The Lifetime Risk of Nursing Home Use

Michael Hurd
RAND, NBER and NETSPAR

Pierre-Carl Michaud
ESG UQAM and RAND

Susann Rohwedder
RAND and NETSPAR

Abstract

This paper estimates the lifetime distribution of stays in nursing homes using 10 waves of HRS data covering the population age 50+. Using both non-parametric and parametric approaches which account for censoring, we estimate that a 50 year old has a 60% chance of ever entering a nursing home before he dies and that, conditional on any stay, the average duration is just over a year. We show that stays at the end of life which are typically not captured in core interviews are very important for assessing lifetime exposure. The HRS performs exit interviews with proxies for those who died. Excluding exit interviews yields lifetime risk under 40%. Being female, white and a non-smoker are associated with higher lifetime risk due to lower (competing) mortality risk and higher nursing home risk at older ages.

Keywords : nursing home, long-term care
JEL Codes : I13, J14

Acknowledgements
This research was supported by the National Institute on Aging, under grant 1R01AG041116. We thank Joanna Carroll for excellent research assistance and David Boisclair for editorial assistance. Any remaining errors are our own.

1. Introduction

The risk of spending for long-term care is one of the most important risks faced by older households because of the long right tail of days spent in nursing homes. However, finding data to estimate the risk has been difficult because of the necessity of following individuals over long periods of time. In this study, we use data from the Health and Retirement Study (HRS) to assess the lifetime distribution of stays in nursing homes and what these indicate for long-term care risks faced by households. While the HRS only samples from the non-institutionalized population at baseline, the follow-up of this longitudinal survey includes all baseline respondents, in particular those who move to a nursing home. As a result, after several waves the HRS will also represent the nursing-home population because of turnover in nursing homes: almost all those in nursing homes at baseline will have died and been replaced by persons initially residing in the community and represented by HRS respondents. We use 10 waves of the HRS including cohorts added after the original HRS cohort, those born in the years 1931-1941. Those additional cohorts were added in years following the initial interview of the HRS cohort in 1992. In addition to the core interviews we use data from the proxy interviews, usually with a spouse or other close relative, for those unable to participate in a given interview wave. Most importantly we use data from exit interviews which are conducted with a proxy after the death of a respondent. Our use of all waves of all relevant HRS cohorts as well as exit interviews allows us to estimate lifetime risk of a nursing-home stay both nonparametrically and with a flexible transition model which allows us to simulate nursing home histories.

Prior results

The types of studies that are most relevant to this study are those that estimate the lifetime chances of ever being in a nursing home (lifetime risk), those that estimate durations of stays in nursing homes, either conditional on a stay or unconditional, and those that estimate the lifetime duration in nursing homes. With respect to the first category, lifetime risk, the estimates range up to 55% (Arling, Hagan, and Buhaug, 1992). A widely cited rate is 37% for those over the age of 65 (Kemper and Murtaugh, 1991), and there are a number of other similar estimates.[1] A common finding is that lifetime risk is higher for women than for men, as for example, in Brown and Finkelstein (2008) 44% for women and 27% for men. As for durations of stay, there are often conditioning events that make comparisons across studies difficult. For example, Dick, Garber and MaCurdy estimate the average length of stay to be 21 months conditional on entering. Arling, Hagan, and Buhaug (1992) estimate that on average individuals who have a nursing home stay will spend 24 months in a nursing home. Conditional on dying in a nursing home, the mean length of stay estimated over the seven waves of AHEAD data was 14 months (Kelly, Anne *et al*., 2010). The average total time spent in nursing homes over a lifetime, counting multiple stays, was estimated to be 2.3 years in Liang *et al*. (1985). Regardless of the exact estimate, the duration of individual stays and of accumulated lifetime stays is substantial, putting individuals that must pay out-of-pocket at considerable financial risk.

2. Data

2.1 The HRS

Our analyses are based on data from the Health and Retirement Study (HRS). The HRS is a biennial longitudinal survey which covers a broad range of topics including income, work, assets, pension plans, health insurance, disability, physical health and functioning, cognitive functioning, and health-care expenditures. Its first wave was conducted in 1992. The target population was the cohorts born in 1931-1941, also called the original HRS cohort (Juster and Suzman, 1995). Additional cohorts were added in 1993 of those born in 1923 or earlier (AHEAD cohort), and in 1998 of those born between 1924 and 1930 (CODA cohort) and of those born between 1942 and 1947 (War Babies), so that in 1998 the HRS represented the population at least 51 years of age. New cohorts have subsequently been added every six years, including, in 2004, Early Baby Boomers born from 1948 to 1953, again making the HRS representative of the population 51 or older.

[1] See Cohen, Tell, and Wallack (1986), Dick, Garber and MaCurdy (1994), and Spillman and Lubitz (2002).

Every cohort is sampled from the non-institutionalized population at baseline so that the population residing in nursing homes is not represented. However, HRS makes extensive efforts to follow respondents after the baseline, including those who move into institutional settings such as nursing homes. Core interviews are conducted every two years. If a person is too frail or cognitively impaired to be interviewed, a proxy interview is conducted instead with a spouse or close relative. If a respondent dies between waves, HRS will attempt to conduct a so-called exit interview with a proxy informant, preferably someone who is knowledgeable about the family and financial situation and the circumstances preceding the person's death. Because nursing-home stays are most prevalent among the frail, cognitively impaired, or those close to death, the information obtained in proxy and exit interviews is critical for assessing the prevalence and incidence of nursing home stays. If despite all efforts HRS cannot make contact with a respondent or any relative, the HRS conducts tracking efforts, with special emphasis on determining whether the respondent may have died. In these efforts HRS cross-checks data sources such as the Social Security Death Index (SSDI) and the National Death Index to ascertain whether the respondent has died.

Table 1 summarizes characteristics of the different cohorts included in the HRS. By the time wave 10 of the HRS was conducted in 2010, 14 percent of the oldest cohort was still alive, as was 79 percent of the youngest cohort. For few respondents in any wave was vital status unknown. The HRS is quite successful in completing exit interviews with a proxy, gathering such for 96 percent of deceased respondents.

For all cohorts in the HRS, an analysis of nursing-home stays will potentially suffer from either left-censoring (do not observe nursing home spells prior to initial wave) or right-censoring (do not follow until death so there could be spells beyond wave 10 in 2010). The importance of such censoring depends on the cohort and the initial age. For example, the subsample of AHEAD respondents who were age 85 at baseline in 1993 have all died by 2010. So there is no right-censoring among this group of respondents, but there is considerable left censoring because no nursing home stays prior to age 85 have been observed for these respondents. Conversely, with younger HRS respondents, there is very little left censoring but considerable right censoring, because by 2010 they are still in their early 70s and the majority of their nursing-home stays will still be in the future and is not observed in the data. Both parametric and non-parametric analyses will address these issues.

2.2 Information on nursing home stays in the HRS

The HRS collects the following information on respondents' nursing home stays in the core HRS interview:

- whether the respondent was residing in a nursing home at the time of interview
- if so, when the person moved to the nursing home (or if the person had stayed there continuously since the previous interview)
- whether the respondent had any (other) nursing home stay since the previous interview
- if so, how many nursing home stays in total
- If one stay: how many nights spent in nursing home;
  if more than one stay: how many nights in total spent in nursing home.

The HRS also asks about the month and year of the nursing home entry and exit for up to three spells which can be used to cross-check or complement the information on the total number of nights spent in the nursing home.

The exit interviews ask for the same information with reference to the time between the last interview the respondent completed and the respondent's time of death. We integrate the information obtained in the exit interviews into our key outcome measures:

Any nursing home stay in previous two years. For respondents who participated in a particular wave *t* and in the immediately preceding wave *t-1* this variable takes the value one if the person was in the nursing home either any time between waves or is currently in a nursing home. For those respondents who died between the two waves, we use the information obtained from the exit interviews. If those indicate that the respondent was in a nursing home any time between the preceding wave *t-1* and the time of death then this variable will take the value one in wave *t*. If someone missed one or more interviews this measure would cover a longer period. For exit interviews which capture the information of those respondents who died between waves the period covered tends to be less than two years.

Number of nights spent in nursing home previous two years. The construction of this variable follows the same principle. It uses the information from the HRS core interview for all those respondents who survive, and the information from the exit interviews for those respondents who died between waves.

Lifetime measures of "any nursing home stay" and the "number of nights spent in a nursing home". These measures cumulate the survey information of any nursing home stay and the number of nights spent in a nursing home in previous two years over all waves up to the last wave collected in the year 2010.

For our analysis, we use sample weights at initial interview.

2.3 Population representativeness with respect to the nursing home population

HRS draws its baseline sample from the non-institutionalized population, but then follows up with all respondents, including when they move to nursing homes. We want to establish how many waves it takes until the HRS survey reaches population representation with respect to the nursing home population. Figure 1 shows the fraction in nursing-home residence by age for each wave of the AHEAD cohort. Because the AHEAD wave 1 sample is drawn from the community, nursing-home stay prevalence upon entering the study is zero at all ages. By wave 2, substantial numbers are living in nursing home, for example, 7.1% of those 86-87 years of age. Nonetheless the curve for wave 2 mostly lies below the curves for later waves suggesting that after two years, nursing homes still had residents that were not represented in the initial AHEAD wave. By wave 3 (1998) or five years after the baseline wave, the prevalence of residing in nursing home was about the same as in later waves, leading us to surmise that by then, at least as far as prevalence is concerned, AHEAD cohort was representative of the entire population, not just the community dwelling population.[2] Thus in addition to left- and right-censoring, non-parametric estimation of the risk of any stay must account for start-up, the fact that the initial waves did not represent adequately the nursing home population.

[2] This statement cannot be true for the entire nursing home population, e.g. those who survive for more than five years in a nursing home.

2.4 The importance of the exit interviews for assessing lifetime nursing home exposure

Some nursing home stays are short-term, beginning and ending between waves. As a result, the measure of nursing-home residence at the time of interview is not suitable for measuring lifetime exposure. Figure 2 adds exposure between waves. It shows, for example, that, among 86-87 year olds in wave 2, 11.6% had nursing-home exposure between wave 1 and wave 2 (including residence at wave 2) but just 7.1% were in residence (Figure 1). This indicates the importance of shorter-term stays. The figure shows that by wave 3, nursing-home exposure between waves was at about the same level as in later waves, again illustrating that the first two waves cannot be used to show nursing-home exposure.

Because of the importance of relatively short-term stays, researchers will underestimate nursing-home exposure if their estimates rely on interviews with respondents who are alive in each wave. Respondents who were living in the community in a wave, experienced a nursing home stay following that wave, and died before the succeeding wave would not be recorded as having a nursing-home stay. In the exit interview, the proxy respondent is asked about nursing home stays since the previous interview. We find that including them increases substantially the estimate of nursing home exposure. Figure 3 illustrates their importance. Consider 86-87 year-olds in wave 2. Adding those who were interviewed in wave 1 and would have been 86-87 years old in wave 2 had they survived shows that nursing-home exposure in that larger group was 15% between waves 1 and 2, rather than 12% among survivors to wave 2 (Figure 2). Thus the use of the exit interview increased nursing home exposure by 3.5 percentage points or 32%.

3. Age-prevalence of nursing-home stays and lifetime exposure

3.1 Age-prevalence of nursing home stays and nursing home use

In the calculations of the age-prevalence we exclude the first two waves of data for each cohort, because of the lack of representation of the nursing home population as discussed above. We pool all remaining waves and cohorts and apply respondent weights. Table 2 shows the age-prevalence for two measures: the fraction residing in a nursing home at the time of interview and the fraction with any nursing home stay in the previous two years. For the latter we included a column incorporating the information from the exit interviews and one without ("core only") to highlight once again the much higher prevalence obtained when including the information from the exit interviews. The exit interviews capture the information of those respondents who died between waves that would otherwise be missed. Because nursing home stays are most prevalent towards the end of life this is an important omission.

The fraction residing in a nursing home at the time of interview is low, less than one percent, at ages less than 70. At ages 70 and older the fraction approximately doubles with every 5-year age band up to age 90-94 when it reaches 23 percent. Among those surviving to age 95 and above the fraction residing in a nursing home is 36 percent. The measure assessed at the time of interview reflects just a moment-in-time whereas the next column measures any nursing home stays that have occurred in the previous two years, and includes the exit interviews. At ages up to 84 the fraction with any nursing home stay in the previous two years is higher than the moment in time measure by a factor of three or more reflecting the importance of short-term stays at relatively younger ages. At the oldest ages it is 20 percentage points higher than the moment in time measure. Among 90-94 year-olds 42 percent resided in a nursing home sometime in the previous two years, and among those age 95 or older 57 did so. The final column has similar statistics but does not use the exit interviews. Overall the exposure to nursing homes is 2.7 percent lower, but at some ages the discrepancy is much greater: at ages 90-94 it is 11.4 percentage points lower.

Table 3 provides the average by age-band of the total number of nights spent in a nursing home in the previous two years, again both with and without consideration of the exit interviews. For the number of nights the differences between the two columns are noticeable, but not particularly large. The explanation is that the exit interviews capture the information for those who died between waves for whom the period covered since the last interview is on average just one year and not two years as for the remainder of the sample.

Focusing on the column that incorporates the exit interviews, the total number of nights spent in a nursing home in the previous two years averaged over the entire sample (unconditional on nursing home stay), approximately doubles every five years between the ages

60 and 95, reaching 130 nights for those age 90-94. Among those age 95 or older the average number of nights in a nursing home is 204.

3.2 Lifetime risk of a nursing-home stay

Using the long panel dimension of the HRS we show in Table 4 estimates of lifetime exposure obtained from the raw data without – for now – addressing the issue of left- or right-censoring. We start with the HRS cohort born 1931-1941 and observed from 1992 until 2010. In the youngest age band of HRS (age 50-54), about 21% had died by 2010. About 10% had a nursing home stay, and the average number of stays was 0.18, indicating that some individuals had multiple stays. The average number of nights was 24 including those with no nursing home exposure. These statistics increase with age. For all HRS cohort respondents the average number of stays was 0.26 and the average number of nights was about 36, indicating that the typical stay was about 135 nights. Among those who died before wave 10, 18 years after wave 1, 29% had a nursing home stay and the average length of stay was 86 nights.

In the AHEAD cohort, mortality was essentially complete for those initially age 80-84. In that group 59% were in a nursing home at some time. The average number of nights was about 300. With the exception of the youngest age group in the AHEAD cohort the difference in nursing home exposure between everyone initially in an age band and those who died before 2010 is not substantial, indicating that right censoring is not very important; that is, among those initially ages 75 or older we are close to observing rest-of-lifetime nursing home risk. In that group 54% used a nursing home at some time before 2010. Thus, among those who survive to ages 75-79, a lower bound on rest-of-life lifetime nursing home exposure is 54%. There is, of course left-censoring, which would increase the lifetime risk of those who survive to that age.

Nonparametric estimation of lifetime risk of any nursing home stay

Our nonparametric estimation of nursing home exposure is based on Figure 4. It combines nursing home exposure and transition probabilities from three cohorts. The main and central cohort is AHEAD wave 1 respondents whose initial ages were 70 to 74. By 2010, at which time the cohort would have been 87 to 91 years old, 67% of that cohort had died and 33% were still alive. Among those who died, 48% were in a nursing home sometime prior to death. Among the 33% who survived 29% had nursing home exposure. To estimate the effect of right censoring among the 71% who survived and had no nursing home exposure, we use the AHEAD wave 1 respondents whose initial ages were 85 to 89. By 2010 all of that cohort had died, and 64% were in a nursing home prior to death but following the initial wave in 1993. These were "fresh" nursing home exposures because AHEAD wave 1 only sampled those in the community. Combining these probabilities, we estimate nursing-home exposure of the initial AHEAD cohort aged 70-74 to be 57% (0.67*0.48+0.33*(0.29+0.71*0.64).[3]

This figure needs to be adjusted in several ways. The initial AHEAD sample in 1993 excluded residents in nursing homes. Some initial AHEAD respondents aged 70-74 had prior nursing-home exposure. Some persons died before reaching ages 70-74 and had nursing home exposure. To make these adjustments we use the HRS cohort. Combining HRS waves for respondents 50-54 years of age, we find that 20% died before reaching age 72 and that 75% of those who died were in a nursing home sometime prior to death. Among survivors, 2.5% were in a nursing home at age 72, and among those not in a nursing home, 4% had previously been in a nursing home. Combining these conditional probabilities with the AHEAD probabilities, we estimate that the lifetime exposure of HRS respondents initially ages 50-54 in 1992 will be 62.6% when the last such respondent has died.

These calculations do not consider nursing-home exposure prior to entering HRS. While we have no data on nursing home prior to the initial wave of HRS, in the subsequent waves of HRS nursing home exposure is infrequent. For example, among those initially age 50-54 in HRS in 1992, 0.2% had nursing home exposure between waves 1 and 2, and an additional 0.6% had nursing home exposure between waves 2 and 3.

4. Methodology of parametric estimation of lifetime exposure

---

[3] This calculation assumes that, conditional on being in the community at wave 1 and on reaching age 85-89 with no intervening nursing home exposure, the probability of nursing home exposure prior to death is the same as the probability of nursing home exposure of those in wave 1 age 85-89.

We develop a simulation model which allows us to compute the lifetime distribution of nursing-home stays and their length. Let $i = 1,..., N$ denote respondents and $t = 1,\dots,T_i$ denote the wave during which an interview takes place. Each wave takes place approximately every two years.[4]

We use two key pieces of information from the HRS in building the model. First, we use reports of any nursing home stays in the previous two years and reports of mortality to construct a combined status variable, d_{it}, which can take four values : 1) alive & living in the community, 2) alive & living in a nursing home, 3) died in the community, 4) died in a nursing home. Obviously, states 3 and 4 are absorbing. Hence, four transitions are possible from each of the two states where the respondent is alive (1 and 2). We define the probability of entering state $j = 1,\dots,4$ at $t+1$ given a current state $k = 1, 2$ at t, a vector of socio-demographic characteristics x_i, and age a_{it} using a multinomial logit :

$$P(d_{it+1} = j \mid x_i, a_{it}, d_{i,t} = k) = \frac{\exp(x_i\gamma_{1,j,k} + \gamma_{a,j,k}(a_{it}))}{\sum_{j'}\exp(x_i\gamma_{1,j',k} + \gamma_{a,j,k}(a_{it}))}.$$

We do not impose parametric restrictions on the functions $\gamma_{a,j,k}$ and instead use age dummies. Because data is scarce at older ages, we use 5-year age groups from age 50 to 100. After obtaining estimates of the parameters by maximum likelihood, we interpolate linearly the age functions at single years of age intervals. We extrapolate for ages between 100 and 110 (maximum age in simulations).

Second, we use reports of the number of days spent in a nursing home between waves. Because the time of entry or exit is unknown and could vary on average depending on the state at t and the state at $t+1$, we estimate separate models of the log of number of visits v_{it} between waves for 1) individuals transiting from "living in the community" to either "living [or]

---

[4] In the next draft, we plan to adjust the models presented below to account for heterogeneity in exposure time (differences in time between interviews).

died in a nursing home", and 2) for those transiting from "living in a nursing home" to the same two destinations. The models estimated take the form

$$\log v_{it} = x_i \beta_{j,k} + \beta_{a,j,k}(a_{it}) + \varepsilon_{it}$$
$$j = 2,4, k = 1,2$$

where ε_{it} is assumed normally distributed with mean 0 and variance $\sigma^2_{\varepsilon,j,k}$. Again, we assume the age functions are given by a set of age dummies (5-year age groups). We use interpolation for intervening years. Given the log formulation of the conditional mean and the assumption of normality, the expected number of visits is given by

$$E[v_{it} \mid x_i, a_{it}, j, k] = \exp(x_i \beta_{j,k} + \beta_{a,j,k}(a_{it}) + \tfrac{1}{2}\sigma^2_{\varepsilon,j,k}).$$

The estimated equations for the transition probabilities and the process for the number of days in a nursing home can then be used to simulate histories of nursing home stays. The initial population for the simulation are those respondents 50 to 55 years of age in the War Babies and Early Baby Boomers cohorts. We draw with replacement 50,000 sets of socio-demographic characteristics x_i. We consider education, race, marital status at age 50, number of children, whether the individual had daughters, and an indicator for whether the respondent was ever a smoker. We then simulate histories using the processes estimated above. Finally, we compute statistics of interest from the simulation using survey weights from the two waves used as the starting point.

4. Results for Parametric Model

4.1 Estimation

We first present estimation results of the transition models. We then present results for the number of days spent in a nursing home between waves. We obtain both sets of estimates using the HRS data as described in section 2 above.

In Table 6, we present multinomial logit estimates of transitions from the community to a nursing home. The reference category is living in the community (i.e. not in a nursing home). Since all age parameters go from negative to positive, this establishes that the fraction alive in the community decreases with age. Not surprisingly, transitions either to a nursing home or death increase in frequency with age. This is shown in Figure 5, first panel, where we see that the average probability of staying in the community decreases from close to 100% for a 50-year-old respondent to less than 50% for a 95-year-old respondent. Before age 65, most transitions out of the community are the result of death outside of nursing homes or residence in a nursing home. After age 75, many more respondents die in a nursing home. The transition rate from the community to a nursing home increases steeply after age 70.

Results from Table 6 indicate that males face transition probabilities significantly different from those females face. Compared to females, males who had been living in the community have a much lower chance of living in a nursing home two years (i.e. one survey wave) later. They also have a significantly higher risk of dying from one wave to the next, reflecting males' lower life expectancy – and, accordingly, higher death rates. This increased death probability is tilted toward death in the community: males are more likely to die in either setting, but more than twice as likely to die outside a nursing home.

Table 6 also shows that education, in particular college education, protects against both mortality and entering a nursing home. This reflects in part the SES-health gradient. Non-white respondents are more likely to die outside a nursing home but are about as likely as white respondents to enter a nursing home. Being married at age 50 also protects against entering a nursing home. Of course, the natural channel for this association is that one spouse may be able to provide help to the other who needs it. In addition, those married at age 50 are less likely to die, either in a nursing home or in the community. Interestingly, not having children does not appear to increase the probability of entering a nursing home (relative to having 1-3 children). On the other hand, having more children (4 or more) appears to increase the probability that the respondent will die outside a nursing home, relative to dying in a nursing home. Having daughters, conditional on the number of children, appears to reduce the probability of dying in a nursing home.

In Table 7, we present multinomial logit estimates of transitions from a nursing home to one of the four states (living in the community, living in a nursing home, dying in the community, dying in a nursing home). The reference category is alive in a nursing home. From the estimates, we see that the exit probability from a nursing home to the community generally decreases with age. In Figure 5, second panel, we see that the estimates imply that this probability goes from 80% at age 50 to less than 10% at age 85. Hence, persistence increases with age, which likely reflects an increase in the severity of disabilities for those in a nursing home. Mortality rates in a nursing home are much higher than in the community. This can be seen in Figure 5. For example, summing the two curves, the probability of dying either in the community or in a nursing home is 8% at age 50, while it increases to more than 57% by age 95.

Table 7 indicates that gender differences are somewhat different for transitions originating in a nursing home. Specifically, and in contrast to the previous state of origin, males do not face a significantly different probability of exiting a nursing home alive. As well, their probability of dying between the two waves is much higher than females', with the difference being tilted again toward dying outside a nursing home. Gender differences are, however, markedly larger than they were for transitions originating in the community. This could perhaps be explained by a more severe disability for males when they finally enter a nursing home.

As for the other characteristics considered and likely to affect transitions out of nursing homes, college education appears to increase the likelihood of return to the community. Education also decreases the probability of dying in a nursing home, as was the case for those individuals who were initially living in the community (Table 6). Being non-white reduces the probability of exiting a nursing home. Family background has no significant effect on exits from nursing homes, excepting those individuals who were married at age 50, who are more likely to leave a nursing home for the community. Finally, being a smoker reduces the probability of returning to the community.

Estimates in Tables 6 and 7 will play out in the simulations. On one hand, some characteristics such as education reduce the probability of entry into a nursing home and increase the exit probability from a nursing home. On the other hand, education also reduces the likelihood of dying, hence prolonging the exposure to nursing home risk. In the end, these opposing forces will yield ambiguous predictions of the effect of education on lifetime prevalence of nursing home stays.

We also look at the intensity of nursing home stays between waves. For this, we turn to Table 8 which reports estimation results for the number of days spent in nursing homes. Each column in the table reports estimates of the effect of variables on the log of the number of days spent in a nursing home for four different pairs of origin and destination states.

The first and second columns present results for transitions that originate from the community. In the first column, respondents are observed living in a nursing home at t+1, while in the second they died in a nursing home by t+1. We see for both these transitions an increase with age in the number of days spent in a nursing home. Figure 6 reports the average expected number of days spent in a nursing home by age. We can observe for both these transitions a small number of days (from roughly 100 days at age 50 to 350 days at age 95) relative to other transitions which originate from nursing homes (see below). This partly reflects the fact that respondents experiencing the two transitions originating in the community will, on average, enter nursing homes in the middle of the time interval between waves. The average number of days is also smaller for those who died between waves. For these two transitions originating in the community, college education and having daughters or at least four children all appear to reduce the number of days spent in a nursing home, whereas being non-white and ever being a smoker appear to increase it.

The third and fourth columns of Table 8 report regression results for transitions that originate from nursing homes. Not surprisingly, the number of days spent in a nursing home is much higher for those respondents, as shown in Figure 6. Furthermore, the number of days increases with age (except for individuals staying in a nursing home at younger ages, for whom the number is initially higher). In terms of characteristics, having a college degree and having daughters both decrease the number of days spent in a nursing home for individuals staying in a nursing home across waves. On the other hand, being non-white increases this number of days.

4.2 Simulation

We simulate the lifetime nursing-home histories for an initial population with characteristics drawn from the pool of respondents 50 to 55 years of age in the War Babies and Early Baby Boomers cohorts. We draw (with replacement) 50,000 observations. We then simulate two-year transitions up to a maximum of 110 years of age.

In Table 9, we present simulation results for 5 outcomes: the simulated probability of any stay, the average number of days in a nursing home (both conditional on having at least one stay and unconditional), the probability of dying in a nursing home, and, finally, the age at which individuals first enter a nursing home. We report these outcomes under two scenarios. The first includes the information from the exit interviews. Hence, we know whether – and when – someone died in a nursing home. For the second scenario we ignore the information from the exit interviews, so that nursing home stays prior to death are unobserved. Hence, we reclassify state $j = 4$ ("Died this wave, in nursing home last 2 years") when computing outcomes to state $j = 3$ ("Died this wave, no nursing home stay last 2 years"). So all individuals who died will be attributed state $j = 3$.

Results are striking, especially when comparing the two scenarios (i.e. with and without exit interviews). When excluding exit interviews, the probability of ever experiencing a nursing home stay is 37.4%, which is consistent with prior literature (Kemper and Murtaugh, 1991).

When using exit interviews however, this probability increases to 58.5%. The average number of days spent in a nursing home increases from 137.8 to 215.2. Conditional on having one stay, the average number of days spent over the lifetime of a 50 year old is 367.6 days, or just over a year. The probability of dying in a nursing home is 48.2%. Finally, the average age at first entry in a nursing home is 76.8 when using exit interviews, and 75.9 without. Overall, exit interviews are crucial in accurately establishing the lifetime prevalence and intensity of nursing home stays.

In Table 10, we present Table 9 results by socio-demographic groups. The lifetime prevalence of nursing home stays differs considerably by gender. Females face a 65.7% probability of having at least one stay, compared to 50.6% for males. This is due to both their longer life expectancy and to the fact that at every age, they face a higher probability of entering a nursing home (perhaps because their husband died before them). Females' average number of days spent in a nursing home is 287.5, compared to 135.9 for males. The probability that females will die in a nursing home is 53.8%, against 42.0% for males. Non-whites have a lower probability of entering a nursing home, in part because of their lower life expectancy.

Differences in terms of education are ambiguous, as predicted above. Individuals with less than a high school diploma have a lower risk of ever entering a nursing home. Most of this difference is explained by their lower life expectancy (74.6 years, compared to 78.0 for those with a high-school diploma and 79.5 for those with a college degree). At the same time, those with college degrees have a slightly lower lifetime exposure to nursing homes, driven in part by a delayed age of entry into nursing homes. Interestingly, individuals without children have a lower risk of ever entering a nursing home. Those who were a smoker at some point in their life are much less likely to ever enter a nursing home, but this effect is also driven by a lower life expectancy (76.1 compared to 81.5 for non-smokers). We find little differences in lifetime prevalence by marital status or according to the presence of daughters.

Because some of the characteristics considered are correlated, the question arises as to whether differences remain over one characteristic when controlling for the others. Hence, we regress the simulated lifetime outcomes from age 50 on individuals' socio-demographic characteristics. This may be interpreted as a reduced-form equation of the transition and intensity equations above. Results are presented in Table 11.

Results reveal that all else equal, males have a 13.6% lower probability of ever entering a nursing home, and also spend 58.7% less time in a nursing home (58.7% less). Education effects are mostly driven by the distinction between having completed high school or not. Individuals with either high school or college degrees have a 8% higher chance of ever staying in a nursing home. Non-whites, individuals without children or with many children, and those who were ever smokers are also less likely to enter a nursing home at any point in their life.

5. Conclusion

In this paper, we use both parametric and non-parametric approaches to calculate the lifetime risk of nursing home stays using rich data from the Health and Retirement Study. Both provide a similar estimate: a 50 year old has roughly a 60% chance of ever staying in a nursing home and conditional on entering the average number of nights spent in a nursing home over the lifetime is just over a year (370 days). This average estimate hides considerable heterogeneity as the distribution is highly skewed.

Our estimates of lifetime risk are larger than those previously reported. One reason is that we account for short-term stays at the end of life using exit interviews in the HRS. In fact, excluding that information, we get an estimate of 37% which is similar to reported estimates in the literature. We also highlight that the estimation of nursing home risk in a survey like the HRS has to take into account that the population is initially does not include nursing home residents since only non-institutionalized respondents are sampled. We take advantage of the fact that stays typically last less than 5 years, and discard the first few waves of each cohort. This allows to sample from a stable nursing home as well as non-nursing home population.

Our results also highlight that there are two competing forces that affect lifetime risk: nursing home risk and mortality risk. Both of these depend in a non-trivial way on socio-demographic characteristics. For example, smokers have a higher risk of entering a nursing home conditional on being alive. But since they also face higher mortality risks, this reduces lifetime exposure to nursing home risk. We find that females, white and non-smokers face the highest risks of ever entering a nursing home.

A number of extensions are worth exploring. First, we focused on fixed socio-demographic characteristics in the simulation model. It would be interesting to model the dynamics of marital status and activity of daily living limitations: their evolution likely affects nursing home risk. Second, we have not dealt with unobserved heterogeneity. This can impact the dynamics of nursing home risk and estimates of lifetime risk in important ways. Third, we can integrate estimates of out-of-pocket nursing home expenses in our estimates in order to assess the financial risks associated with nursing home stays.

References

Arling G, Hagan S, Buhaug H. The feasibility of a public-private long-term care financing plan. Med Care 1992; 30:699.

Aykan, Hakan, "Effect of Childlessness on Nursing Home and Home Health Care Use," Journal of Aging and Social Policy 15:1, p33-53 [2003]

Brown, Jeffrey R., and Amy Finkelstein. The Interaction of Public and Private Insurance: Medicaid and the Long-Term Care Insurance Market. American Economic Review 2008; 98(3): 1083-1102.

Cohen MA, Tell EJ, Wallack SS. The lifetime risks and costs of nursing home use among the elderly. Medical Care 1986; 41(6):785-792.

Dick A, Garber AM, MaCurdy T. Forecasting Nursing Home Utilization of Elderly Americans, in Studies in the Economics of Aging, ed. David Wise, University of Chicago Press, 1994, pp. 365 - 394

Kelly, Anne; Conell-Price, Jessamyn; Covinsky, Kenneth; Cenzer, Irena; Chang, Anna; Boscardin, W. J.; Smith, Alexander, "Length of Stay for Older Adults Residing in Nursing Homes at the End of Life," Journal of the American Geriatrics Society 58:9, p1701-1706 [2010]

Kemper P and Murtaugh CM. Lifetime Use of Nursing Home Care. NEJM 1991; 324:595-600.

Liang J, Liu X, Tu E, Whitelaw N. Probabilities and Lifetime Durations of Short-Stay Hospital and Nursing Home Use in the United States, 1985. Medical Care 1996; 34(10):1018-1036.

Robinson, Jim. 1996. “A Long-Term-Care Status Transition Model”. Society of Actuaries.

Spillman BC and Lubitz J. New estimates of lifetime nursing home use: have patterns of use changed? Medical Care 2002; 40(10):965-975.

Table 1: Cohorts included in the HRS defined by birth year

	AHEAD	CODA	HRS	War Babies	Early Baby Boomers
Birth years	1890-1923	1924-1930	1931-1941	1942-1947	1948-1953
Baseline interview	1993	1998	1992	1998	2004
Age at baseline	70 or older	68-74	51-61	51-56	51-56
Age in 2010 if still alive	87 or older	80-86	69-79	63-68	57-62
N	7,758	4,210	10,413	3,488	3,624
Overall status as of wave 10 (%)					
Alive and in survey as of wave 10	14.48	46.82	57.23	72.28	78.84
Dead with exit interview	76.31	38.86	21.84	9.55	4.39
Dead without exit interview	3.48	2.47	3.25	1.18	0.39
Alive, but no response to wave 10	2.23	7.17	8.39	11.47	14.40
Unknown - dropped from sample	3.51	4.68	9.30	5.53	1.99

Notes:
AHEAD stands for Asset and Health Dynamics, the initial name of the 1993 survey.
CODA stands for Children of the Depression.

Table 2. Prevalence of nursing home stay by age. At time of interview and in previous two years. Weighted.

	All cohorts					
	Ns			Fraction		
Age of person including their wave of death	Live in Nursing Home at Interview - CORE only	Nursing Home stay, previous 2 years	Nursing Home stay, previous 2 years - CORE only	Live in Nursing Home at Interview - CORE only	Nursing Home stay, previous 2 years	Nursing Home stay, previous 2 years - CORE only
<55	769	807	768	0.001	0.007	0.001
55-59	16,886	17,180	16,855	0.001	0.008	0.006
60-64	24,719	25,299	24,688	0.003	0.014	0.010
65-69	20,551	21,306	20,526	0.007	0.027	0.019
70-74	14,313	15,168	14,290	0.014	0.057	0.038
75-79	14,458	15,759	14,435	0.026	0.083	0.056
80-84	10,889	12,458	10,867	0.053	0.151	0.103
85-89	6,651	8,240	6,637	0.124	0.268	0.195
90-94	2,746	3,876	2,740	0.226	0.420	0.306
95+	797	1,336	793	0.361	0.567	0.445
All	112,779	121,429	112,599	0.025	0.073	0.046

Source: Authors' calculations.

Note: Data from all HRS cohorts used, except the cohort added in HRS 2010. For each cohort the first two waves are excluded in keeping with the previous finding that population representation of the population, including the nursing home population is not achieved until the third wave.

"previous 2 years" refers to time since last interview, which for most respondents is approximately two years.

"CORE only" excludes exit interviews.

Table 3. Number of nights spent in nursing home, by age. All Cohorts. Weighted.

Age of person including their wave of death	All cohorts			
	Ns		Mean	
	Nights in nursing home, previous 2 years	Nights in nursing home, previous 2 years - CORE only	Nights in nursing home, previous 2 years	Nights in nursing home, previous 2 years - CORE only
<55	807	768	0.71	0.60
55-59	17,171	16,849	0.87	0.73
60-64	25,278	24,679	2.32	1.65
65-69	21,276	20,511	4.22	3.60
70-74	15,141	14,273	8.95	6.98
75-79	15,700	14,404	15.88	12.65
80-84	12,370	10,818	33.73	26.99
85-89	8,127	6,567	66.43	57.63
90-94	3,788	2,687	130.45	117.08
95+	1,312	780	203.96	204.75
All	120,970	112,336	17.07	12.52

Note: “CORE only” excludes exit interviews.

“previous 2 years” refers to time since last interview, which for most respondents is approximately two years.

Table 4. HRS cohort. Mortality and cumulative nursing home frequencies, unconditional and conditional on dying between waves 2 and 10

Age wave 1	N	Died waves 2-10	Any stay 2-10	Number stays	Number nights
50-54	3852	0.21	0.10	0.18	24.1
55-59	4122	0.26	0.14	0.31	38.5
60-64	1336	0.34	0.18	0.37	59.7
All	9310	0.25	0.13	0.26	35.6
Conditional on dying in between waves 2 and 10					
50-54	791	1.00	0.27	0.56	81.4
55-59	1059	1.00	0.30	0.67	79.3
60-64	449	1.00	0.30	0.72	110.4
All	2299	1.00	0.29	0.64	86.1

Note: Includes exit interviews.

Table 5. AHEAD cohort. Mortality and cumulative nursing home frequencies, unconditional and conditional on dying between waves 2 and 9

Age wave 1	N	Died waves 2-9	Any stay 2-9	Number stays	Number nights
70-74	2676	0.67	0.41	0.84	148.2
75-79	2031	0.83	0.54	1.14	247.4
80-84	1493	0.94	0.59	1.31	304.4
85-89	723	0.99	0.64	1.51	343.7
90-94	231	0.99	0.63	1.20	299.7
95+	68	1.00	0.56	0.97	299.9
All	7239	0.82	0.52	1.10	233.9
Conditional on dying between waves 2 and 9					
70-74	1802	1.00	0.48	0.98	162.5
75-79	1678	1.00	0.57	1.21	253.1
80-84	1400	1.00	0.61	1.35	308.0
85-89	715	1.00	0.64	1.52	342.8
90-94	229	1.00	0.64	1.21	302.4
95+	68	1.00	0.56	0.97	299.9
All	5905	1.00	0.56	1.21	251.5

Note: Includes exit interviews.

Table 6. Multinomial Logit Estimates: Transition from Community

Variables	Alive in NH	Died, not NH	Died in NH
age (65 omitted)			
50	-1.921***	-1.210***	-1.832***
	(0.177)	(0.106)	(0.261)
55	-1.296***	-1.083***	-1.431***
	(0.114)	(0.084)	(0.180)
60	-0.593***	-0.488***	-0.608***
	(0.103)	(0.078)	(0.158)
70	0.691***	0.250***	0.832***
	(0.099)	(0.085)	(0.146)
75	1.183***	0.773***	1.343***
	(0.088)	(0.073)	(0.130)
80	1.820***	1.156***	1.991***
	(0.086)	(0.075)	(0.127)
85	2.403***	1.675***	2.883***
	(0.091)	(0.083)	(0.129)
90	2.800***	2.258***	3.737***
	(0.116)	(0.112)	(0.144)
95	2.791***	3.029***	4.221***
	(0.239)	(0.199)	(0.225)
Male	-0.459***	0.364***	0.154***
	(0.047)	(0.041)	(0.060)
education (less than high school omitted)			
high school	-0.0948*	-0.402***	-0.146**
	(0.053)	(0.049)	(0.069)
College	-0.180***	-0.511***	-0.413***
	(0.055)	(0.049)	(0.073)
non-white	-0.0973	0.240***	-0.00666
	(0.072)	(0.057)	(0.096)
married at age 50	-0.333***	-0.250***	-0.290***
	(0.053)	(0.051)	(0.073)
number of children (1-3 omitted)			
no children	-0.397	0.079	-0.109
	(0.262)	(0.209)	(0.280)
4+ children	-0.0287	0.0878**	-0.156**
	(0.046)	(0.042)	(0.063)
has daughters	0.00874	-0.0348	-0.139*
	(0.060)	(0.057)	(0.075)
ever smoker	0.241***	0.649***	0.477***
	(0.044)	(0.045)	(0.062)
Constant	-3.624***	-3.661***	-4.665***
	(0.111)	(0.099)	(0.158)
Observations	104,799		
Log-likelihood	-27093		
degrees freedom	54		
chi-square	9910		

Notes : Multinomial logit parameter estimates and standard errors, *** p<0.01, ** p<0.05, * p<0.1. The base category is living in the community. NH refers to nursing homes.

Table 7. Multinomial Logit Estimates : Transition from Nursing Homes (NH)

Variables	alive, not NH	died, not NH	died in NH
age (65 omitted)			
50	1.611**	0.189	-0.885
	(0.688)	(1.512)	(1.292)
55	0.652**	-0.15	-1.211**
	(0.289)	(0.628)	(0.505)
60	0.202	-0.211	-1.138***
	(0.234)	(0.496)	(0.353)
70	0.0224	-0.0265	-0.113
	(0.214)	(0.430)	(0.254)
75	-0.185	0.421	-0.0514
	(0.199)	(0.373)	(0.231)
80	-0.371*	0.195	0.529**
	(0.191)	(0.367)	(0.211)
85	-1.108***	0.211	0.470**
	(0.197)	(0.361)	(0.209)
90	-1.513***	0.231	0.876***
	(0.246)	(0.402)	(0.225)
95	-2.656***	0.178	0.857***
	(0.500)	(0.520)	(0.274)
Male	0.163	0.743***	0.450***
	(0.109)	(0.176)	(0.108)
education (less than high school omitted)			
high school	0.266**	-0.322	-0.245**
	(0.120)	(0.200)	(0.111)
College	0.508***	-0.0491	-0.289**
	(0.124)	(0.198)	(0.117)
non-white	-0.308*	0.146	0.0255
	(0.165)	(0.258)	(0.160)
married at age 50	0.304**	0.276	0.0672
	(0.119)	(0.211)	(0.115)
number of children (1-3 omitted)			
no children	-0.674	0.332	0.204
	(0.645)	(0.855)	(0.519)
4+ children	-0.00534	-0.00014	0.0594
	(0.105)	(0.179)	(0.104)
Daughter	0.138	0.199	0.126
	(0.135)	(0.232)	(0.126)
ever smoker	-0.310***	0.0831	0.12
	(0.102)	(0.177)	(0.100)
Constant	-0.119	-2.451***	-0.632**
	(0.240)	(0.443)	(0.249)
Observations	3,086		
log-likelihood	-3610		
degrees freedom	54		
Chi-square	555.5		

Notes : Multinomial logit parameter estimates and standard errors, *** p<0.01, ** p<0.05, * p<0.1. The base category is living in a nursing home.

Table 8. Regression Estimates for # of days in Nursing Homes (NH)

	from community		from nursing home	
Variables	living in NH	died in NH	living in NH	died in NH
age (65 omitted)				
50	-0.308	-0.502	0.946	-0.768
	(0.319)	(0.477)	(0.795)	(0.957)
55	-0.0242	-0.116	0.333	-0.556
	(0.186)	(0.275)	(0.368)	(0.473)
60	-0.233	-0.138	0.0803	0.351
	(0.158)	(0.219)	(0.268)	(0.345)
70	0.025	-0.0373	0.465**	0.00768
	(0.138)	(0.191)	(0.225)	(0.238)
75	0.221*	0.0737	0.433**	0.174
	(0.126)	(0.174)	(0.205)	(0.205)
80	0.487***	0.473***	0.633***	0.00344
	(0.122)	(0.169)	(0.197)	(0.189)
85	0.694***	0.512***	0.828***	0.233
	(0.128)	(0.168)	(0.194)	(0.186)
90	1.046***	0.438**	0.857***	0.348*
	(0.160)	(0.182)	(0.213)	(0.190)
95	0.922***	0.412	1.234***	0.507**
	(0.313)	(0.257)	(0.261)	(0.221)
male	-0.0905	-0.190**	-0.0724	-0.126
	(0.071)	(0.084)	(0.107)	(0.087)
education (less than high school omitted)				
high school	-0.0266	-0.131	-0.147	-0.1
	(0.079)	(0.095)	(0.105)	(0.091)
college	-0.206**	-0.184*	-0.321***	-0.108
	(0.081)	(0.101)	(0.113)	(0.099)
non-white	0.442***	0.00324	0.480***	0.0821
	(0.093)	(0.116)	(0.127)	(0.114)
married at age 50	-0.0993	0.0943	0.00698	-0.176*
	(0.081)	(0.101)	(0.104)	(0.093)
number of children (omitted 1-3)				
no children	-0.366	-0.0999	0.388	-0.479
	(0.359)	(0.328)	(0.530)	(0.392)
4+ children	-0.222***	0.0688	0.0623	-0.119
	(0.070)	(0.090)	(0.099)	(0.088)
daughter	-0.224**	-0.207*	-0.260**	-0.099
	(0.092)	(0.108)	(0.120)	(0.101)
ever smoker	0.114*	-0.0457	-0.0938	-0.0432
	(0.068)	(0.088)	(0.095)	(0.083)
Constant	3.829***	3.491***	5.283***	5.207***
	(0.162)	(0.204)	(0.231)	(0.214)
Observations	2,579	1,478	1,272	1,226
log-likelihood	-4882	-2694	-2349	-2074
degrees of freedom	18	18	18	18

R-square	0.0662	0.0398	0.0608	0.034

Notes : Regression of the log of number of days in nursing homes between interviews. Each column refers to a different specification, estimated on the sample of respondents either living in the community or in a nursing home at t and who either survive in a nursing home at t+1 or die in a nursing home by t+1. Standard errors below point estimates, *** p<0.01, ** p<0.05, * p<0.1.

Table 9. Simulated Lifetime Exposure to Nursing Homes (NH) With and Without Exit Interviews

	With Exit Interviews	Without Exit Interviews
Prob(any stay)	0.585	0.374
E[# days]	215.2	137.8
E[# days| any stay]	367.6	368.2
Prob(Dies in NH)	0.482	
Age first entry NH	76.8	75.9

Notes : Simulated outcomes from age 50, for 50,000 respondents drawn randomly from the pool of HRS respondents aged 50-55 in the War Babies and Early Baby Boomers cohorts. Sample weights used.

Table 10. Simulated Lifetime Exposure by Socio-demographic Characteristics

Sub-groups of population	Prob(any stay)	E[# days]	Prob(Dies in NH)	E[Age first entry NH]	E[Age of death]
Gender					
Female	0.657	287.5	0.538	76.7	79.5
Male	0.506	135.9	0.42	76.9	77.0
Race					
White	0.600	215.4	0.495	76.9	78.7
Non-white	0.506	214.4	0.414	76.0	76.4
Education					
Less than high school	0.484	185.1	0.397	74.5	74.6
High school	0.61	233.8	0.516	76.3	78.0
College	0.597	210.7	0.482	77.7	79.5
Number of children					
No children	0.512	161.2	0.427	78.0	77.7
1-3 children	0.602	223.6	0.497	76.9	78.6
4+ children	0.551	196.1	0.446	76.9	77.6
Ever smoker					
No	0.651	259.5	0.541	78.7	81.5
Yes	0.539	184.6	0.442	75.1	76.1
Married at age 50					
No	0.586	253.2	0.487	74.7	76.3
Yes	0.582	204.6	0.481	77.4	78.9
Has daughters					
No	0.592	243.4	0.501	76.5	78.3
Yes	0.583	209.4	0.478	76.9	78.3
Total	0.595	219.7	0.489	76.9	78.4

Notes : Simulated outcomes from age 50, for 50,000 respondents drawn randomly from the pool of HRS respondents aged 50-55 in the War Babies and Early Baby Boomers cohorts. Sample weights used.

Table 11. Regression of Lifetime Measures on Socio-demographic Characteristics

variables	P(any stay)	E[log days|any stay]
male	-0.136***	-0.587***
	(0.004)	(0.022)
education (less than high school omitted)		
high school	0.0858***	-0.0283
	(0.006)	(0.031)
college	0.0726***	-0.0423
	(0.006)	(0.031)
non-white	-0.0765***	0.0727***
	(0.005)	(0.027)
married at age 50	-0.00302	-0.131***
	(0.006)	(0.027)
number of children (1-3 omitted)		
no children	-0.0531***	-0.325***
	(0.020)	(0.105)
4+ children	-0.0313***	-0.028
	(0.005)	(0.023)
has daughters	0.00491	-0.127***
	(0.006)	(0.029)
ever smoker	-0.0891***	-0.131***
	(0.004)	(0.021)
Constant		5.337***
		(0.045)
Observations	50,000	29,152

Notes : Simulated outcomes from age 50, for 50,000 respondents drawn randomly from the pool of HRS respondents age 50-55 in the War Babies and Early Baby Boomers cohorts, are regressed on other observed characteristics. The first model for the probability of any stay in a nursing home is estimated using a logit model and average marginal effects on the probability of any stay are reported. In the second model, the log of the number of days spent in a nursing home, conditional on having at least one stay, is regressed on the same characteristics. Standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figures

Figure 1. Prevalence of respondent being in nursing home at the time of interview, by age and wave. AHEAD cohort

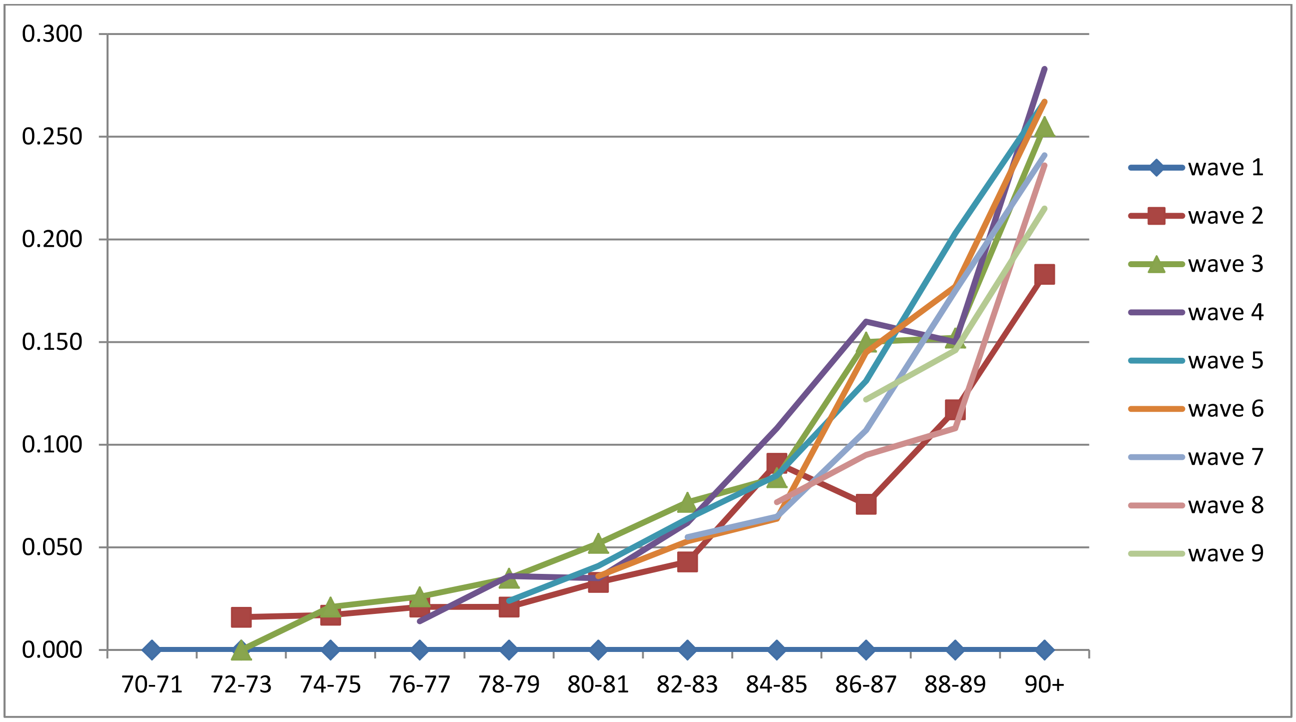


**Figure 2. Fraction with any nursing home stay in last two years, by age.
AHEAD cohort. No exit interviews**

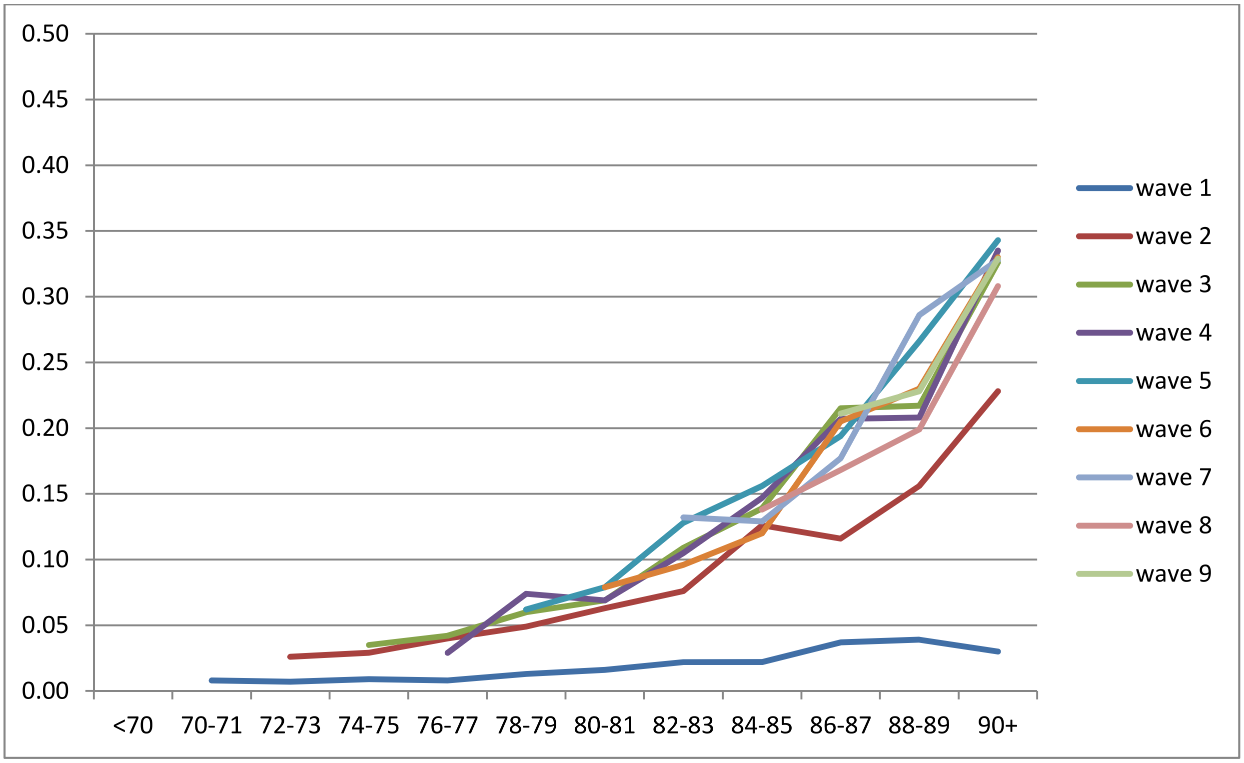


**Figure 3. Fraction with any nursing home stay in last two years, by age.
AHEAD cohort. With exit interviews.**

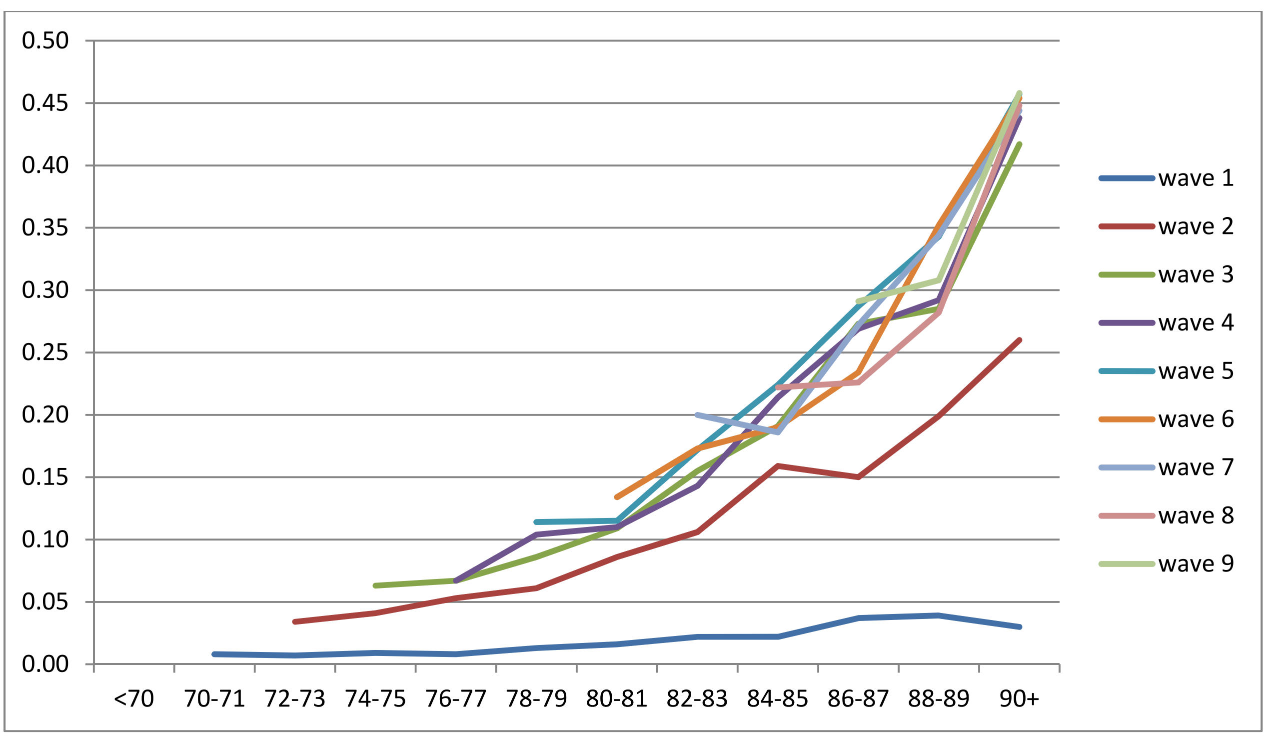
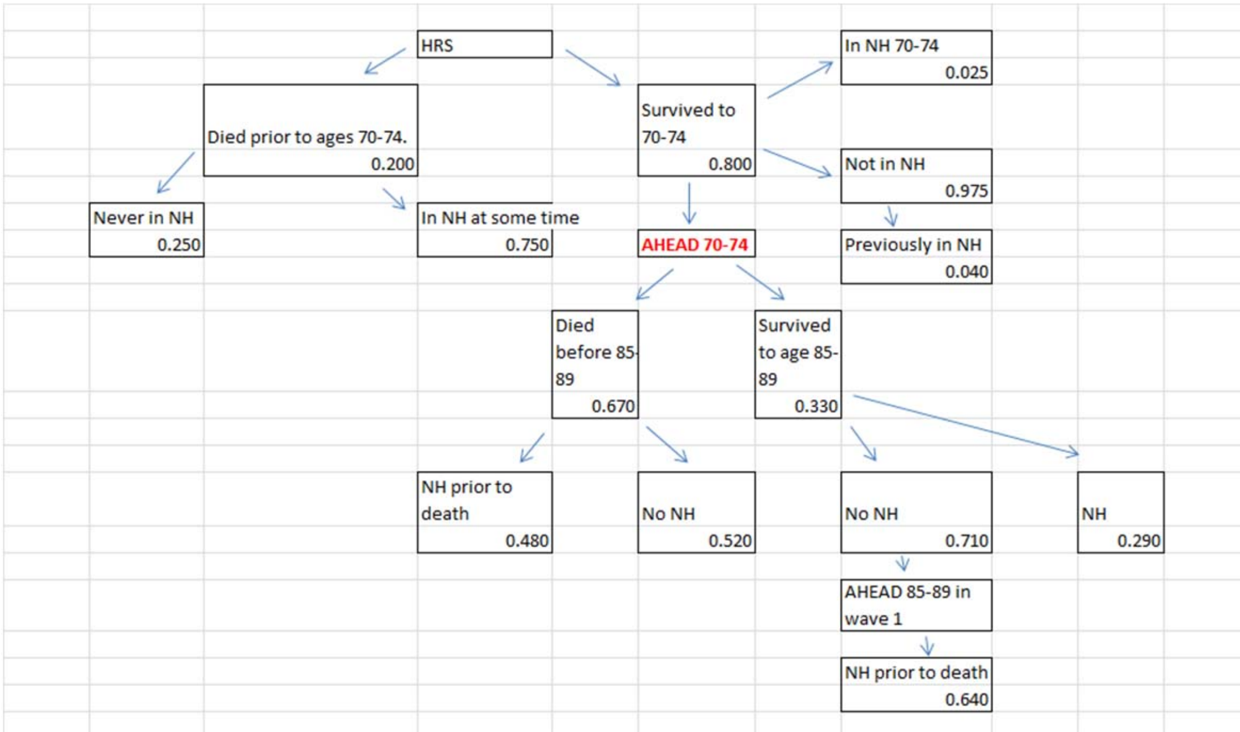


Figure 4. Unconditional probabilities of nursing home exposure and transitions



Note: NH = nursing home

Figure 5. Transition Probabilities across States

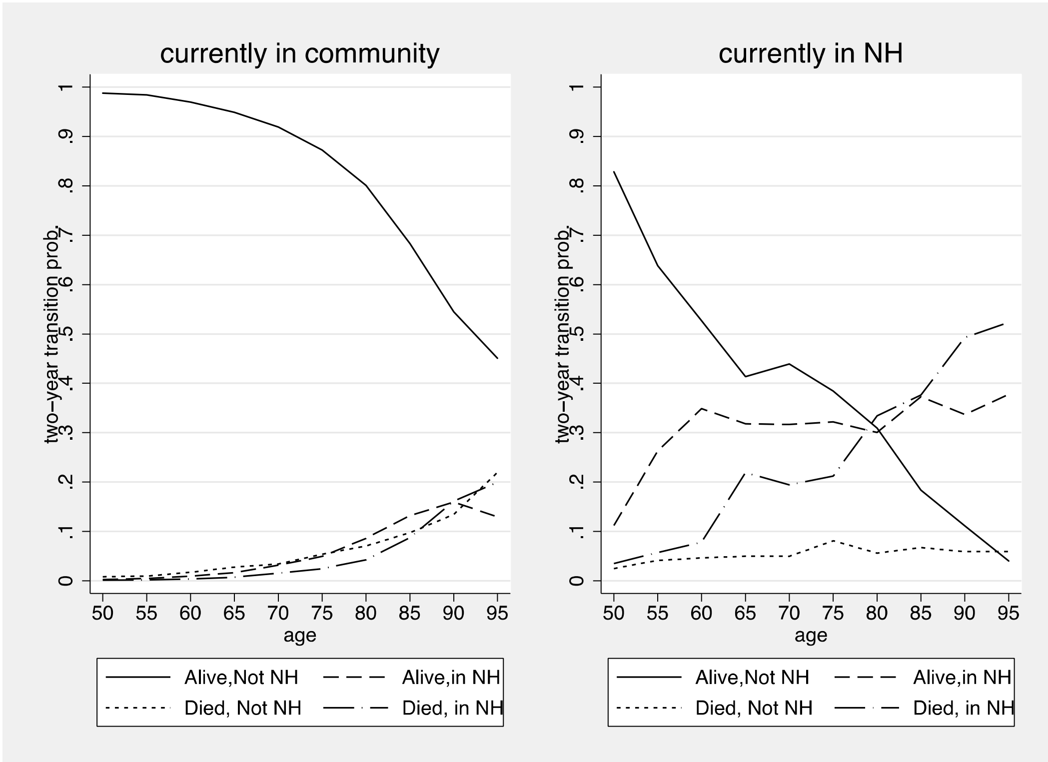


Figure 6. Expected Number of Days in NH Conditional on Being in NH Next Wave