
informal care-giving on nursing home risk. Given a large literature documenting the
higher propensity of daughters to provide informal care relative to sons (McGarry, 1999;
Wolf et al., 1997), we use a variety of family structure characteristics to instrument for
receipt of unpaid help with ADLs in 1993. We argue that conditional on marital status
and the number of children, the gender composition and other characteristics of those
offspring are, for all practical purposes, exogenous, and are thus valid instruments for
receipt of informal care. We find that OLS estimates show a positive association between
nursing home entry and the receipt of informal care. However, when we account for
the endogeneity of the receipt of informal care using instrumental variables, we find
that nursing home entry and informal care are negatively related – that is, they are
substitutes.

Empirical evidence that informal care can reduce nursing home use has important
policy implications. Evaluations of demonstration projects like the Channeling Demon-
stration in the past few decades concluded that investments in home and community based
long-term care could not reduce total long-term care expenses by reducing nursing home
care. Nevertheless, state governments are continually looking for ways to reduce Medicaid
expenditures. Given that most disabled elderly people would prefer to receive care in the
community rather than be institutionalized, our finding suggests that an effective and possibly
efficient way to reduce nursing home expenses may be to subsidize informal caregivers.
This is particularly important as the number of disabled elderly continues to increase, and
Medicaid places a larger strain on state budgets.

The remainder of the paper is organized as follows. In the next section we summarize
some of the previous literature on this question. Section 3 provides a description of the
AHEAD data we use in our analysis. Section 4 presents OLS estimates of the associational
relationship between informal care and nursing home use. In Section 5 we use instrumen-
tal variables strategies to account for the endogeneity of informal care. Two-Stage Least
Squares (TSLS) and bivariate probit results are presented in this section. Section 6 con-
cludes.
2. Previous literature on informal care and nursing home risk

Many papers have examined the risk of nursing home entry and nursing home utilization rates using both national longitudinal data sets and data from smaller demonstrations and experiments. Though it is beyond the scope of this paper to review all of this work, we review papers that have specifically looked at the effect of home-based care on nursing home use.1

The literature has found conflicting and inconclusive evidence on the causal effect of home based care receipt on nursing home risk. A number of studies use the presence of potential caregivers – a spouse, child or sibling – as a proxy for informal care receipt and find mixed results. Freedman (1996) finds that being married reduces the hazard of nursing home entry by 41%, having a living daughter reduces it by 27% and having a sibling reduces it by 21%.

Newman and Struyk (1990) look at caregiving directly and find that informal support by a spouse is associated with a somewhat lower risk of nursing home entry and that formal support by others is associated with a higher risk of nursing home entry. However, when they interact these measures of home based care with the presence of disability, they find no effect of home based care on nursing home entry risk. Boaz and Muller (1994) find that family support significantly reduces the risk of long stays, but a negligible effect on the risk of short nursing home stays in the 1982–1984 National Long-Term Care Survey (NLTCS), a survey of the disabled elderly. Using the same data, however, Ettner (1994) finds that the potential supply of informal caregivers does not affect future nursing home use. In contrast, she does find that recipients of Medicaid subsidies for paid home care were less likely to use a nursing home.

Some of the difficulties of estimating the effects of home based care on nursing home entry should have been alleviated by data from The National Long-Term Care Channeling Demonstration Project, a major demonstration project funded by the Department of Health and Human Services in the 1980s. The project tested the feasibility and cost-effectiveness of alternative home and community based long-term care as alternatives to nursing home care among a sample of very frail older individuals. Despite the use of an experimental design, the “treatment” of intense care in the community did not result in lower rates of nursing home entry among the treatment group (DHHS, 1987).

One of the shortcomings of the literature is that, until recently, researchers have not had longitudinal data on a nationally representative of individuals, regardless of disability levels. In the Channeling demonstration, for example, the sample was one that was already quite frail. While these were individuals who are at highest risk of nursing home entry, they were already a highly selected sample in that they had managed to remain in the community until then. The same concern applies to the NLTCS data used by many studies, which is a sample of community-dwelling individuals disabled at baseline. In this paper, we use data from the Asset and Health Dynamics of the Oldest Old Study (AHEAD), described in detail in the next section. Unlike many earlier studies, it allows us to follow nursing home transitions, among a sample that was nationally representative at baseline.

3. Data

The Asset and Health Dynamics of the Oldest Old Study (AHEAD) first collected data on 8222 community dwelling older Americans in 1993. The goal of the survey was to collect nationally representative data on persons born in 1923 or older, and their spouses. Thus, while spouses of married respondents were interviewed, irrespective of their age, the nationally representative portion of the sample consists only of “age eligible” respondents—the 7443 non-institutionalized persons who were born before 1923. The study re-interviewed respondents in 1995, 1998 and every 2 years thereafter, until the death of the respondent. The study began as a representative sample of non-institutionalized older Americans, but since respondents are tracked over time as they age, whether they enter nursing homes or not, the data are ideal for studying the transitions into nursing homes out of the general population.2

Our analysis uses data from AHEAD questions on health, disability, income, wealth and family. The presence of information on this wide range of topics for a representative sample of older Americans is one of the AHEADs unique features. Other national surveys with health data on the elderly, including the National Long-Term Care Survey and the National Health Interview Survey, have poor coverage of economic or family structure variables; are not longitudinal in form; or survey only disabled individuals.

The age distribution of the AHEAD sample means that over the various waves, many respondents die or move into institutions. If a sample member is unwilling or unable to complete an interview because they are in a nursing home or ill, the study interviews a proxy respondent. If a sample member has died since the last survey, a final “exit interview” is conducted with a pre-designated proxy respondent. Nursing home use is recorded in four places in the AHEAD. First, the interviewer notes whether the respondent is in a nursing home at the time of the interview. Second, respondents who are alive (or their proxies) report if they spent a night in a nursing home since the last interview. Third, proxies for respondents who have died report whether the deceased died in a nursing home, and fourth, they report whether the deceased spent any other time since the prior wave in a nursing home.

Given this paper’s focus on the determinants of nursing home entry, we restrict attention throughout to persons from the nationally representative part of the sample whose patterns of nursing home use can be accurately determined. Our analysis sample thus consists of persons who (a) were from the nationally representative portion of the sample; (b) each completed the initial survey in 1993; (c) either completed two of the three interviewed after 1993, or died during the period but were responded by a proxy interview. Of the total 7443 persons meeting the age-eligibility condition, 6855 met the interview requirements. This group constitutes the analysis sample used throughout the paper.3

Table 1 shows the nursing home use in the analysis sample across the different waves of the survey. Twenty-six percent of respondents used a nursing home at some point between

2 In 1999, 87% of nursing home residents were over age 75, suggesting that though we miss some nursing home transitions by not having a sample that is even younger at baseline, we miss few of them (NCHS http://www.cdc.gov/nchs/agingact.htm).
3 We compare the means of observables for the approximately 600 persons who do not meet the interview requirements to means for the analysis sample. For virtually all of the key variables in the analysis, the persons
Table 1
Nursing home use in AHEAD, by wave

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of respondents at start of period</td>
<td>6855</td>
<td>6855</td>
<td>6102</td>
<td>5178</td>
</tr>
<tr>
<td>Proportion using nursing home at any point during interval (%)</td>
<td>25.9</td>
<td>8.5</td>
<td>15.0</td>
<td>16.0</td>
</tr>
<tr>
<td>Number of respondents alive at end of period</td>
<td>4357</td>
<td>6102</td>
<td>5178</td>
<td>4357</td>
</tr>
<tr>
<td>Nursing home use rate (%)</td>
<td>15.7</td>
<td>6.2</td>
<td>10.1</td>
<td>12.3</td>
</tr>
<tr>
<td>Number of respondents dead at end of period</td>
<td>2498</td>
<td>753</td>
<td>924</td>
<td>821</td>
</tr>
<tr>
<td>Nursing home use rate (%)</td>
<td>43.8</td>
<td>27.4</td>
<td>42.3</td>
<td>50.4</td>
</tr>
</tbody>
</table>

Note: AHEAD (n = 8222) was sample of community dwelling residents ages 70+ and their spouses in 1993. Our sample is drawn from the 7443 respondents who were from the nationally representative part of the sample, who met particular interview requirements.

1993 and 2000. Reading across the first row reveals that much of this nursing home use occurred in the later waves of the survey; a larger fraction of people entered institutions over the last two waves of the survey that across any other pairs of waves. This result is consistent with the idea that nursing home entry risk rises with age.

The second and third rows of numbers in table distinguish respondents by whether they were alive at the end of the indicated survey period. Two patterns are evident. The first – that use increases with age for both groups – we have already noted. Second, the table shows that over any particular interval, institutionalization was much more likely for persons who had died by the end of the period than it was for those who were still alive at period’s end. This pattern is as we would expect; those who had died by the end of any interval were more likely to have been in worse health and thus in greater need of institutionalized care. This result is also consistent with results from previous work. In our data, 43% of deceased respondents used a nursing home at some point after 1993. This estimate matches that from the National Mortality Followback Survey of Kemper and Murtaugh (1991), who project that 43% of individuals who turned 65 in 1990 will enter a nursing home at some time before their death.


Anderson’s (1968, 1995) influential behavioral model of health utilization suggests there are many factors apart from the receipt of informal care which determine individuals’ likely entry into nursing home. To gauge the marginal impact of informal care, we use the rich information available in the AHEAD to control for all of these other factors. Anderson’s analysis suggests that these other factors may be separated into three distinct categories.

excluded from the analysis matched the analysis sample. Thus, there was no significant difference in the number of ADLS, whether or not the person received informal care in the first wave of data in 1993, or marital status. For the few variables where a difference in the means existed, this difference was quite small. For example, we find that the excluded cases were 0.8 years younger than the analysis sample, and are marginally more likely to report poor health. On the whole, these mean comparisons suggest that their absence from the analysis sample raises little concern that any coefficient of interest may be biased as a result.
One set of variables – determinants of access – concern individuals’ ability to overcome the financial hurdles associated with entering a nursing home. These financial resources include things like income, net worth, Medicaid coverage, and ownership of long-term care insurance (LTCI) – all of which we control for in the various regression estimates that follow.

The literature has not consistently found an effect of financial resources on nursing home risk. Cutler and Sheiner (1994) find that home owners in the National Long-Term Care Survey (NLTCS), are less likely to enter nursing homes over a 2 year period, but that income and asset levels do not affect the risk of entry, after controlling for health differentials. Ettner’s (1994) study using the NLTCS also fails to find an effect of income. Garber and MaCurdy (1990) use data from the National Long-Term Care Channeling Demonstration of highly disabled individuals in the early 1980s, and also find that income does not affect the risk of nursing home entry.

By contrast, using the NLTCS Headen (1993) finds that individuals with pension income and real estate (other than their home) have 20 and 37%, respectively, lower hazard of nursing home entry between 1982 and 1984, suggesting that greater financial resources lower the risk of nursing home use. Finkelstein and McGarry (2003) find that long-term care insurance is not associated with greater future use of nursing homes because the families that buy these policies are more cautious and therefore face lower risks.

In fact, there are sound theoretical reasons why the different effects of financial resources may operate in opposite directions. On the one hand, nursing homes are expensive, costing $55,000 per year, on average, or about $150 per day (Kassner, 2004). Greater income should thus be associated with greater use. However, the eligibility rules of the Medicaid program and the fact that nursing home placement is often considered a less desirable long-term care outcome (i.e. a “bad” as opposed to a “good”), make the relationship between wealth and nursing home placement more complicated.

The effect of having Medicaid – another important “access” variable – is also somewhat complicated. Families on Medicaid face close to a zero cost of nursing home use because Medicaid pays for nursing home stays and there are essentially no deductibles. In many states, Medicaid will not pay for any home or community based alternative to nursing home care. Under such scenarios, individuals with Medicaid coverage at the start of any period should be more likely to enter nursing homes, holding all else constant. In some states however, Medicaid may pay for home and community based care for some individuals. In these cases, Medicaid coverage will not necessarily increase nursing home entry.

There is considerably less controversy about factors that determine individuals’ need for long-term care – the other key variable in Anderson’s taxonomy. The greatest risk factors of nursing home use are, of course, lower physical and mental health status. Older age and some of the conditions associated with it – difficulty with activities of daily living (ADLs), lower body limitations, greater hospital use, stroke, cancer, and hip fractures – are all associated with higher nursing home risk (Jette et al., 1995; Wolinsky et al., 1992). Our analysis controls for several risk factors, including whether a doctor has ever diagnosed the respondent as having had a stroke, diabetes, or psychiatric problems. We also control

---

4 In 2000, Medicaid expenditures for home and community based care totaled $4.9 billion whereas Medicaid spent $31 billion on institutional care (HCWM, 2000).
for respondents’ scores on a word recall test administered during the survey to measure memory acuity.

Finally we include a count of the number of activities of daily living with which the respondent reports having difficulty. Difficulty with ADLs should predict future nursing home placement both because it is a strong indicator of disability and long-term care need, and because standard long-term care insurance policies only pay benefits when the insured person has difficulty with two to three ADLs (Alecxih and Lutzky, 1996). The AHEAD data asks about six ADLs: walking, dressing, bathing, eating, rising out of bed, and toileting.

The third set of variables emphasized by Anderson are the various demographic factors which predispose individuals to make use of nursing homes. These include age, race, ethnicity, female gender, and education.

Table 2 provides key summary statistics for the analysis sample. About one quarter of the sample receives informal care in the baseline year, with an additional 5% receiving formal, paid care at home. At baseline in 1993, 9% of the sample had Medicaid, and 14% of respondents reported having private long-term care insurance. Mean net worth was about $192,000 and mean income was about $26,000. Having difficulty walking was the most common ADL impairment (24%) and difficulty eating or using the toilet the most rare (5%). Half of respondents were married at baseline and the mean number of living children was 2.7, split roughly evenly between sons and daughters.

Table 2
Descriptive statistics of AHEAD sample in 1993

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receives informal home care</td>
<td>0.25</td>
<td>0.44</td>
</tr>
<tr>
<td>Receives paid home care</td>
<td>0.05</td>
<td>0.22</td>
</tr>
<tr>
<td>Measures of access</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.52</td>
<td>0.50</td>
</tr>
<tr>
<td>#Sons</td>
<td>1.32</td>
<td>1.33</td>
</tr>
<tr>
<td>#Daughters</td>
<td>1.35</td>
<td>1.34</td>
</tr>
<tr>
<td>#Kids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial resources</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>25730</td>
<td>29667</td>
</tr>
<tr>
<td>Networth</td>
<td>192090</td>
<td>349233</td>
</tr>
<tr>
<td>Long-term care insurance</td>
<td>0.14</td>
<td>0.35</td>
</tr>
<tr>
<td>Medicaid</td>
<td>0.09</td>
<td>0.28</td>
</tr>
<tr>
<td>Demographic variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED = HS Grad</td>
<td>0.41</td>
<td>0.49</td>
</tr>
<tr>
<td>ED = some college+</td>
<td>0.14</td>
<td>0.34</td>
</tr>
<tr>
<td>Female</td>
<td>0.60</td>
<td>0.49</td>
</tr>
<tr>
<td>Non-white race</td>
<td>0.10</td>
<td>0.30</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.04</td>
<td>0.19</td>
</tr>
<tr>
<td>Age</td>
<td>77.47</td>
<td>5.89</td>
</tr>
<tr>
<td>Measures of need</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty walking</td>
<td>0.24</td>
<td>0.42</td>
</tr>
<tr>
<td>Difficulty bathing</td>
<td>0.12</td>
<td>0.33</td>
</tr>
<tr>
<td>Difficulty eating</td>
<td>0.05</td>
<td>0.22</td>
</tr>
<tr>
<td>Self rated health score (1 = Exc., 5 = poor)</td>
<td>3.03</td>
<td>1.18</td>
</tr>
<tr>
<td>n</td>
<td>6855</td>
<td></td>
</tr>
</tbody>
</table>
Table 3
OLS regression of nursing home use between 1993 and 2000, among individuals ages 70+ and community dwelling in 1993

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informal care receipt</td>
<td>0.12** (0.01)</td>
<td>0.04** (0.02)</td>
</tr>
<tr>
<td>Married</td>
<td>−0.08** (0.01)</td>
<td>−0.07** (0.01)</td>
</tr>
<tr>
<td>Education = HS Grad</td>
<td>0.01 (0.01)</td>
<td>0.05** (0.01)</td>
</tr>
<tr>
<td>Education &gt; HS Grad</td>
<td>0.01 (0.02)</td>
<td>0.07** (0.02)</td>
</tr>
<tr>
<td>Female</td>
<td>0.01 (0.01)</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>Non-white race</td>
<td>−0.03** (0.02)</td>
<td>−0.07** (0.02)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>−0.12** (0.02)</td>
<td>−0.12** (0.02)</td>
</tr>
<tr>
<td>Age</td>
<td>0.07** (0.02)</td>
<td>0.06** (0.02)</td>
</tr>
<tr>
<td>Age squared</td>
<td>−0.0003** (0.0001)</td>
<td>−0.0003** (0.0001)</td>
</tr>
<tr>
<td>Self rated health</td>
<td>0.01** (0.0055)</td>
<td></td>
</tr>
<tr>
<td>Long-term care insurance</td>
<td>0.0007 (0.0147)</td>
<td></td>
</tr>
<tr>
<td>Medicaid</td>
<td>−0.0109 (0.0219)</td>
<td></td>
</tr>
<tr>
<td>Networth</td>
<td>0.0000** (0.0000)</td>
<td></td>
</tr>
<tr>
<td>Log income</td>
<td>−0.0214** (0.0070)</td>
<td></td>
</tr>
<tr>
<td>Control for ADLs and health conditions</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.13</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Dependent variable = 1 if respondent was in a nursing home at any point 1993–2000.
** Denotes statistical significance at the 5% level.

To assess the marginal effect of informal care and control variables on the likelihood of nursing home use, we estimate OLS models for the probability of entering a nursing home between 1993 and 2000 of the form:

\[ NH_i = \beta_1 IC_{i1993} + \beta_2 Access_i + \beta_3 Need_i + \beta_4 X_i + \gamma_i, \]  

where $NH_i$ equals one if the respondent has stayed in a nursing home at any time since the first interview in 1993. The variable $IC_{i1993}$ equals one if the respondent was receiving informal home care at the time of the first interview in 1993. Respondents are asked whether they receive assistance with any activities of daily living. A respondent is coded as receiving informal care if they have a helper with at least one ADL. “Access” and “Need” are the two vectors of variables described above, and $X$ is a vector of demographic controls including measures for the respondent’s marital status, and number of children.

If informal care and nursing home use are substitutes, the estimate of $\beta_1$ should be negative, unless the receipt of care is correlated with unobserved factors that independently raise nursing home use. In the regressions, we use baseline (1993) values for these variables rather than time varying ones observed when we observe nursing home use in future waves, because these later values may be endogenous to nursing home residence.

Table 3 presents the results from the estimation of different versions of (1). We also estimated (1) by logistic regression, but do not present those results. On the whole,

---

5 Having children and having daughters in particular is also associated with lower nursing home risk (Freedman, 1996; Garber and MaCurdy, 1990).

6 These results are available upon request.
the results from the logit models are statistically and substantively similar, but we focus on the OLS estimates because of their close connection to the TSLS estimates to follow.

Contrary to the belief that informal care is a substitute for nursing home use, the results indicate that the receipt of informal care in 1993 is associated with a higher probability of entering a nursing home by 2000. There is a 12% point higher probability when we control only for basic demographics (column 1), and a 4.3% point difference when we control fully for all of the access, risk, and other variables (column 2). Both estimates are statistically significant. The estimate obtained when we control for all the access, risk and other variables suggests that receiving informal care is associated with roughly a 16% higher risk of nursing home entry, relative to the mean.

Family and financial resources appear to be associated with lower nursing home risk. Being married reduces the risk of nursing home entry by 6% points. This is consistent with previous results from Lakdawalla and Schoeni (2003) and Freedman (1996). Although Medicaid and private long-term care insurance have no noticeable effect on nursing home risk, higher income and higher net worth are both associated with significantly though not substantially lower risk. An increase in net worth of $100,000 is associated with a 3.5% point lower nursing home probability, and at a mean income of about $25,000, an increase in annual income of $10,000 is associated with just under a 1% point reduction in nursing home use probability. The negative coefficients strongly suggest what many believe but have not been able to show empirically — that nursing homes are an inferior good. The estimated negative effect of net worth could also be due to families “spending down” to qualify for Medicaid coverage of nursing home expenses. In this case, families would be strategically spending or bequeathing their assets so that they could qualify for Medicaid in the future.

As the literature has consistently found, we find that greater individual need for long-term care is generally associated with higher nursing home risk. We do not present the point estimates in the table, but find that three of the six ADLs have positive coefficients. We also find that cognitive impairments, as measured by the word recall score or a report of having memory problems significantly increase the risk of nursing home use. Finally, having diabetes, a condition that is associated with higher risk of heart disease, blindness, and amputation, increases the probability of nursing home use by 5% points. We do show the results for a summary measure of health – how the person rates his or her own health.

---

7 To conserve space, we do not present all of the point estimates in the table, but they are available from the authors upon request.
8 Medicaid eligibility can occur in a number of ways. Nursing home residents who are covered by Medicaid may have had Medicaid before they entered the nursing home, due to low income and wealth. This is the Medicaid coverage we observe at baseline. They may also qualify for Medicaid upon entering a nursing home by way of “Medically Needy” rules, and a number of other rules if the ratio of nursing home expenses to their income is high enough. We cannot control for this empirically.
9 To minimize this sort of “gaming” of the system, Medicaid considers asset levels for the 3 year period prior to application for Medicaid.
10 When we estimate the equation using number of ADL difficulties rather than enter each ADL separately (not reported), we estimate that an additional ADL impairment significantly increases nursing home probability by two percentage points.
in 1993. The table shows that worse (higher) self-rated health score significantly increases the probability of future nursing home use.\footnote{We also control for recent declines in health and find that this, too, is associated with greater nursing home use.}

In summary, virtually all of the point estimates in Table 3 are to be expected, with the noteworthy exception of the estimated effect of informal care. Because informal care receipt signals a need for long-term care, there is some reason to think that it would predict future nursing home use. However, the fact that it is associated with a greater risk of nursing home entry, \textit{after} we control for the key factors that should affect nursing home utilization -- health and disability -- is inconsistent with the presumed theoretical relationship of substitutability between informal care and institutional care. Thus, even with longitudinal data, there is no evidence that informal care can reduce the risk of nursing home entry.

It is quite likely that survey measures of health and disability do not perfectly capture variation in long-term care needs. If this is the case, the measure of at-home care receipt in (1) would pick up the effect of unobserved bad health characteristics. If individuals who receive informal care are those with greater unobserved disability (holding constant observable health characteristics), the estimate of $\beta_1$ would be biased upward and this could explain why the OLS estimate suggests that receipt of informal care increases the risk of nursing home entry.

One way to address this problem is to use an instrument that is correlated with the receipt of informal care but not directly correlated with nursing home risk. In the next section, we discuss an instrumental variables approach for measuring the effect, net of any bias arising from endogeneity of care receipt.

5. Two-stage least squares model of nursing home entry

A large literature documents the higher propensity of daughters, and unmarried daughters in particular, relative to sons, to provide informal care to their parents (McGarry, 1999; Wolf et al., 1997; Coward and Dwyer, 1990). We exploit variation in the gender, marital status, and location of one’s offspring as instruments for informal care receipt. Conditional on marital status and the number of children, the gender composition of those children should be exogenously determined for a family of a given size. Though the gender of one’s children has been found to affect nursing home entry risk, we believe it primarily does so indirectly, through the channel of informal care. Individuals with a spouse or children are less likely to enter a nursing home because these family members provide alternative types of care in the home. Although the marital status and geographic location of one’s children may be endogenous to the parents’ health at a point in time, we use baseline values for these variables and argue that they are valid instruments for our test of the effect of informal care.\footnote{To test whether children’s location endogenously responds to parents’ health in our sample, we estimated a logit equation of the probability that a child lives close by (within 10 miles) as a function of all of the health variables and other controls used in Eq. (1). We find almost no evidence that children are more likely to live close by when their parent is in bad health. Of all the health variables, only arthritis is associated with a higher probability of having a child live close by.}
Table 4
First stage: OLS estimates of effect of child characteristics on receipt of informal care

<table>
<thead>
<tr>
<th>Instrument Formulation 1</th>
<th>Instrument Formulation 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has a daughter close by</td>
<td>Has a daughter close by</td>
</tr>
<tr>
<td>0.03** (0.01)</td>
<td>0.01** (0.01)</td>
</tr>
<tr>
<td>Has an unmarried daughter</td>
<td>Has an unmarried daughter</td>
</tr>
<tr>
<td>0.00 (0.01)</td>
<td>-0.01* (0.01)</td>
</tr>
<tr>
<td>Has a married daughter</td>
<td>Has an married daughter</td>
</tr>
<tr>
<td>-0.02* (0.01)</td>
<td>-0.01 (0.03)</td>
</tr>
<tr>
<td>Has a son close by</td>
<td>Has a son close by</td>
</tr>
<tr>
<td>0.04** (0.01)</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>Has an unmarried son</td>
<td>Has a married son</td>
</tr>
<tr>
<td>0.00 (0.01)</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>Has a married son</td>
<td>% Children that are married daughters</td>
</tr>
<tr>
<td>0.01 (0.01)</td>
<td>0.00 (0.02)</td>
</tr>
<tr>
<td>% Children that are married sons</td>
<td>% Children that are married sons</td>
</tr>
<tr>
<td>0.01 (0.01)</td>
<td>0.01 (0.02)</td>
</tr>
</tbody>
</table>

$R^2$ 0.35 0.35  
$F$ statistic for coefficient = 0  5.94 5.42  
$P$-value for test statistic 0.00 0.00

Dependent variable = 1 if respondent received unpaid help with health or personal care in 1993.

* Denotes statistical significance at the 10% level.

In 1993, AHEAD respondents reported whether each of their children live within 10 miles of them. If they do, we code them as being “close by”. Of the roughly 1.3 average children of each gender, more live far away (0.8) than live close by (0.5), and this pattern is about the same for sons and daughters. We hypothesize that respondents whose children live close to them will be more likely to receive informal care than others, since the time cost of providing care is lower for these children. It is also possible that married children may have differential propensities to provide care. For example, marriage may make a child less likely to provide assistance because of competing responsibilities. It could also result in a higher probability of receiving care, because of the addition of another family member as a potential caregiver (a daughter-in-law or son-in-law). In AHEAD, 30% of children are married daughters, 14% are unmarried daughters, 31% are married sons, and 12% are unmarried sons.

The effect of the gender distribution, marital status and location of a person’s children on the likelihood of that person receiving informal care is assessed on the first stage results shown in Table 4. We present two sets of results. In one (Instrument Formulation 1), we create a set of dummy variables denoting, for example, whether a person has any unmarried daughters, or any sons close by. In the other (Instrument Formulation 2), we code the variables as raw numbers and percentages. In these first stage regressions, we control fully for all of the variables used in the regression in Table 3. The results in the table are from OLS regressions.

In both panels, the results show that the presence and number of children (measured as of 1993) are correlated with receipt of informal care.\(^{13}\) For example, having a daughter or son who lives nearby is associated with a three or 4% point higher probability of receipt of informal care. This translates to an increase of 12–16%, given that a quarter of the sample

\(^{13}\) The $F$-statistic from the test that the coefficient estimates is jointly zero, is 5.94 in Instrument Formulation 1 and 5.42 in Instrument Formulation 2, resulting in the ratio $1/F=0.17$ and 0.18. We report this statistic to increase confidence in the strength of our instruments. Bound, Jaeger, and Baker (1995) express concern of weak instruments as the ratio approaches one (i.e. for low $F$).
Table 5 contains TSLS estimates of Eq. (1), where we instrument for the receipt of informal care using the two sets of instruments outlined above in turn. Column 1 presents results for our base specification. All of the health, demographic and other variables listed in Table 3 are controlled in these regressions, but we suppress the point estimates to conserve space. Using the two sets of instruments, we find that informal care in 1993 is associated with significantly lower risk of nursing home entry between 1993 and 2000. The estimates suggest that receipt of informal care reduces the probability of any nursing home use by 39–49% points.

Given the mean of nursing home entry of 0.25, the magnitude of the TSLS estimates may appear very large. Are they plausible? We think that they are, especially when one considers that this estimate is from a sample of individuals who are, on average, at high risk of entering a nursing home because on their health and disability. More importantly, Angrist and Imbens (1994), have shown that TSLS models estimate what has been termed the “Local Average Treatment Effect” (LATE). That is, the TSLS results are estimated off of those persons who would have entered nursing home but for the gender and location composition of their children. The fact that the estimates in the first column of the table are estimated off of the behavior of this population, we would expect the point estimate to be relatively large.

To better interpret the magnitudes of the point estimates, consider what our results imply about an unmarried woman in 1993, who was 80 years old, had difficulty walking and dressing but not with other ADLs, rated her self-rated health as fair and had the average values for all other variables. Given the coefficient estimates, this woman has an 85% chance

14 Much of the literature finds that daughters are more likely to provide care than sons. Though this could mean that individuals with daughters are more likely to receive care, it need not be the case. For example sons may only provide care if they have no sisters.
of entering a nursing home by 2000 if she receives no informal care. The TSLS estimate from Formulation 2 predicts that receipt of informal care reduces her risk to 46%.

Despite our belief that the estimates are reasonable, the strong contrast between the OLS and TSLS estimates, and the large magnitude of the latter argue for robustness tests to assess the strength of our qualitative results. Estimates from these robustness tests are presented in columns 2–4 of Table 5.

One source of concern with our instrumental variables strategy may be that child location is endogenous. Children with sick parents may relocate to a residence closer to their parents. This concern is mitigated for several reasons. For one thing, we observe the location of the children in 1993 and nursing home use is observed through 2000. Thus child location is exogenous to unexpected changes in health status after 1993. Of course, some of the respondents may have exhibited some need for care by 1993 and child location in 1993 may have been in response to these anticipated changes, making the child location measures endogenous. We tested whether children’s location endogenously responds to parents’ health in our sample, and found no evidence of this.15

To further assess the importance of possible endogeneity of 1993 child location, column 2 of Table 5 reports TSLS estimates when we exclude nursing home visits between 1993 and 1995 from the analysis. By focusing only on nursing home use between 1995 and 2000, it is less likely that child location in 1993 would have been in response to unanticipated health in these future periods. The table shows that our results do not change qualitatively with this restriction.

The most important reason for our confidence that the results are not driven by endogenous child location is that any bias arising from this should bias our estimated effects towards the OLS estimates of a positive correlation between informal care receipt and nursing home use. Children are more likely to live close by if their parents are in bad health. Parents in bad health are more likely to enter nursing homes, and if this worse health is unobservable, it would bias the TSLS estimate of informal care towards a positive finding.

A more serious concern with the instrumental variables results is that the gender composition and location of children affect nursing home risk in ways other than by providing informal care. Children may provide financial assistance to parents to help them pay for formal home-care. In the AHEAD however, very few children provide financial assistance to their parents (McGarry and Schoeni, 1997). In addition only 5% of our sample receives paid home care (whereas 25% receive informal home care) so this is also unlikely to bias our results much. We actually find that having additional children, especially close by, is associated with a reduced probability of receive formal home care, likely due to the fact that informal and formal care are substitutes.

Another concern is the fact that there is substantial heterogeneity in nursing home use. Many individuals may enter a nursing home for a short period of time, for post-acute care after hospitalization, whereas others enter nursing homes for permanent residence. One way to differentiate these types of episodes is by the length of stay. Medicare pays for up to 100 days in a nursing home care after a hospitalization as “post-acute care.” As a result,

---

15 We estimated a logit equation of the probability that a child lives close by (within 10 miles) as a function of all of the health variables and other controls used in Eq. (1). Of all the health variables, only arthritis is associated with a higher probability of having a child live close by (results available by request).
Table 6
Second stage estimates from bivariate probit and two-stage probits of nursing home use as a function of informal care receipt (table presents point estimates, standard errors in parentheses, and marginal effect in curly brackets)

<table>
<thead>
<tr>
<th></th>
<th>Instrument Formulation 1</th>
<th>Instrument Formulation 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bivariate probit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Informal home care</td>
<td>−0.26 (0.37)</td>
<td>0.06 (0.30)</td>
</tr>
<tr>
<td>Rho</td>
<td>0.23 (0.21)</td>
<td>0.05 (0.17)</td>
</tr>
<tr>
<td>Log pseudo-likelihood</td>
<td>−17959339</td>
<td>−17931031</td>
</tr>
<tr>
<td>Modified two-stage probit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pr(Informal Care</td>
<td>X = X̄ )</td>
<td>−1.76** (0.523)</td>
</tr>
<tr>
<td>Pr(Informal Care</td>
<td>Z = Z̄ )</td>
<td>0.25 (0.42)</td>
</tr>
</tbody>
</table>

Estimation includes controls listed in Table 3. Excluded instruments are dummy variables for presence of: daughter living close by, married daughter, unmarried daughter, son living close by, married son, unmarried son. Pr(Informal Care | Z = Z̄ ) is the predicted probability of informal care receipt based on the coefficient estimates of the first stage probit equation, when the values of the excluded instruments are set to the mean. Thus, it picks up variation in informal care receipt due solely to variation in the covariates. Pr(Informal Care | X = X̄ ) is the predicted probability of informal care receipt based on the coefficient estimates of the first stage probit equation, when the values of all the covariates are set to the mean. Thus, it picks up variation in informal care receipt due solely to variation in the excluded instruments.

** Denotes statistical significance at the 5% level.

we consider episodes shorter than 100 days as post-acute care and episodes that are greater than 100 days as a long-stay. In columns 3 and 4, respectively, we report TSLS estimates of the effect of informal care on the probability of a long-stay nursing home, and on the probability of post-acute care. Not surprisingly, we find no effect on post-acute visits but large effects on long-stays. This suggests that informal care is a substitute for long-term nursing home residence but not for short-term nursing home care.

Our preferred TSLS models are widely used in the economics literature for discrete outcomes models of the sort we estimate. Nonetheless, it is a natural question how sensitive our results are to the linear specification implicit in the TSLS models. In Table 6 we present results for bivariate probit models.

In the first column, we present two sets of standard bivariate probit results. In the first of these, the excluded instruments in the informal care receipt equation are those denoted “Instrument Set 1” in Table 5. In the second, we use the instruments labeled “Instrument Set 2” in Table 5. The results show that the standard bivariate probit appears to yield results that do not strongly support the earlier set of results. In particular, the point estimate for the effect of informal care is not statistically significant in either specification, as is negative only when the first set of instruments are used.

In recent work, Altonji et al. (2004) point out that comparison of TSLS and bivariate probit results is not straightforward. In particular, they note that variation in the latter class of models comes not only from variation in the instruments (as is true with TSLS) but also from non-linearities. They suggest an approach for distinguishing between the two sources of variation. Their approach allows a comparison between the two classes of models that is more revealing than a comparison of the point estimates for the endogenous regressor in the two sets of models.

Their recommended two-step approach is as follows. In the first step, a probit model for the endogenous regressor (informal care in our context) is estimated on the full vector of
controls variables \((X_i)\) and the instruments \((Z_i)\). Next, a second probit model is estimated – this for the relevant outcome variable, or nursing home entry in our paper. In the standard application of the bivariate probit, the regressors in this second probit are the full set of controls and the predicted probabilities for the endogenous regressor from the first stage, \(\Phi(\hat{\beta}_1 X_i + \hat{\beta}_2 Z_i)\). In the Altonji et al. modification, these predicted probabilities are computed, in turn, with the value of the instruments and the value of the controls set to the sample means – that is \(\Phi(\hat{\beta}_1 X_i + \hat{\beta}_2 Z_i)\) and \(\Phi(\hat{\beta}_1 \bar{X} + \hat{\beta}_2 \bar{Z})\). When the instruments are set to the sample mean, the two step procedure estimates that portion of the standard bivariate probit estimate due to non-linearities; when the controls but not the instruments are set to the sample mean, the two step procedure estimates the portion of the bivariate probit estimate attributable solely to variation in the instruments. This last is most closely analogous to the TSLS estimate under the linear specification.

In the last two columns of Table 6 we present the results for this two step procedure for both of the instrument sets. The standard errors for these results are bootstrapped. Both sets of results show that when we isolate solely the variation due to differences in the value of the instruments, effect of informal care on nursing home entry is quite consistent with the TSLS results. Under both characterizations of the instrument set, the point estimates are strongly statistically significant, and the estimated marginal effect is qualitatively very similar to our preferred TSLS estimates. These results show clearly that when variation in the instruments in held constant, the non-linearities in the controls causes the point estimate to be statistically indistinguishable from 0. It is this effect that produces the overall point estimate to be zero in the standard bivariate probit results. On the whole, we think that these results are strongly consistent with our argument that informal care lowers nursing home use, once endogeneity of care receipt is accounted for.

6. Discussion

Using an instrumental variables strategy, we find that receipt of informal home care substantially reduces the probability of future long-term nursing home use. This result is in sharp contrast with OLS estimates from the same set of data which show a positive correlation between informal care receipt and future nursing home use.

The cohort of individuals presently entering nursing homes has more children on average than the baby boomers that will be entering nursing homes some decades from now. Thus, aging baby boomers may be less likely to receive informal care to assist them as they age. Our findings suggest this will lead to an increase in nursing home use rates.

The finding that home based long-term care reduces the probability of nursing home entry has important implications for public policy and the private market for long-term care insurance. Private long-term care insurance policies, most of which traditionally only paid for institutional care in a nursing home, could reduce costs by adding coverage for home based care, by reimbursing private home care providers or paying family caregivers. Concerns about moral hazard for home based care would continue to be a concern in pricing.

---

16 The two regressions are often estimated simultaneously in the standard bivariate probit model, but the procedure can be broken into two steps. The standard errors must be bootstrapped if this approach is followed.
these policies effectively however. Greater public financing for home and community based long-term care, and subsidies or payments to family caregivers may be an effective way to reduce nursing home use. One example of this is the National Family Caregiver Support Program which was authorized by the Older Americans Act in 2000 in an effort to reduce nursing home utilization through the support of family caregivers. Because nursing homes are generally a more costly source of long-term care, such subsidies could reduce overall public expenditures on long-term care and potentially ease growing pressures on burgeoning state Medicaid budgets.

Acknowledgements

The authors are grateful to Joe Altonji, Sandra Decker, Lina Walker, and Tony Webb for thoughtful comments and suggestions.

References

Boaz, R.F., Muller, C., 1994. Predicting the risk of “permanent” nursing home residence: the role of community help as indicated by family helpers and prior living arrangements. Health Services Research 29 (1).