10 Forecasting Nursing Home Utilization of Elderly Americans

Andrew Dick, Alan M. Garber, and Thomas A. MaCurdy

Nursing home care accounts for most expenditures for long-term care, yet little is known about the risk that the elderly face of having a prolonged nursing home stay. The issue is of growing and fundamental importance: the design of policies to insure against the expenditures associated with institutionalization depends on the distribution of nursing home utilization. It is not sufficient to know the expected value of nursing home utilization, because many insurance plans have a fixed ceiling on the nursing home benefit and their risk of paying the maximum benefit is the likelihood that an enrollee's utilization will exceed the ceiling. Similarly, for public-private partnerships, in which private insurers might cover the first 12 or 24 months of nursing home care, while the federal government provides coverage above the private limit, the expenditures borne by insurers and by the federal government depend on the exact distribution of utilization. Insurers are not the only ones with an interest in nursing home utilization: before individuals can make an informed decision about purchasing long-term care insurance, saving to cover the costs of institutionalization, or changing living arrangements, they need to know how much nursing home care they are likely to use. To address these issues, this paper describes the likelihood and duration of nursing home admissions experienced by Americans after age 65. Our analysis generates predictions for a representative popu-

---

Andrew Dick is an instructor in the Public Policy Analysis Program in the Department of Political Science at the University of Rochester. Alan M. Garber is HSR&D Senior Research Associate of the Palo Alto Department of Veterans Affairs, associate professor of medicine at Stanford University, and a research associate of the National Bureau of Economic Research. Thomas A. MaCurdy is professor of economics at Stanford University and a research associate of the National Bureau of Economic Research.

This research was supported in part by grants R29 AG07651 and P01 AG05842 from the National Institute on Aging and by grant 12671 from the Robert Wood Johnson Foundation. Andrew Dick was supported by grant T32 HS00028 from the Agency for Health Care Policy and Research.

lation, not for one selected to be at high risk of institutionalization, and should therefore be of direct relevance to the design of long-term care policies.

The existing literature contains remarkably little information about either the number of times older Americans are admitted to nursing homes or their cumulative utilization of nursing home care. Much of the literature on the risk of institutionalization, for example, has used a variety of methods to assess the probability that an individual at a given age will enter a nursing home before dying (Palmore 1976; Vicente, Wiley, and Carrington 1979; Wingard, Jones, and Kaplan 1987; Branch and Jette 1982; McConnel 1984; Murtaugh, Kemper, and Spillman 1990). Other studies explore the duration of nursing home admissions (Keeler, Kane, and Solomon 1981; Lewis, Cretin, and Kane 1985; Spence and Wiener 1990), but not the number of admissions or overall utilization. The few studies that analyze overall utilization typically are based on nonrandom samples of the elderly (Liu, Coughlin, and McBride 1991).

Failure to address cumulative utilization in representative samples undoubtedly reflects the inadequacy of much existing data. No individual data sets have sufficiently complete longitudinal data to infer comprehensive measures of nursing home utilization for a nationally representative, random sample of elderly Americans. Ideally such a study would enroll a large number of elderly individuals, track them for several years, and obtain complete information on the number and timing of nursing home stays during the period of observation. Data sets such as the Longitudinal Study on Aging fail to track admission and discharge dates for every nursing home stay and have no means of determining the total number of days an individual spends in a nursing home. Data sources on institutional populations typically have good information about the duration of nursing home stays but cannot be used to determine either the risk of admission or the time subsequently spent in the community.

Perhaps the most natural method for determining patterns of nursing home utilization after a given age is to obtain complete retrospective data on prior nursing home use among a random sample of the elderly who are above that given age at the time of their death. However, it is very difficult to obtain such data, since it is rare to have both a representative sample and complete retrospective data. In the most prominent study of this kind, Kemper and Murtaugh (1991) used the 1986 National Mortality Followback Survey of next of kin to assess prior nursing home utilization of a sample of decedents. They employed a weighting scheme to reconstruct a random sample of the national population. Unfortunately, next of kin's recollections were the sole basis for estimating lifetime nursing home use. The information was recorded as one of four categories of use (less than 3 months, 3–12 months, 1–5 years, and more than 5 years). The memories of next of kin about prior utilization, particularly if it began several years before death or if the decedent had multiple nursing home admissions, may well be inaccurate. Furthermore, the study did not estimate number of admissions and other aspects of patterns of use. Thus, this otherwise well-designed study suffered from important data limitations.

The length of the period during which a person is at risk of entering a nursing home poses a related challenge to prospective data. A person who reaches age 65 faces a gradually increasing risk of entering a nursing home and may be at risk for 30 years. Yet most of the longitudinal data sources cover only one- to two-year windows. Retrospective studies based on time at death do not face this challenge, but obtaining accurate retrospective data for a period of many years, particularly for very old decedents, is often impossible.

Because of these challenges, there is no single data source for estimating the distribution of utilization of nursing home care for a random sample of Americans after the age of 65. In this paper, we overcome these challenges by combining data from two major data sets based on nationally representative samples of the elderly. To exploit the information from these data sources, which are longitudinal but cover short periods (a two-year window), we develop a transition probability model and an empirical framework that allows us to infer the relevant measures of utilization.

The empirical model is designed specifically to forecast the distribution of utilization for a cohort of 65-year-olds until their deaths, although with trivial changes it can be used to forecast utilization for any older cohort. Our interest is not only in predicting total nursing home utilization, but also the pattern of utilization; thus we estimate the likelihood of multiple admissions and determine the distribution of both the durations of nursing home admissions and the intervening episodes of community residence.

The transition probability model focuses on the two states that any living study subject can occupy at any time: community residence or nursing home residence. Our model treats hospital admissions as part of community residence. A spell in the community can terminate either in admission to the nursing home or in death; similarly, a nursing home admission can end either by discharge to the community or by death. Our model builds on transition probabilities, which are the monthly probabilities of a change in status, and exit probabilities, which are the conditional probabilities that if an exit occurs it will be to a particular state (death or change of residence). We account for demographic characteristics, for history of prior nursing home utilization, and for duration effects. We measure durations in months after the start of spells, with changes in age (rather than time) as the relevant metric. We thereby produce estimates capable of forecasting transitions at all ages 65 years and older, using only the data available from a two-year observation window provided by our major data sources. The estimated model enables us to perform simulations that characterize several aspects of nursing home utilization.

In the following section, we discuss the general empirical model underlying our approach. We next turn to a description of our data sets, the National Long-Term Care Survey (NLTCS) and the National Nursing Home Survey (NNHS). In the section following, we describe the implementation of the empirical model, the mechanisms used to deal with the multiple types of patients included in the two surveys, and the empirical results. We then turn to simulation results derived from the model, and close with a discussion of the implications of the analysis.

10.1 Empirical Framework

Our duration model incorporates three mutually exclusive states: residence in the community, residence in a nursing home, and death. Two sets of transition probabilities within and across states fully describe behavior in the model, one set for each state other than death, which represents the absorbing state. We structure the transitions in the model as follows. Persons who reside in the community face a probability each period of remaining in the community; the alternative event is exiting to a different state. Given an exit, one of two mutually exclusive events can occur: the individual either dies or enters a nursing home. We structure the transition probabilities for residence in a nursing home in an analogous manner. That is, we estimate the probability that an individual remains in the nursing home for the next month. If the person exits, we assess the probability that the exit is to the community, the alternative being death. Of course, as the absorbing state, death completes an individual's experience. Hence, we specify four basic probabilities to describe lifetime nursing home experiences. Given knowledge of these four probabilities, we can calculate the probability of observing various patterns of lifetime experiences. After introducing specifications for these probabilities, we describe a simulation procedure to summarize lifetime nursing home experiences.

10.1.1 Model Characterization

Define $P(i \rightarrow j) \equiv P(i \rightarrow j | X, H, \tau)$ as the probability that an individual transits from state i to state j in the next month, conditional on the values of X, H, and τ. X is a covariate vector made up of demographic characteristics; H is a vector of variables summarizing an individual's nursing home history prior to the start of the current spell in state i; and τ is the current number of consecutive months in state i. The primary covariate included in H is an indicator variable signifying whether the individual has had a nursing home admission between age 65 and the current age. In the duration model considered here, i and j take values representing each of the three states: community (C), nursing home (N), and death (D). Hospitalizations are included as part of state C. With three states, nine transition probabilities are required to fully characterize experiences in the model, i.e., the probabilities $P(i \rightarrow j)$ for i, j = C, N, D. For the trivial case of the absorbing state death, $P(\mathrm{D} \rightarrow \mathrm{D}) = 1$ always, so $P(\mathrm{D} \rightarrow \mathrm{C}) = P(\mathrm{D} \rightarrow \mathrm{N}) = 0$. Consequently, it is the remaining six transition probabilities that are of interest.

Consider the transition probabilities associated with residence in the community: $P(\mathrm{C} \rightarrow j)$ for j = C, N, D. Because the sum of these three probabilities equals one, the specification of two probabilities provides a complete characterization. The quantity $1 - P(\mathrm{C} \rightarrow \mathrm{C})$ is the probability that the individual exits the community in the next month. Given the event that a community spell ends, let $P(\mathrm{D}|\text{C-Exit}) \equiv P(\mathrm{D}|\text{C-Exit}, X, H, \tau)$ denote the probability that an individual dies; this probability explicitly conditions on exiting the community after a spell lasting τ months with history H for an individual with characteristics X. Thus, $1 - P(\mathrm{D}|\text{C-Exit})$ represents the likelihood that the individual enters a nursing home given termination of a community spell. Specifying the quantities, $P(\mathrm{C}\rightarrow\mathrm{C})$ and $P(\mathrm{D}|\text{C-Exit})$ implies the following relationships among the three transition probabilities for community residents:

$$\begin{aligned} &P(\mathrm{C}\rightarrow\mathrm{C}) = P(\mathrm{C}\rightarrow\mathrm{C}),\\ (1)\qquad &P(\mathrm{C}\rightarrow\mathrm{N}) = [1 - P(\mathrm{C}\rightarrow\mathrm{C})][1 - P(\mathrm{D}|\text{C-Exit})],\\ &P(\mathrm{C}\rightarrow\mathrm{D}) = [1 - P(\mathrm{C}\rightarrow\mathrm{C})]P(\mathrm{D}|\text{C-Exit}). \end{aligned}$$

Because these quantities depend on an individual's current and recent history, they vary from month to month. Knowledge of $P(\mathrm{C}\rightarrow\mathrm{C})$ and $P(\mathrm{D}|\text{C-Exit})$ provides complete information regarding the experience of an elderly community resident in any month of his life.

Nursing home residence is characterized by an analogous pair of probabilities. The quantity $P(\mathrm{N}\rightarrow\mathrm{N})$ is the transition probability that an individual remains in a nursing home for the next period, conditional on values of X, H, and τ; thus $1 - P(\mathrm{N}\rightarrow\mathrm{N})$ is the probability that the individual leaves a nursing home in the next period. Given termination of a nursing home spell, let $P(\mathrm{D}|\text{N-Exit})$ be the probability that an individual dies. From the two probabilities, $P(\mathrm{N}\rightarrow\mathrm{N})$ and $P(\mathrm{D}|\text{N-Exit})$, one can infer the three transition probabilities associated with nursing home residence via the formulae

$$\begin{aligned} &P(\mathrm{N}\rightarrow\mathrm{N}) = P(\mathrm{N}\rightarrow\mathrm{N}),\\ (2)\qquad &P(\mathrm{N}\rightarrow\mathrm{C}) = [1 - P(\mathrm{N}\rightarrow\mathrm{N})][1 - P(\mathrm{D}|\text{N-Exit})],\\ &P(\mathrm{N}\rightarrow\mathrm{D}) = [1 - P(\mathrm{N}\rightarrow\mathrm{N})]P(\mathrm{D}|\text{N-Exit}). \end{aligned}$$

Thus, knowledge of $P(\mathrm{N}\rightarrow\mathrm{N})$ and $P(\mathrm{D}|\text{N-Exit})$ provides complete information regarding experience at any time in the life of an elderly nursing home resident.

10.1.2 Formulating Duration Distributions

The specification of the transition probabilities $P(\mathrm{C}\rightarrow\mathrm{C})$ and $P(\mathrm{N}\rightarrow\mathrm{N})$ provides all the information needed to formulate duration distributions for nursing home and community spells, where spell duration refers to the number of continuous months in a state. Consider the distribution describing the length of stays in state i. With $P(i\rightarrow i) \equiv P(i\rightarrow i|X, H, \tau)$ representing the probability of remaining in state i at least another month, given continuous occupancy of τ months and given the values of X and H, the duration distribution for spells is given by

$$(3)\qquad f_i(\tau) = S_i(\tau - 1)\,[1 - P(i\rightarrow i)],$$

with

$$S_i(\tau - 1) = \prod_{t=1}^{\tau-1} P(i \rightarrow i). \tag{4}$$

The quantity $S_i(\tau - 1)$ is the survivor function, representing the probability that a spell lasts at least $\tau - 1$ months. Both X and H are fixed at the time of entry into the state; X includes such factors as gender and age, and H incorporates measures of prior experience. The probability that an exit occurs at τ, given that a spell has lasted at least $\tau - 1$ months, is defined as

$$h_i(\tau) = \frac{f_i(\tau)}{S_i(\tau - 1)} = 1 - P(i \rightarrow i). \tag{5}$$

Negative (positive) duration dependence exists if h increases (decreases) as a function of duration τ.

10.1.3 Analytically Characterizing Nursing Home Experiences

The above elements serve as building blocks for several approaches to summarizing lifetime nursing home experiences. First, one could compute the likelihood of observing any lifetime experience as an appropriate product of the transition probabilities in equations (1) and (2). Formula (3) shows how to calculate the probability of observing a spell of any length at any time in an individual's life. The likelihood of observing any sequence of spells can be calculated as the product of the probabilities of observing the given spells and the appropriate $P(\mathrm{D}|i\text{-Exit})$.

Consider an individual who lives in the community at age 65, remains in the community for three months, and enters a nursing home for two months before dying. The probability of observing this pattern of lifetime experiences is

$$f_{\mathrm{C}}(3)\,[1 - P(\mathrm{D}|\mathrm{C}\text{-Exit}, X, H, 3)]\,f_{\mathrm{N}}(2)\,P(\mathrm{D}|\mathrm{N}\text{-Exit}, X, H, 2),$$

where the covariates X and H are properly updated to reflect changes with age and history. The same likelihood can be calculated via the six transition probabilities as

$$\left[\prod_{t=1}^{2} P(\mathrm{C} \rightarrow \mathrm{C}|X, H, t)\right] P(\mathrm{C} \rightarrow \mathrm{N}|X, H, 3)$$
$$\left[\prod_{t=1}^{1} P(\mathrm{N} \rightarrow \mathrm{N}|X, H, t)\right] P(\mathrm{N} \rightarrow \mathrm{D}|X, H, 2).$$

Summarizing the distribution of lifetime nursing home utilization within this framework requires computing the probability of observing every possible pattern of lifetime experiences and then calculating the cumulative duration of each state. For example, calculating the expected number of months of nursing home residence after a particular age requires a summation over probabilities associated with every possible pattern of experiences. Rather than proceeding with this cumbersome approach, we implement an alternative that uses a familiar simulation procedure to summarize lifetime experiences.

10.1.4 Simulating Nursing Home Experiences of the Elderly

We can simulate the lifetime experiences of an individual by randomly drawing discrete variables indicating a person's state of residence in each month, using the transition and exit probabilities to determine the values of these variables given the realized experiences to date. The simulation method that we employ makes use of predicted values of the probabilities of equations (1) and (2). For an individual with characteristics X, we construct predicted values of the four probabilities, $P(i \rightarrow i)$ and $P(\mathrm{D}|i\text{-Exit})$ for $i = \mathrm{C}$ and N, recomputing for each period as we update the values of X, H, and the current spell duration, τ. By evaluating these probabilities in every month of a lifetime we simulate a person's entire set of lifetime experiences.

Consider a 65-year-old male living in the community. Calculate the predicted probability that he will exit the state in the next month, and draw a random variable from a uniform distribution between zero and one. If the realization is greater than the predicted probability that the individual exits the community, assign him to the community state for another period, update τ, and repeat the process until a random variable is drawn such that he does exit the community. When an exit occurs, calculate the predicted probability that the man exits to death, and again draw a random variable. If the realization of the random variable is less than the predicted probability of death, assign the man to death for the period, completing his experience; otherwise, assign the man to the nursing home and repeat the process using predicted probabilities associated with nursing home residence. The simulation appropriately updates the values of the covariates, and hence the transition probabilities, as the individual ages.

Repeating the above procedure numerous times and recording sequences of monthly experiences for a large number of hypothetical individuals allows us to characterize the nursing home experiences for a population of 65-year-old males residing in the community. With suitable adjustments in the values of the covariates, the same procedure can be applied to simulate lifetime experiences for various populations.

10.1.5 Initial Conditions

Although our empirical model is designed to predict the distribution of nursing home utilization at ages 65 and older, a small number of 65-year-olds are already in nursing homes. A comprehensive description of nursing home utilization from age 65 to death must account for their nursing home experience. To incorporate utilization from this group in our forecasts, we include a distinct component of our empirical model that estimates the likelihood that an individual resides in an institution upon reaching age 65, and the distribution of the number of months that elapse from the date of the 65th birthday until death or discharge from the nursing home. Along with race and sex, these quantities constitute the initial conditions of our forecasting model. We estimate the probability of nursing home occupancy by directly calculating the proportion of our nationally representative sample that resides in a nursing home upon reaching age 65. We estimate the distribution of subsequent duration by calculating the empirical survivor function from our sample. This approach is equivalent to using nonparametric procedures to estimate the initial conditions.

10.2 The Data

To estimate the parameters of the empirical model, we require longitudinal data on individuals residing in the communities and on individuals residing in nursing homes. In developing the model, we assume that it is important to distinguish experiences that represent *recurrent* nursing home and community spells from initial spells. We do not expect persons residing in the community following a nursing home discharge, for example, to have the same survival or subsequent nursing home utilization as a person of the same age who was never in a nursing home. Insofar as second and higher-order spells differ from first spells, it is important to distinguish them in the modeling. This is only possible if the data permit this distinction.

This poses a significant challenge to our analysis, since as far as we know, no nationally representative data set offers a random population of individuals in second or higher-order spells in sufficient numbers and with sufficient information about them to carry out the analysis. Hence we combine data from two distinct data sets, the NLTCS and the NNHS, to estimate the parameters of the empirical model. We use the NLTCS to estimate components of our empirical model needed to describe community spells and their mode of termination. We use the NNHS for the corresponding analyses of nursing home admissions and forecast the patterns of utilization by applying the empirical model, which is estimated from the combined data.

10.2.1 The National Long-Term Care Survey

The NLTCS offers suitable data for the analysis of community spells because it includes a stratified random sample of spells over a two-year window for the entire population of elderly, along with information about nursing home entry. Other data sets have similar information on nursing home utilization, but offer either selected populations or inadequate spell information, making comparable analysis impossible. For example, the National Long-Term Care Demonstration enrolled only a sample of high-risk elderly suffering at least one "unmet need," and the National Nursing Home Survey enrolled only nursing home residents. Its population coverage makes the NLTCS appropriate for estimating community survival rates.

Although the NLTCS was designed to provide information on the need for long-term care services of the functionally impaired community-based elderly population, it also gathered information on other groups of elderly people. One of its goals was to discover characteristics that are associated with utilization of home-based long-term care, and so to anticipate future needs for other services for the chronically disabled elderly. To address these issues, the NLTCS identified and interviewed a sample of chronically disabled individuals who resided in the community. They also kept and reported critical information on the remaining segments of the elderly population.

The first survey was administered in 1982. From a randomly selected sample of 50,000 Medicare recipients in the (Medicare) Health Insurance Skeleton Eligibility Write-off (HISKEW) in 1982, about 6,000 community-dwelling disabled elderly were identified by a telephone screening process. Although this is the only population that was surveyed in 1982, another 2,000 elderly individuals were determined to be institutionalized. Another 5,000 individuals were identified who would turn age 65 by the 1984 follow-up interview date (the 1984 aged-in population). The remaining 25,000 screened-out were nondisabled and living in the community. Thus, although we only have detailed information on a limited subset of the elderly in 1982, we are able to identify the vital status and place of residence (community or nursing home) of all individuals included in the 1982 and 1984 files.

The 1984 NLTCS makes use of three detailed questionnaires: one administered to the 1984 community population, a second administered to the institutional population, and a third administered to the next of kin of those who died between the 1982 and 1984 interview dates. All of the 6,000 disabled community residents from 1982 were automatically reinterviewed regardless of their 1984 functional status, as were the 2,000 individuals in the 1982 institutional population. All of the aged-in cohort, and roughly half of those screened-out in 1982, were rescreened in 1984 following the same process as in 1982. Thus, the 1984 NLTCS contains detailed information on a broad community-based elderly population (of roughly 6,000), an institutionalized elderly population (of roughly 1,700), and a recently deceased elderly population (of roughly 2,500). In addition, information collected in the screening process is provided for the screened-out population—individuals who were not functionally impaired (roughly 14,000).

The weights we use are the 1984 final screener cross-sectional weights, which are calculated as the reciprocal of the probability of inclusion in the sample when the sample is stratified across characteristics such as age, gender, and race. Summing these weights produces a national cross-sectional estimate of the U.S. population above age 64. Thus, although some subgroups may be oversampled, by appropriate weighting we can generalize our results to the entire elderly population.

NLTCS Nursing Home Data

NLTCS is an incomplete data source, for the purposes of our analysis, because it has incomplete information about nursing home utilization for some subgroups. While the community and deceased questionnaires offer complete retrospective histories, the institutional questionnaire focuses only on the period between 1982 and the current interview period. Information concerning utilization prior to the 1982 interview date, such as admission and discharge dates and the number of prior admissions, was not collected. Thus the duration and number of previous nursing home and community spells is unknown. Furthermore, for spells beginning after 1982 but prior to the current admission, only the admission dates were collected, making it impossible to determine their durations. Consequently, the NLTCS cannot be used to compile complete nursing home histories for this important component of the elderly population. Although this does not impede our estimation of the likelihood of experiencing a first nursing home spell, it seriously impairs our ability to measure durations of nursing home spells.

We compiled information covering 2,774 nursing home episodes, of which 1,604 came from the 1984 institutional questionnaire and were therefore still in progress at the end of the study, or right-censored. Another 1,170 spells came from the remaining questionnaires. The information from the retrospective components of these questionnaires limits spell durations to less than 100 months. In addition, the deceased questionnaire, which allows us to calculate long spells if they end in death or just prior to death, includes a heavily used "Don't Know" category, further confounding our efforts to create a random sample of nursing home spells. As a result, while 5 percent of the 2,774 spells are longer than 100 months, all but two of them are censored, making it impossible to observe either the exact duration or outcome of long spells. Table 10.1 summarizes the nursing home spell information calculated from the NLTCS.

NLTCS Community Spell Data

We also use the NLTCS data to characterize duration and exit probabilities for the repeat community spells, or the community spells that follow nursing home discharges. Although the retrospective character of the NLTCS makes it possible to calculate spell data regarding experiences prior to the 1982 interview date, inclusion of repeat community spells initiated prior to the 1982 interview date would introduce a (positive) selection bias. The bias occurs because repeat community spells ending in discharge to death prior to the 1982 interview date are excluded, so probability of inclusion in the sample is enhanced for those with longer spells. Furthermore, since short spells ending prior to 1982 are included in the sample universe only if they do not end in death, we would underrepresent short repeat community spells that end in death. Thus, if we included pre-1982 experiences, our estimates of the probability of nursing home readmissions (i.e., the probability that a repeat community spell terminates in nursing home admission) would be biased upward. In

Table 10.1 **LTCS Nursing Home Spell Summary Statistics**

	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
Institutional questionnaire					
Spell length (months)	1,601	34.53	36.82	1	229
Male = 1	1,601	0.23	0.42	0	1
White = 1	1,601	0.94	0.24	0	1
Age (months)	1,601	973.18	91.06	781	1,234
Other questionnaires					
Spell length (months)	1,169	5.24	10.04	1	112
Male = 1	1,169	0.33	0.47	0	1
White = 1	1,169	0.94	0.24	0	1
Age (months)	1,169	957.93	88.13	781	1,194
All spells					
Spell length (months)	2,770	22.17	32.17	1	229
Male = 1	2,770	0.27	0.45	0	1
White = 1	2,770	0.94	0.24	0	1
Age (months)	2,770	966.74	90.13	781	1,234
Censored = 1	2,770	0.58	0.49	0	1

order to avoid these selection biases, we include in our analysis only those community spells that begin between the two survey interview dates. Thus, we include all spells that commence after the 1982 interview date and conclude with exits either to death or to repeat nursing home admissions or are right-censored at the 1984 interview date. Many spells are therefore right-censored and it is impossible to characterize fully repeat community experiences of more than about 28 months. Table 10.2 summarizes the NLTCS data used to characterize the return community spells.

To measure the probabilities that an individual will enter a nursing home for the first time or will die, we use the information regarding these transitions between the 1982 and 1984 interview dates. Inclusion of nursing home admissions prior to the 1982 interview date would introduce selection biases analogous to those for repeat community spells discussed above. We drop all individuals who experienced first spells prior to the 1982 interview and an additional 254 individuals from the institutional questionnaire, for whom we cannot determine when the first spell occurred. Table 10.3 summarizes NLTCS data used to calculate community spells prior to a first nursing home episode after age 65.

10.2.2 The 1985 National Nursing Home Survey

The 1985 NNHS, which is a nationally representative data set developed to provide comprehensive information about nursing home services, offers extensive data about the lengths of nursing home spells, their modes of termination, and prior nursing home usage. Our analysis uses the discharged resident file of the NNHS, which provides a random sample of nursing home spells that

Table 10.2 LTCS Return Community Spell Summary Statistics

	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
Spell length (months)	472	6.33	7.15	1	28
Male = 1	472	0.33	0.47	0	1
White = 1	472	0.94	0.23	0	1
Age (months)	472	965.24	87.15	788	1,190
Censored = 1	472	0.51	0.50	0	1

Table 10.3 NLTCS Summary Statistics for First Community Spells

	1984 Community Cohort	1984 Institutional Cohort	1984 Deceased Cohort	1984 Screened-out Cohort	Total
Total individuals	5,934	1,690	2,475	14,145	24,244
Individuals dropped					
Ambiguous history	0	254	0	0	0
Transition pre-1982 date	154	1,099	140	24	326
Inaccurate transition date	192	24	147	22	385
Total individuals remaining	5,588	313	2,188	14,099	22,188
Transitions with state changes	64	313	2,188	35	2,700
Censored spells	5,424	0	0	14,064	19,488
Transitions without state changes	161,547	5,870	31,837	389,500	588,754

	Mean	Standard Deviation	Minimum	Maximum
Observations per individual	23.99	7.15	1	33
Fraction with exit	0.12	0.33	0	1
Male = 1	0.41	0.49	0	1
White = 1	0.91	0.29	0	1
Age at 1982 interview date (months)	868.48	85.66	781	1,443

are completed within the 12 months prior to the survey date. All spells are noncensored, and outcomes for each spell are known. The file includes basic demographic characteristics, health status, place of residence prior to admission and after discharge, services received, and indications of the outcome of care. Weights associated with each observation are derived as the reciprocal of the probability of inclusion in the sample for each stratum along which the selection is based.

The discharged resident file includes information on 5,317 nursing home discharges. We compile information on 4,705 nursing home episodes beginning after age 65, of which 2,354 are first spells. The remaining 612 spells began prior to the individual's sixty-fifth birthday. Table 10.4 summarizes characteristics of the population included in the file. A comparison of tables 10.1 and 10.4 reveals the difficulties in utilizing the nursing home spell information from the NLTCS. The spell characteristics from both data sets for all spells are similar; however, the mean of the completed spells in the NLTCS is only slightly greater than 5 months, while the mean of the censored spells is more than 34 months.

10.2.3 Definition of Variables

All of our analyses incorporate demographic characteristics as covariates. The variable MALE is set to one for males and zero for females. The variable AGE refers to the age of the individual in months beyond his or her sixty-fifth birthday. The variable WHITE is set to one if the individual is white and zero otherwise.[1]

We account for history dependence by specifying the spell number of a particular state. We set the dichotomous dummy variable CM1 equal to one for those community spells that occur before a first nursing home admission, or for which the only prior nursing home admissions were completed before reaching age 65. The covariate PRE65 is set to one to account for differing experiences of those with completed nursing home spells prior to age 65. The dichotomous variable NH1 equals one for first nursing home admission and zero otherwise.

10.2.4 Initial Conditions

The NLTCS provides the nationally representative sample from which we calculate the quantities needed to infer initial conditions. This sample consists of a random selection (stratified and weighted) of individuals reaching age 65 between the 1982 interview data and the 1984 interview date. Although the NLTCS did not record information on these individuals in 1982, the individuals were screened in 1984 and, according to the result of the screen, either administered the institutional questionnaire, the community questionnaire, or their next of kin were administered the deceased questionnaire. For respondents not meeting the criteria for any of these survey instruments, the screening information was reported. Thus, we can identify the characteristics of each individual on his or her sixty-fifth birthday.

Few of these individuals had nursing home admissions prior to age 65. For

---

[1] The NLTCS does not distinguish respondents of Hispanic origin. The NNHS contains a separate variable for Hispanic origin, but since Hispanic is not a racial classification, it categorizes some as white and some as black. Because there are few Hispanics in the NNHS and because the NLTCS data set does not identify them at all, we use the race variable that does not contain information regarding Hispanics. Thus, blacks, Asians, Pacific Islanders, and Native Americans, including Eskimos and Aleuts, make up the population in which WHITE equals zero.

Table 10.4 NNHS Nursing Home Spell Summary Statistics

	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
First nursing home spells					
Spell length (months)	2,348	21.22	34.82	1	321
Male = 1	2,348	0.33	0.47	0	1
White = 1	2,348	0.93	0.25	0	1
Age (months)	2,348	979.87	90.77	781	1,298
End in death	2,348	0.48	0.50	0	1
Second nursing home spells					
Spell length (months)	1,012	19.93	30.63	1	239
Male = 1	1,012	0.32	0.47	0	1
White = 1	1,012	0.96	0.20	0	1
Age (months)	1,012	985.96	89.74	781	1,267
End in death	1,012	0.41	0.49	0	1
Third and above nursing home spells					
Spell length (months)	1,337	13.53	21.01	1	231
Male = 1	1,337	0.30	0.46	0	1
White = 1	1,337	0.94	0.24	0	1
Age (months)	1,337	988.34	91.27	781	1,221
End in death	1,337	0.15	0.36	0	1
All spells					
Spell length (months)	4,697	18.75	30.74	1	320
Male = 1	4,697	0.32	0.47	0	1
White = 1	4,697	0.94	0.24	0	1
Age (months)	4,697	983.59	90.75	781	1,298
End in death	4,697	0.37	0.48	0	1
Number of spells	4,697	2.15	1.76	1	22

those who had not completed the admission, we estimate the corresponding distribution of duration in the nursing home beyond the age of 65.

We construct weighted nonparametric estimates of the likelihood of every possible combination of covariates occurring. Because the weights convert the data to national totals by age, gender, and race, our procedure estimates initial covariate values for a nationally representative population. We draw our combinations of initial covariates using these estimates, thus reproducing the characteristics of a random sample of the entire 65-year-old U.S. population.

10.3 Empirical Specifications and Results

To implement the empirical formulation described in section 10.1, we need to specify two sets of transition probabilities, $P(i \rightarrow i)$, where i = N, C: the probabilities of remaining in either the community or a nursing home; and the exit probabilities $P(\mathrm{D}|i\text{-Exit})$, giving the probabilities that the community or nursing home spells terminate in death. In this section, we first introduce empirical specifications for the transition probabilities $P(\mathrm{C} \rightarrow \mathrm{C})$ and $P(\mathrm{N} \rightarrow \mathrm{N})$.

We then introduce specifications for the exit probabilities $P(\mathrm{D}|i\text{-exit})$ ($i = \mathrm{C}, \mathrm{N}$). Finally, we present the estimates that result from implementing these specifications.

10.3.1 Parameterizations of Transition Probabilities

The presence of duration dependence implies that $P(i \rightarrow i)$ varies with the number of continuous months in a state (τ). The empirical specification needs to be flexible enough to account for general types of duration dependence, which may be nonmonotonic. Return episodes in the community may show increasing duration dependence for small values of τ, as short stays (say in a hospital) end either in death or in return to a nursing home, and decreasingly thereafter. Eventually any spell will exhibit decreasing duration dependence for large values of τ, especially at very advanced ages. Standard empirical specifications typically do not accommodate nonmonotonic duration dependence, nor do they allow for interaction between the form of duration dependence and covariates X and H. To allow for such flexibility, we specify the following logistic model for the probabilities $P(i \rightarrow i)$:

$$P(i \rightarrow i) = \frac{1}{1 + e^{Z_1\beta + \theta(\tau,\, Z_2)}}, \tag{6}$$

where Z_1 and Z_2 are vectors of variables included in X and H, β is a parameter vector, and the function $\theta(\tau, Z_2)$ captures the properties of duration dependence.

Spline models are an attractive approach for modeling duration effects, since they fit the data with a flexible and smooth function of duration. Implicit in conventional spline models, which fit polynomial functions to a series of intervals over duration, is a tradeoff between smoothness and goodness of fit. Fit can be improved by increasing the number of polynomial functions, but nondifferentiability at the boundaries requires a sacrifice in smoothness. Limiting the number of intervals or the order of the polynomial functions yields a smoother curve but diminishes the capabilities of detecting complicated forms of duration dependence.

In our approach, we specify $\theta(\cdot)$ as the general function:

$$\theta(\tau, X_2) = \sum_{j=1}^{K_\theta} [\Phi_{1j}(\tau) - \Phi_{2j}(\tau)]\,(\alpha_{0j}Z_2 + \alpha_{1j}\tau + \alpha_{2j}\tau^2 + \alpha_{3j}\tau^3). \tag{7}$$

K_θ represents the number of polynomial functions entered in the specification. The quantity $\Phi_{ij}(\tau)$ denotes the cumulative distribution function of a normal random variable having mean μ_{ij} and standard deviation σ_{ij}, and the coefficients α_{ij} represent parameter vectors. This "overlap-spline" specification allows the polynomial functions to be operative as the linear combination of cumulative normal functions (i.e., $\sum \Phi_{1j}(\tau) - \Phi_{2j}(\tau)$) takes nonzero values. Thus, each polynomial function begins and ends gradually, according to the properties of the cumulative normal. As the value of σ decreases, the polynomial has a more abrupt impact; as σ approaches zero, this specification approaches that of conventional spline models without equality constraints at interval boundaries. While our overlap-spline specifications do not require equality constraints, they do allow for several polynomials to be operative over an interval. Unlike conventional splines, they are also continuously differentiable at all points.

This specification of $\theta(\cdot)$ allows covariates in Z_2 to have different effects at different ages. Covariates in Z_1 from model (6) are similar to covariates in a typical proportional hazards model in that they simply shift the hazard rate up or down at all times.

To allow the experiences in first community spells and first nursing home spells to differ from subsequent spells, we estimate distinct transition probabilities for first and later episodes. The literature often refers to this form of history dependence as "occurrence dependence."

10.3.2 Parameterizations of Exit Probabilities

The probabilities $P(\mathrm{D}|i\text{-Exit})$ are also likely to display history dependence. For example, if short nursing home spells are predominantly short recuperative episodes, they are more likely to end in return to the community. Longer spells, especially for the very old, may more frequently end in death. Therefore, the functional specification for the exit probabilities, like the transition probabilities, should be flexible enough to detect dependence on previous spell duration as well as interactions of characteristics, X, with duration. We adopt the specification

$$(8) \qquad P(\mathrm{D}|i\text{-Exit}) = \frac{1}{1 + e^{-Z_1\gamma - \Gamma(\tau,\, Z_2)}},$$

where Z_1 and Z_2 are vectors of variables included in X and H, γ is a parameter vector, and the function $\Gamma(\tau, Z_2)$ captures dependence on the duration of the previous spell, τ. The function $\Gamma(\cdot)$ takes a form analogous to the overlap-spline function $\theta(\cdot)$:

$$(9) \qquad \Gamma(\tau, Z_2) = \sum_{j=1}^{K_\Gamma} [\Phi_{1j}(\tau) - \Phi_{2j}(\tau)]\, (Z_2\gamma_{2j}),$$

where $\Phi_{ij}(\tau)$ represents the cumulative normal distribution function, as before, and τ is the spell length of the previous spell. K_Γ represents the number of polynomial functions entered in the specification. If $\sigma_{ij} = 0$ for all i and j, the model can be interpreted as interacting a series of dummy variables in duration with the covariates in Z_2, where the values of the μ_{ij} terms determine the range over which the dummy variables apply. In this specification, the Z_2 variables can have different effects as duration changes for the spell just ended. For example, the model could detect relatively high mortality rates for the very old, given a short nursing home admission. This phenomenon would be reflected in the γ parameters. Our overlap-spline specification allows for the effect of the dummy variables to change gradually with time, according to the properties of the cumulative distribution functions, $\Phi_{ij}(\tau)$.

For the analysis reported below, we fix the values of K_θ, K_Γ, and of each μ_{ij} and σ_{ij}. We selected their values after extensive exploratory data analysis; the properties of the empirical hazard rates principally determine the formulation of the function $\theta(\cdot)$, and the properties of the empirical exit probabilities guided the formulation of the function $\Gamma(\cdot)$ (i.e., the values of the exit probabilities as a function of duration determine $\Gamma(\cdot)$).

10.3.3 Estimation of Community Transition Probabilities

We accommodate history dependence (distinct from duration dependence) in estimating $P(\mathrm{C}\rightarrow\mathrm{C})$ by introducing the covariate CM1 in the specification (6). CM1 is a binary variable that indicates that the individual has had no nursing home experiences since age 65. Incorporating it into equation (6) allows the duration characteristics of first community spells (CM1 = 1) to differ from those of return community spells (CM1 = 0). We fully interact CM1 with all other variables to fully distinguish the characteristics of first and higher-order spells in the community. Because AGE and τ are exactly collinear for first community spells, we substitute AGE for τ when CM1 = 1 in equation (6) and estimate $P(\mathrm{C}\rightarrow\mathrm{C}|X, H, \mathrm{AGE})$ at each age. To avoid length-biased sampling when estimating the conditional exit probabilities, we include only those transitions that occur between interview dates in the NLTCS.

We estimate the parameters of $P(\mathrm{C}\rightarrow\mathrm{C})$ in equation (6) using weighted maximum likelihood to account for the nonrandom sample design. After exploring a variety of specifications for equation (6), we adopted a formulation for CM1 spells with MALE, WHITE, and PRE65 in Z_1 and an intercept in Z_2. In this case, the best combination of smoothness and fit resulted in the overlap-spline parameters of $K_\theta = 1$ with (μ_{11}, σ_{11}) and (μ_{21}, σ_{21}) set to any combinations such that, for every observation in the data,

$$\Phi_{11}(\mathrm{AGE}) - \Phi_{21}(\mathrm{AGE}) = 1.$$

These values are not unique: any set of pairs of (μ_{11}, σ_{11}), with μ_{11} small enough, and (μ_{21}, σ_{21}), with μ_{21} large enough, will produce the same result. This is equivalent to dropping the overlap-spline terms altogether and entering an intercept, MALE, WHITE, PRE65, AGE, AGE^2, and AGE^3 in Z_1.

The specification for return community spells (CM1 = 0) includes MALE, WHITE, AGE, and AGE^2 in Z_1 and an intercept in Z_2.[2] Table 10.5 contains the coefficient estimates for the community spell duration model.

---

[2] It also sets $K_\theta = 3$ with $(\mu_{11}, \sigma_{11}) = (-5, 0.5)$, $(\mu_{12}, \sigma_{12}) = (3, 1.0)$, $(\mu_{21}, \sigma_{21}) = (3, 1.0)$, $(\mu_{22}, \sigma_{22}) = (16, 2.0)$, $(\mu_{31}, \sigma_{31}) = (16, 2.0)$, and (μ_{32}, σ_{32}) set to any combination such that, for every observation in the data, $\Phi_{32}(\tau)$ equals zero for all τ. We also restrict $\alpha_{21} = \alpha_{22} = \alpha_{j3} = 0$ for j = 1, 2, and 3, eliminating the linear and quadratic terms in duration for the second overlap-spline, the quadratic term in duration for the third overlap-spline, and the third-order term from all overlap-splines.

Table 10.5 **Parameter Estimates of Community Spell Duration Probabilities**

A. First Community Spells

	Coefficient	Standard Error
Z_1 Variables		
MALE	−0.40633	0.042886
WHITE	−0.04472	0.070229
PRE65 (prior nursing home)	0.263979	0.450623

Spline: All Months

	Coefficient	Standard Error
Z_2 Variables		
Intercept	7.255503	0.117733
AGE	−0.01756	0.001858
AGE^2	0.000047	0.00001
AGE^3	−6.5E-08	1.60E-08

B. Return Community Spells

	Coefficient	Standard Error
Z_1 Variables		
AGE	−0.00045	0.005521
AGE^2	0.000005	0.000013
MALE	0.490293	0.279404
WHITE	0.308954	0.373481

	Spline					
	1–3 Months		3–16 Months		≥16 Months	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
Z_2 Variables						
Intercept	1.978263	1.316833	−4.22992	0.671357	−13.1521	3.154348
t	−3.76324	1.508056			0.420809	0.134577
t^2	0.652585	0.423117				

The results for first community spells indicate that duration dependence is monotonically decreasing and accelerating. That is, $P(C \rightarrow C | X, H, \text{AGE})$ always decreases and at an increasing rate as AGE increases. Men have significantly shorter spells throughout, but neither race nor prior nursing home experience significantly alter spell lengths.

Second and higher-order community spells (CM1 = 0) display nonmonotonic duration dependence, rapidly increasing over the first few months, slowly increasing over the next year, and decreasing and accelerating for the remaining nine months. The age at which these spells are experienced has little impact on the form of duration dependence and the magnitude of $P(\mathrm{C}\rightarrow\mathrm{C}|X, H, \tau)$. The variable MALE is of borderline significance, reducing $P(\mathrm{C}\rightarrow\mathrm{C}|X, H, \tau)$ throughout τ at all ages. Individuals have only a 62 percent change of surviving in the community the first month, while the chance of surviving each of months 3 through 22 exceeds 95 percent.

10.3.4 Estimation of Nursing Home Transition Probabilities

We control for history dependence in a similar manner when estimating model (6) for nursing home residence. In order to allow estimates for first nursing home admissions (NH1 = 1) to vary from those of second and higher-order nursing home admissions (NH1 = 0), we interact the dummy variable NH1 with the other covariates. We also include measures of age. The final specification of equations (6) and (7) for nursing home spells includes MALE and WHITE in Z_1; an intercept, AGE, and AGE^2 in Z_2.[3] Table 10.6 contains the coefficient estimates of equations (6) and (7) for the nursing home spells.

First and repeat nursing home spells display the same general forms of duration dependence, sharply increasing over the first several months, then leveling off and slowly decreasing. The coefficient of MALE is significantly different from zero for both types of admissions, but it is almost 60 percent larger for first admissions. The decreased duration of nursing home stays for men is most pronounced for first admissions. The variable WHITE does not have a significant impact on the lengths of either first or subsequent admissions.

10.3.5 Estimation of Community Exit Probabilities

In estimating the probability that a community spell will end in death rather than in nursing home entry, we control for history dependence in several ways. First, in order to allow the characteristics of first community spell exits to

---

[3] This specification sets the spline parameter to $K_\theta = 3$ with, for NH1 = 1,

$$(\mu_{11}, \sigma_{11}) = (-5, 0.5), \quad (\mu_{21}, \sigma_{21}) = (4, 0.5),$$
$$(\mu_{12}, \sigma_{12}) = (4, 0.5), \quad (\mu_{22}, \sigma_{22}) = (20, 1.0), \text{ and}$$
$$(\mu_{13}, \sigma_{13}) = (20, 1.0),$$

and for NH1 = 0,

$$(\mu_{11}, \sigma_{11}) = (-5, 0.5), \quad (\mu_{21}, \sigma_{21}) = (4, 1.0),$$
$$(\mu_{12}, \sigma_{12}) = (4, 1.0), \quad (\mu_{22}, \sigma_{22}) = (22, 4.0), \text{ and}$$
$$(\mu_{13}, \sigma_{13}) = (22, 4.0),$$

and (μ_{23}, σ_{23}) is set to any combination such that, for every observation in the data, $\Phi_{23}(\tau)$ equals zero, for all types of nursing home spells. We restrict α_{23} to be zero for first nursing home spells, thus eliminating the quadratic term in the third spline for these spells. We also restrict α_{3j} to be zero for both types of nursing home spells, thus eliminating the third-order term in duration.

Table 10.6 **Parameter Estimates of Nursing Home Spell Duration Probabilities**

A. First Nursing Home Spells

	Coefficient	Standard Error
Z_1 Variables		
MALE	0.278368	0.072025
WHITE	0.003483	0.11065

	Spline					
	1–4 Months		4–20 Months		≥20 Months	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
Z_2 Variables						
AGE	−0.00087	0.001819	−0.00003	0.000023	0.000248	0.000213
AGE^2	4.75E-07	0.000004	6.23E-08	5.35E-08	−7.1E-08	5.19E-07
Intercept	0.101047	0.317329	−1.59763	0.473573	−4.11875	0.254231
t	−0.81425	0.296768	−0.17948	0.081135	0.00007	0.001238
t^2	0.056743	0.071692	0.004761	0.003546		

B. Second or Higher-Order Nursing Home Spells

	Coefficient	Standard Error
Z_1 Variables		
MALE	0.17539	0.068032
WHITE	0.090283	0.102742

	Spline					
	1–4 Months		4–22 Months		≥22 Months	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
Z_2 Variables						
AGE	−0.00057	0.001946	0.000006	0.00002	−0.00008	0.000237
AGE^2	0.000002	0.000005	1.11E-10	4.81E-08	8.92E-07	5.87E-07
Intercept	−0.75217	0.342505	−2.4842	0.476187	−3.60134	0.33752
t	−0.46025	0.285716	−0.03874	0.071455	−0.00492	0.006101
t^2	0.023516	0.069963	−0.00051	0.002806	0.000032	0.000033

differ from those of return community spells, we fully interact the dummy variable CM1. We also control for the duration of the current spell, τ, by introducing the function $\Gamma(\cdot)$ as discussed above. Finally, to determine how $P(\mathrm{D|C\text{-}Exit})$ changes with age, we interact a quadratic polynomial in age with overlap-splines in duration of the previous spell.

The duration of first community spells (CM1 = 1) is always equal to the age of the individual; thus, we substitute AGE for τ and estimate P(D|C-Exit, X, H, AGE) for these observations. The final specification of models (8) and (9) for exits from community spells with CM1 = 1 includes MALE, WHITE, and PRE65 in Z_1 and an intercept, AGE, and AGE^2 in Z_2.[4] Because τ is replaced by AGE in the overlap-spline specification, we are allowing for different polynomials in AGE to be operative over different age intervals, determined by the properties of the overlap splines. We use this specification explicitly to allow exit characteristics to be different for the very old. For return community spell exits (CM1 = 0), the final specification of equations (8) and (9) includes WHITE in Z_1 and an intercept, MALE, AGE, and AGE^2 in Z_2.[5] The overlap-spline intervals are determined over the values of previous spell duration, τ. We fit separate polynomials in AGE over the two intervals determined by the overlap-spline properties, one for short previous spells and one for long previous spells. Table 10.7 contains the coefficient estimates of equations (8) and (9) for community spell exits.

The results, as displayed in table 10.8, indicate that most people at all ages die on their first community exit. Females are more likely than males, and whites are more likely than nonwhites, to enter nursing homes on termination of a first community spell. For men at age 65, there is about an 85 percent chance that a community spell will end in death; the probability is about 75 percent for females. The percentage of individuals who terminate community spells by entering nursing homes gradually increases with age, until it reaches a peak (at about age 85) of almost 30 percent for males and 53 percent for females. The probability then decreases sharply with age. The risk of entering a nursing home is greatest between the ages of 78 and 95, with more than 20 percent of male exits and more than 30 percent of female exits at each age.

Return community spell exits (CM1 = 0), as displayed in table 10.8, differ from first community exits. The likelihood that a community spell will end in death increases steadily through age. Long return community spells, defined as more than six months in the community, are much more likely to end in nursing home admissions at young ages, with fewer than 10 percent of 65-year-olds dying on exit, while almost 80 percent of men and 60 percent of women aged 95 die on exit. Mortality rates for short spells vary less with age, ranging from a low of around 60 percent dying at age 65 to about 65 percent dying at age 95. In other respects, the return community spells are similar to first spells; whites are more likely than blacks to enter nursing homes on exiting a return community spell. Females exiting long return community spells

---

[4] This specification sets the spline parameter $K_\Gamma = 2$ with $(\mu_{11}, \sigma_{11}) = (-5, 0.5)$, $(\mu_{12}, \sigma_{12}) = (200, 10.0)$, $(\mu_{21}, \sigma_{21}) = (200, 10.0)$, and (μ_{22}, σ_{22}) set to any combination such that, for every observation in the data, Φ_{22}(AGE) equals zero.

[5] We set the spline parameter $K_\Gamma = 2$ with $(\mu_{11}, \sigma_{11}) = (-5, 0.5)$, $(\mu_{12}, \sigma_{12}) = (6, 0.5)$, $(\mu_{21}, \sigma_{21}) = (6, 0.5)$, and (μ_{22}, σ_{22}) set to any combination such that, for every observation in the data, $\Phi_{22}(\tau)$ equals zero.

Table 10.7 Parameter Estimates of Community Spell Exit Probabilities

A. First Community Exit Probabilities

	Coefficient	Standard Error
Z_1 Variables		
MALE	−0.703840	0.095611
WHITE	0.482762	0.170774
PRE65 (prior nursing home)	2.491183	1.163058

	Spline			
	1–200 Months		≥200 Months	
	Coefficient	Standard Error	Coefficient	Standard Error
Z_2 Variables				
Intercept	−1.60232	0.19079	−3.87718	2.064165
AGE			0.024363	0.014405
AGE^2	0.000018	0.000006	−0.00005	0.000025

B. Return Community Exit Probabilities

	Coefficient	Standard Error
Z_1 Variable		
WHITE	0.389739	0.2181

	Spline			
	1–6 Months		≥6 Months	
	Coefficient	Standard Error	Coefficient	Standard Error
Z_2 Variables				
MALE	0.2181	0.342772	−0.77631	0.931566
Intercept	−0.801	0.697308	2.682614	1.561231
τ	−0.00238	0.001887	−0.00906	0.005279

are more likely than males to enter nursing homes, but females exiting short return community spells are more likely than males to die.

10.3.6 Estimation of Nursing Home Exit Probabilities

The same structure serves to account for history dependence in nursing home exits. We allow first nursing home admissions (NH = 1) to have different exit characteristics from second and higher-order nursing home admissions (NH1 = 0) by fully interacting the dummy variable NH1. The final specifica-

Table 10.8 **Probability of Dying Given a Community Spell Exit**

A. Exit from First Community Spells

Age	Male	Female
70	0.853	0.742
75	0.828	0.704
80	0.778	0.634
85	0.702	0.538
90	0.702	0.539
95	0.766	0.618
100	0.863	0.757
105	0.944	0.892

B. Exit from Second or Higher-Order Community Spells

	Male		Female	
Age	Short Spell	Long Spell	Short Spell	Long Spell
70	0.547	0.394	0.600	0.225
75	0.627	0.223	0.677	0.114
80	0.677	0.213	0.722	0.108
85	0.698	0.353	0.742	0.196
90	0.695	0.701	0.739	0.512
95	0.665	0.956	0.712	0.906
100	0.608	0.998	0.658	0.995
105	0.517	1.000	0.571	1.000

tion is identical for both types of admissions and has WHITE in Z_1, and an intercept, MALE, AGE, and AGE^2 in Z_2. We set the overlap-spline parameters to distinguish between very short (less than three months) and other nursing home spells.[6] Table 10.9 contains the coefficient estimates of equations (8) and (9) for nursing home exits.

Table 10.10 contains results for exits from first nursing home spells. Most spells at early ages end in return to the community, while the probability of a spell ending in death increases monotonically with age. The probability that a nursing home spell will end in death increases rapidly with age for long nursing home spells. Although age is a significant determinant of the probability for long spells, it is not significant for short spells. The variable MALE is significant for short spells, but not for long spells; consequently, women are much more likely than men to return to the community when exiting short nursing home spells. The probability that a nursing home spell will end in death does not vary with race.

Table 10.10 also contains results for exits from second or higher-order nurs-

---

[6] The overlap-spline parameters are $K_{\Gamma} = 2$, $(\mu_{11}, \sigma_{11}) = (-5, 0.5)$, $(\mu_{12}, \sigma_{12}) = (2.5, 0.5)$, $(\mu_{21}, \sigma_{21}) = (2.5, 0.5)$, and (μ_{22}, σ_{22}) is set to any combination such that, for every observation in the data, $\Phi_{22}(\tau)$ equals zero.

Table 10.9 **Parameter Estimates of Nursing Home Spell Exit Probabilities**

A. First Nursing Home Exit Probabilities

	Coefficient	Standard Error
Z_1 Variable		
WHITE	0.010358	0.200626

	Spline			
	1–2.5 Months		≥2.5 Months	
	Coefficient	Standard Error	Coefficient	Standard Error
Z_2 Variables				
MALE	−0.9564	0.246972	−0.10943	0.150629
Intercept	1.410593	0.477743	1.670388	0.383976
τ	−0.0004	0.004379	−0.00809	0.003043
τ^2	−2.2E-06	0.00001	0.000004	0.000006

B. Second or Higher-Order Nursing Home Exit Probabilities

	Coefficient	Standard Error
Z_1 Variable		
WHITE	0.013078	0.259445

	Spline			
	1–2.5 Months		≥2.5 Months	
	Coefficient	Standard Error	Coefficient	Standard Error
Z_2 Variables				
MALE	−0.49471	0.257421	0.029287	0.160889
Intercept	2.959729	0.753246	1.867464	0.424702
τ	−0.00763	0.006547	−0.00019	0.003156
τ^2	0.000006	0.000014	−8.8E-06	0.000007

ing home spells. For these spells, the probability of death increases monotonically with age for both short and long spells. Race remains insignificant, and gender is only of borderline significance for short spells, in which men have higher mortality rates.

10.4 Simulation Results

The assessment of alternative policies toward financing or delivering long-term care requires full characterization of the distribution of nursing home utilization. Measures of central tendency, such as the mean or median, are of

Table 10.10 **Probability of Dying Given a Nursing Home Spell Exit**

A. Exit from First Nursing Home Spells

	Male		Female	
Age	Short Spell	Long Spell	Short Spell	Long Spell
70	0.393	0.249	0.200	0.229
75	0.405	0.340	0.207	0.316
80	0.420	0.436	0.218	0.409
85	0.439	0.529	0.232	0.501
90	0.463	0.612	0.249	0.586
95	0.490	0.683	0.270	0.658
100	0.521	0.739	0.295	0.718
105	0.556	0.784	0.325	0.765

B. Exit from Second or Higher-Order Nursing Home Spells

	Male		Female	
Age	Short Spell	Long Spell	Short Spell	Long Spell
70	0.115	0.134	0.073	0.137
75	0.162	0.147	0.105	0.150
80	0.216	0.169	0.144	0.173
85	0.274	0.205	0.187	0.209
90	0.331	0.257	0.232	0.263
95	0.385	0.331	0.276	0.338
100	0.431	0.431	0.316	0.438
105	0.468	0.551	0.349	0.559

limited value in estimating utilization, since insurance is often concerned with providing protection against catastrophic events rather than covering expected utilization for the average enrollee. In some respects, then, it is as important to learn the likelihood that an individual will have a very lengthy nursing home admission as it is to know the expected value of utilization. These quantities can be obtained from our empirical analysis, but doing so requires simulations that construct the distributions of nursing home and community spells.

Our simulations follow the procedure described in section 10.1.4 and are based on random draws of 5,000 hypothetical individuals from the estimated distribution of initial covariate values. Key results from the simulation appear in table 10.11, which lists summary statistics for the simulation population, measures of nursing home utilization, and durations of community spells and survival.

Because we drew the simulated population from a distribution representing the composition of the U.S. 65-year-old population in 1984, the hypothetical population should be similar to the actual population. Forty-four percent of the simulated population are men, and 91 percent are white. Fewer than 0.4 percent had been discharged from a nursing home before reaching age 65, and

Table 10.11 **Simulation Results**

	Mean	1%	5%	10%	25%	Median	75%	90%	95%	99%
MALE = 1	0.44									
WHITE = 1	0.91									
PRE65 = 1 (prior nursing home)	0.0038									
Fraction residing in nursing home at 65th birthday	0.0062									
Fraction with at least one nursing home admission	0.35									
Number of nursing home admissions	0.53	0	0	0	0	0	1	2	2	4
Male	0.51	0	0	0	0	0	1	1	2	4
Female	0.54	0	0	0	0	0	1	2	2	4
Total nursing home utilization (months)	7.47	0	0	0	0	0	2	23	50	111
Male	5.80	0	0	0	0	0	2	17	37	87.28
Female	8.77	0	0	0	0	0	3	29.1	61	119.91
Nursing home utilization for those with at least one admission (months)	21.06	1	1	1	2	6	27	66	88.3	139
Male	16.11	1	1	1	1	4	20	50	70.5	127.5
Female	25.03	1	1	1	2	9	34	74.5	100.75	155.45
Time before first nursing home admission (months)	207.98	0	44	74.4	142.5	212	280	330.2	361	414.26
Male	191.58	0	37.5	60	127	195	261	312	336.5	381.1
Female	221.13	0	49	87.5	154	227	292.75	341.5	376	426.3

Spell length of first nursing home admissions (months)	14.45	1	1	1	1	3	14	48	69	135
Male	11.05	1	1	1	1	2	9	34	54.05	102.21
Female	17.19	1	1	1	1	3	18	57	83	136.35
Spell length of nonfirst nursing home admissions (months)	12.27	1	1	1	2	4	14	33	53.4	103.96
Male	11.11	1	1	1	1	4	12	27.8	53.6	98.52
Female	13.01	1	1	1	2	5	15	36	53.95	112.11
Time spent in community after first nursing home admission, for those with at least one nursing home admission (months)	11.92	0	0	0	0	1	25	33	53.4	103.96
Male	8.67	0	0	0	0	1	9.5	29	43	65.1
Female	14.53	0	0	0	0	1	27	39.5	60	102
Time from onset of first spell until death, for those with at least one nursing home admission (months)	32.98	1	1	1	3	18	49	91	114	172.25
Male	24.77	1	1	1	2	9	35	69	93.5	167
Female	39.55	1	1	1	5	28	61.5	100	132	187.12
Age at death (months after 65th birthday)	207.50	9	43	67	129	204	284	350	380	436
Male	186.14	6	38	61	115	183	252	317	352	394.07
Female	224.17	13	48.45	75	142.25	225	309	366.1	398	446

0.62 percent were residents in nursing homes on their sixty-fifth birthdays. Thus 99 percent of the simulated population had no prior nursing home admissions at age 65.

The distribution of age at death is similar to figures obtained from the 1984 life tables for the United States, particularly in the middle of the distribution. The median age at death from the life table and from the simulation is 82. The simulations tend to predict slightly lengthier survival than do the life-table figures for individuals in the lower half of the distribution (up to just over one year).

Perhaps the most common statistic used to describe nursing home utilization is the probability that an individual who reaches age 65 will *ever* enter a nursing home. The simulation reveals that 35 percent have at least one nursing home admission; this compares with figures in the literature ranging from 25 to 50 percent. Although a substantial minority (10 percent) have more than one admission, very few have large numbers of admissions; only 0.5 percent have more than four.

More important is the likelihood of a lengthy admission. Many of the people admitted to nursing homes spend little time there. Almost 25 percent of those with at least one admission spend only one month in nursing homes during their lifetime. Of the individuals with some nursing home utilization, only half have more than six months of accumulated nursing home residency. But a substantial minority have lengthy stays—nearly 25 percent of all women who enter a nursing home will spend three years there. Thus the distribution of nursing home utilization is highly skewed; most elderly persons never enter a nursing home, many enter for short times, and a substantial minority spend years in an institution. This minority accounts for a large fraction of nursing home utilization.

When nursing home admission occurs, it tends to be at advanced ages. For men, the median age at a first admission exceeds the median age of death; most community spells that terminate at young ages end in death, while nursing home admissions occur in the middle of the age distribution. The median age of first admission is about age 81 for men and age 84 for women; half of all individuals who enter nursing homes do so for the first time between the ages of 76 and 88.

Inasmuch as 42 percent of first nursing home admissions end in death, entering a nursing home signals a high mortality rate. However, if the admission terminates in discharge alive, the outlook is not grim, at least for women. Leaving the admission alive marks the onset of the first return community spell, which on average lasts slightly more than a year for women and less than nine months for men. Of those who have return community spells, females accumulate far more months of residency. Half of the females with return community spells experience at least two years in the community, while the corresponding value for males is around nine months. The large difference in male and female return community experiences persists throughout the distributions.

From an insurer's point of view, a potential subscriber with a history of health service utilization is likely to have above-average claims. In this respect, it is interesting to examine the utilization pattern of individuals with at least one prior admission. The simulations reveal that the mean duration of repeat nursing home admissions is longer than first admissions for men, but not for women. The upper percentiles of spell durations for women are substantially greater for first admissions than for repeat admissions. For men and women combined, fewer than 10 percent of repeat spells are more than three years long, while 10 percent of first spells are at least four years long. The skew in utilization revealed in these simulations suggests that there should be a strong role for insurance; a large percentage of people are at risk for very limited nursing home stays, but a nonnegligible minority will have lengthy admissions that could lead to catastrophic expenditures.

10.5 Conclusions

The elderly have a substantial risk of entering nursing homes. According to our analysis, most of the admissions are short, and multiple admissions are uncommon. Median nursing home utilization is only six months for those with at least one admission, and only 25 percent of 65-year-olds can expect to spend more than two months in a nursing home during the remainder of their lives.

We find that very heavy utilization is somewhat less likely than Kemper and Murtaugh (1991) reported in the only published study that attempts to measure cumulative nursing home utilization in a random sample of elderly Americans. They reported a similar lifetime risk of nursing home utilization for 65-year-olds—37 percent in their study, compared to 35 percent in ours. They also found, as we did, that about one-third of persons who have any admissions will spend less than three months in nursing homes. However, in their study, 21 percent of persons with a nursing home admission, compared to 28 percent in our study, spent 3–12 months in a nursing home. The differences in findings are most striking for persons institutionalized for several years; 12 percent of the patients with any nursing home utilization in our simulations, compared to 17 percent of their subjects, spent five or more years in a nursing home. Kemper and Murtaugh estimate that about 48 percent of people with any admissions will have at least one year of utilization, compared to our estimate of 40 percent. Their results may differ from ours because they are based on recall alone, which may be particularly inaccurate for long stays, while ours are based on direct measurements of length of stay from two nationally representative data sources. Like Kemper and Murtaugh, though, we find that few of the elderly have prolonged stays and that those who do account for most nursing home utilization. Thus there is a nonnegligible but small risk of "catastrophic" nursing home admissions and a very high likelihood of experiencing a brief admission.

Our projections do not take into account changes in mortality rates over time, changes in the organization and supply of institutional care, or behavioral responses to changes in financing. If long-term care insurance is purchased more frequently, or if government agencies begin to finance or supply long-term care, utilization is likely to change. We previously showed that for Medicare-financed nursing home admissions, the exhaustion of Medicare nursing home benefits leads to a sharp rise in the discharge rate, suggesting that utilization is responsive to price (Garber and MaCurdy 1991). Since insurance constitutes a subsidy to covered services, it is likely that the development of insurance would lead to increased utilization as compared with what we observe in these samples. Thus our estimates provide evidence for the belief that there is a substantial potential market for long-term care insurance. Changing social arrangements, economic factors, and the dissemination of insurance could promote even greater use of long-term care in the future.

References

Branch, L. G., and A. M. Jette. 1982. A prospective study of long-term care institutionalization among the elderly. *American Journal of Public Health* 72:1373–79.

Garber, A. M., and T. MaCurdy. 1991. Nursing home discharges and exhaustion of Medicare benefits. NBER Working Paper no. 3639. Cambridge, Mass.: National Bureau of Economic Research.

Keeler, E. B., R. L. Kane, and D. H. Solomon. 1981. Short- and long-term residents of nursing homes. *Medical Care* 19:363–69.

Kemper, P., and C. M. Murtaugh. 1991. Lifetime use of nursing home care. *New England Journal of Medicine* 324:595–600.

Lewis, M. A., S. Cretin, and R. L. Kane. 1985. The natural history of nursing home patients. *Gerontologist* 25:382–88.

Liu, K., T. Coughlin, and T. McBride. 1991. Predicting nursing-home admission and length of stay: A duration analysis. *Medical Care* 29:125–41.

McConnel, C. E. 1984. A note on the lifetime risk of nursing home residency. *Gerontologist* 24:193–98.

Murtaugh, C. M., P. Kemper, and B. C. Spillman. 1990. The risk of nursing home use in later life. *Medical Care* 28:952–62.

Palmore, E. 1976. Total chance of institutionalization among the aged. *Gerontologist* 16:504–7.

Spence, D. A., and J. M. Wiener. 1990. Nursing home length of stay patterns: Results from the 1985 National Nursing Home Survey. *Gerontologist* 30:16–20.

Vicente, L., J. A. Wiley, and R. A. Carrington. 1979. The risk of institutionalization before death. *Gerontologist* 19:361–67.

Wingard, D. L., D. W. Jones, and R. M. Kaplan. 1987. Institutional care utilization by the elderly: A critical review. *Gerontologist* 27:156–63.