2009-07-09

Estimating the Effect of Disability on Medicare Expenditures

David Morris Burk
Brigham Young University - Provo

ESTIMATING THE EFFECT OF DISABILITY ON MEDICARE EXPENDITURES

by

David M. E. Burk

A project submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Statistics

Brigham Young University

August 2009

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a project submitted by

David M. E. Burk

This project has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	H. Dennis Tolley, Chair
Date	G. Bruce Schaalje
Date	John Lawson

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the project of David M. E. Burk in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date	H. Dennis Tolley Chair, Graduate Committee

Accepted for the Department

Scott D. Grimshaw
Graduate Coordinator

Accepted for the College

Thomas W. Sederberg
Associate Dean, College of Physical and Mathematical Sciences

ABSTRACT

ESTIMATING THE EFFECT OF DISABILITY
ON MEDICARE EXPENDITURES

David M. E. Burk

Department of Statistics

Master of Science

We consider the effect of disability status on Medicare expenditures. Disabled elderly historically have accounted for a significant portion of Medicare expenditures. Recent demographic trends exhibit a decline in the size of this population, causing some observers to predict declines in Medicare expenditures. There are, however, reasons to be suspicious of this rosy forecast. To better understand the effect of disability on Medicare expenditures, we develop and estimate a model using the generalized method of moments technique. We find that newly disabled elderly generally spend more than those who have been disabled for longer periods of time. Also, we find that increases in expenditures have risen much more quickly for those disabled Medicare beneficiaries who were at the higher ends of the expenditure distribution before the increases.

ACKNOWLEDGEMENTS

I am grateful to my Mom, Dad, brother, Abby, Leonard, Olivia, the late Dagny, and all my teachers in this program and throughout the years—especially Professor Tolley and Gnat—who directly or indirectly have helped me to put this project together.

CONTENTS

TABLES

Table

FIGURES

Figure

1 Introduction

Rising costs for health care represent the single greatest challenge to balancing the federal budget. Spending on health care and related activities will account for nearly 18% of the gross domestic product in 2009—an expected total of $2.5 trillion; under current policies, that share is projected to exceed 20% in 2018. Federal spending accounts for roughly one-third of those totals (Elmendorf 2009). Two-thirds of the federal health budget is spent on Medicare, a health insurance program enacted in 1965 to provide coverage for Americans age 65 and over. In 2006, Medicare spending totaled an estimated $381.9 billion (Orszag 2007). This total neglects the uncompensated costs associated with informal care provided primarily by families. Hence, as the size of the older population grows, tracking the proportion who need assistance with daily tasks has become an important and highly policy-relevant exercise.

Expenditures for disabled elderly constitute a large portion of Medicare spending. Life expectancy has steadily been on the rise, from 70 to 77, over the last 50 years (National Center for Health Statistics 2009). Considerable attention has been devoted to whether this rise in life expectancy has been accompanied by a contraction in the portion of disabled elderly. After a period of some ambiguity on this point due to various data sources, observers have recently agreed on a declining trend in the portion of disabled elderly (Freedman et al. 2004). This has led some to have positive outlooks for the long-term solvency of the Medicare program (Manton and Gu 2001).

There is cause for hesitation, however. In Table 1 we present summary statistics on the proportion of disabled elderly (generated from the National Long-term Care Survey, as reported by Manton, Gu, and Lamb 2006). Disabilities are considered to be of two types—disability in Instrumental Activities of Daily Living (IADL) and, more severely, in Activities of Daily Living (ADL) (see, for example, Mosby 2006). The former category refers to need for assistance in activities such as preparing meals, doing housework, shopping, and managing money. The latter refers to need

Table 1: Proportion of disabled persons by type of disability (National Long-term Care Survey, 1982–2005). Annual declines in proportion with any disability are accompanied by a slight increase in the proportion with severe disabilities.

Year	1982	1984	1989	1994	1999	2004/2005
Total disabled	26.5	26.2	24.8	23.2	21.2	19.0
IADL, one or two ADL	12.5	12.9	11.1	10.5	9.6	8.0
Three+ ADL	6.4	6.3	6.8	6.3	6.7	7.0
Institutionalized	7.5	7.0	6.9	6.3	4.9	4.0
Per annum declines, %		0.6	1.1	1.3	1.8	2.2

for assistance in such activities as dressing, bathing, toileting, moving from and into a chair, and eating. IADL disabilities generally are less severe than ADL disabilities, and require less expensive care. Also, having more instances of a disability—needing assistance in multiple activities rather than just in one—is more expensive.

Table 1 shows that while there has been a decline in the total number of disabled, the costliest category has experienced an increase. Explanations for such trends abound. For example, the decline in disabled persons may be due to the advent and propagation of assistive medical devices that allow people with what previously would have been considered minor disabilities to be effectively undisabled. Due to improved end-of-life care, severely disabled may be living longer than previously, causing their number to increase. More research needs to be done on these mechanisms.

The purpose of this paper is to investigate the relationship between disability and Medicare expenditures, and thereby shed light on the claim that the overall decline in the portion of disabled elderly implies lower Medicare expenditures for the disabled. We proceed by modeling the variation in Medicare expenditures by demographic and disability characteristics. We estimate the model using data from the Medicare Current Beneficiary Surveys from a 12-year period.

2 Model

We model variation in Medicare expenditures as being determined by life-cycle profile and cohort of an individual through the empirical relationship

$$y_i = g(\mathbf{x}_i) + u_i$$

where y_i measures expenditures for individual i and $\mathbf{x}_i$ represents a column vector of characteristics of individual i. The deterministic function g measures trends and life-cycle profiles, and the residual u_i reflects deviations from these trends. Hereafter, for notational convenience we drop the subscript i.

The model g consists of two components, one which depicts the cohort expenditure function for those who are not disabled, and one for those who are disabled. That is,

$$g = \boldsymbol{\beta}(\mathbf{x}) + \boldsymbol{\gamma}(\mathbf{x}), \tag{1}$$

where $\boldsymbol{\beta}(\mathbf{x})$ depicts each cohort's expenditure function for those who are not disabled, and $\boldsymbol{\beta}(\mathbf{x}) + \boldsymbol{\gamma}(\mathbf{x})$ depicts the expenditure function for those who are disabled. In particular, we define $\boldsymbol{\beta}(\mathbf{x})$ as a function of age (denoted by a) and cohort (c), where cohort is indicated by the year of birth:

$$\boldsymbol{\beta}(a, c) = \beta_0(c) + \beta_1(c)a + \beta_2(c)a^2$$

where

$$\beta_j(c) = \beta_{0j} + \beta_{1j}c + \beta_{2j}c^2, \quad j = 0, 1, 2. \tag{2}$$

Thus expenditure is assumed to follow a quadratic function in age and cohort.

The component $\boldsymbol{\gamma}(\mathbf{x})$ captures the impact of disability on expenditures. We wish to allow disability to have an immediate impact, as well as a component that grows (or, possibly, decreases) over time. Let p denote the age at which a disabled individual acquired his or her disability. Then $\boldsymbol{\gamma}(\mathbf{x}) = \boldsymbol{\gamma}(a,p)$ is the additional annual cost of care for an individual with current age a, who became disabled at age $p < a$. We assume the form of the model to be

$$\boldsymbol{\gamma}(a,p) = \delta\left[\gamma_{10} + \gamma_{11}a\right] + \Phi_q(a-p)\left[\gamma_{20} + \gamma_{21}a\right]$$

where Φ_q is a sigmoid function such that

$$\Phi_q(a-p) \approx \begin{cases} 0 & \text{when } (a-p) < 0 \\ 0 \text{ increasing to } 1 & \text{when } 0 \le (a-p) \le q \\ 1 & \text{when } (a-p) > q \end{cases} \tag{3}$$

where q is the duration of disability in years at which we regard expenditures as stabilizing with respect to duration of disability. For now, we somewhat arbitrarily let $q = 5$, though this can easily be adjusted.

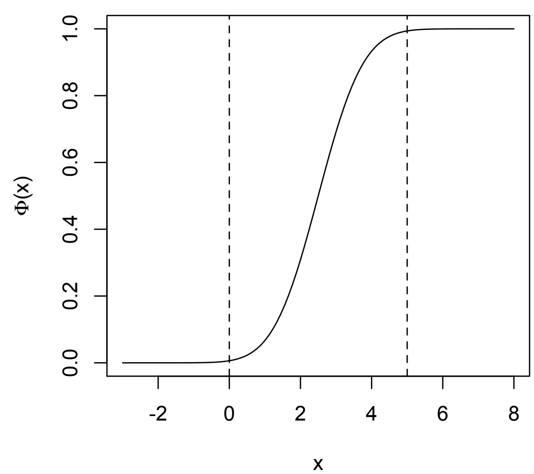

Figure 1: Effect of disability as a function of years since onset of disability, $\Phi_q(a-p)$ with $q = 5$. The specification of this function includes no effect before the onset of disability, gradual onset of disability for the first five (q) years, and full effect of disability after five years. The normal distribution function displayed here (mean 1, variance 2.5) meets these specifications.

The first term of the right hand side of the expression for $\boldsymbol{\gamma}$ in equation 3 captures the immediate impact of disability; the second captures the gradual effect. The function Φ_q above is such that the gradual effect of disability increases until the duration of disability has reached five years, and then flattens out (see Figure 2). A specific function that meets these requirements is the cumulative distribution function of a normal distribution with a mean, say, of 1 and variance, say, of, 2.5. Of course, a function could be chosen that gradually increases over 3 or 10 or any arbitrary number of years.

Thus the model is written

$$\begin{aligned} y &= g(\mathbf{x}) + u \\ &= \boldsymbol{\beta}(\mathbf{x}) + \boldsymbol{\gamma}(\mathbf{x}) + u \\ &= \beta_0(c) + \beta_1(c)a + \beta_2(c)a^2 + \delta\left[\gamma_{10} + \gamma_{11}a\right] + \Phi_q(a-p)\left[\gamma_{20} + \gamma_{21}a\right] + u, \end{aligned} \tag{4}$$

making the appropriate substitutions from Equations 2 and 3.

The model does not contain an explicit measure of the severity of the disability. This is primarily due to data limitations. In the face of these data limitations, however, the duration of disability, which the model includes, seems a sensible proxy for severity of disability. We estimate the model separately for males and females, to allow the effects to vary by sex.

3 Data

We use data from the Medicare Current Beneficiary Survey (MCBS). The MCBS is a continuous, multipurpose survey of a representative national sample of the Medicare population, conducted by the Center for Medicare and Medicaid Services (CMS), the federal agency that oversees the Medicare program. The data were compiled through surveys and official administrative records. They provide complete expenditure data on all health care services paid by Medicare. Survey-reported data include information on the use and cost of all types of medical services, as well as beneficiary personal characteristics such as health status and physical functioning. Medicare claims data includes use and cost information on various types of medical services, including hospital, outpatient, and in-home services. Individuals in the MCBS are followed annually for at most four years.

We use data from the MCBS for each year between 1992 and 2003. We use individual-level data on annual Medicare expenditure amount, age, year, and dichotomous disability status. From the age and year variables we constructed a cohort variable. As the minimum age of eligibility for Medicare services generally is 65 years old, all individuals are at least of age 65. Observations at ages above 95 were excluded due to data sparseness at these ages. The data set consists of 124,085 unique observations over the 12-year period.

Even though there are repeat observations on a few individuals, we treat the data as a snapshot of cross-sectional data. That is, we regard the data as observations on distinct individuals. We do not exploit the panel nature of the data because there are insufficient observations. Ignoring the fact that the data do actually contain repeat observations on a few individuals exacerbates heteroscedasticity of the residuals. Our estimation procedure accounts for this.

Accordingly, we have 124,085 individual-level observations with the variables age, sex, expenditures, disability status, and cohort (the year the individual was born). Additionally, each observation contains a frequency weight, indicating how many times that observation should appear in the data set. Imputing this frequency weight gives a total of 412 million observations, or roughly 32 million in year 1992 increasing to nearly 36 million in year 2003. All the summary statistics and analysis in this paper incorporate these frequency rates. Figure 2 presents expenditure levels at .1 quantile, the median, and the .9 quantile, across age and for the cohorts born in 1910, 1920, and 1930.

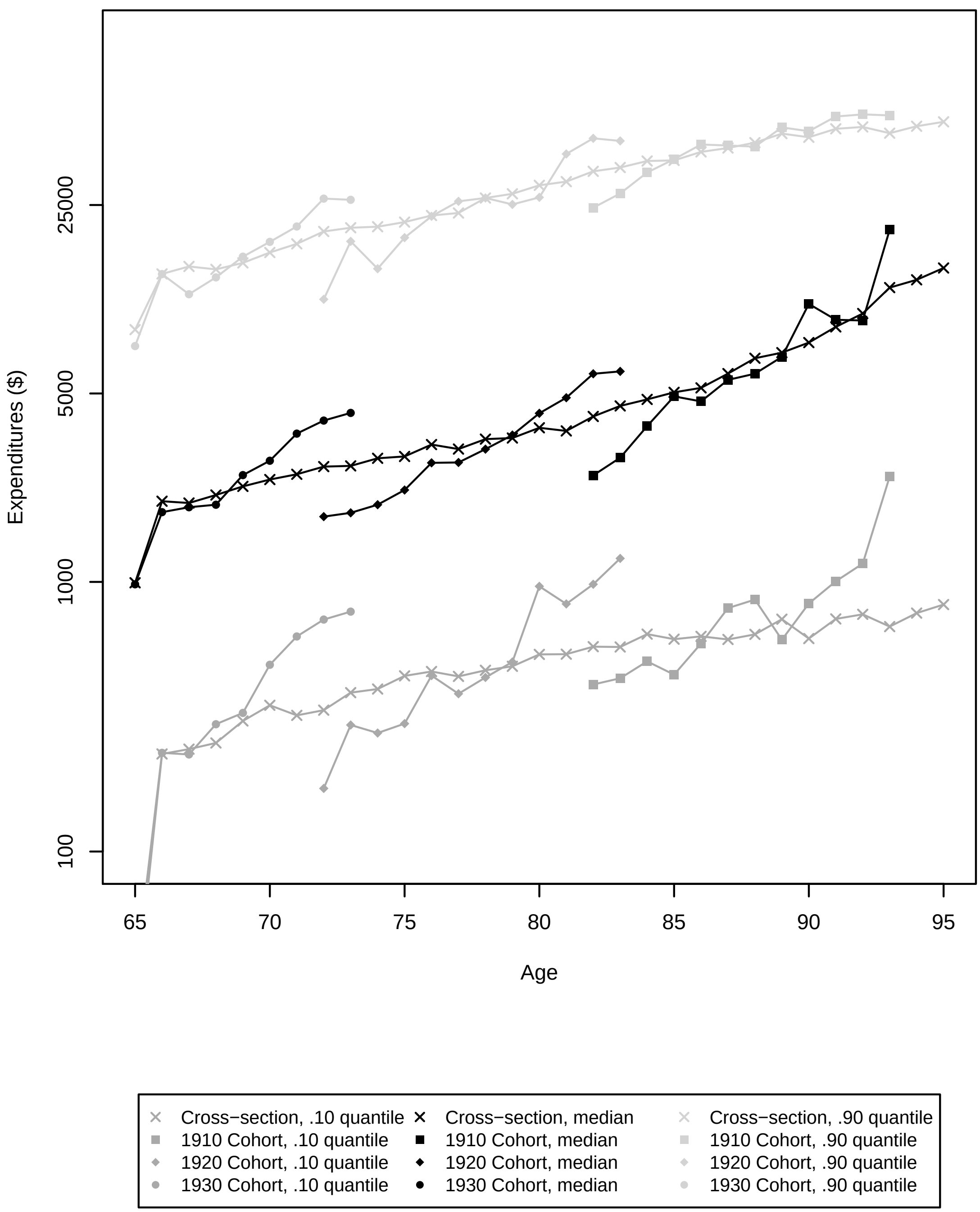

Figure 2: MCBS expenditures by age and cohort. The tenth, fiftieth, and ninetieth percentiles of expenditure are graphed against age for the aggregated cross-sectional data and for the 1910, 1920, and 1930 birth cohorts. As age increases, the expenditures increase—for both the cross-sectional data and within each birth cohort.

Table 2: Summary statistics for MCBS Data. Table entries represent mean (first quartile, third quartile).

	Female	Male	Overall
Age	75.8	77.7	76.9
(years)	(70,81)	(71,84)	(70,83)
Expenditures	9014	10210	9721
($)	(1016,8913)	(1257,10580)	(1158,9831)
Proportion Disabled	0.041	0.048	0.045

Summary statistics of the age and expenditures are in Table 2. The density plot in Figure 3 makes apparent that log expenditures (scaled down by a factor of $10,000$), are roughly distributed normally.

These data omit the age at which a disabled individual became disabled, a variable necessary to implement the model described in the previous section. Accordingly, this is a value we estimate, as described in the next section. For this estimation we exploit the time-series nature of the data.

We assume that no one is disabled before entering the data set. Therefore, we regard those 65-year-olds who are disabled as having incurred their disability that year. Additionally, we assume that once a person becomes disabled, he or she cannot become nondisabled.

4 Estimation

Here we discuss the estimation procedure we implement for the model. We describe the estimation of the age at which disability was acquired as well as the Generalized Method of Moments, the estimation procedure we use to estimate the parameters of the model.

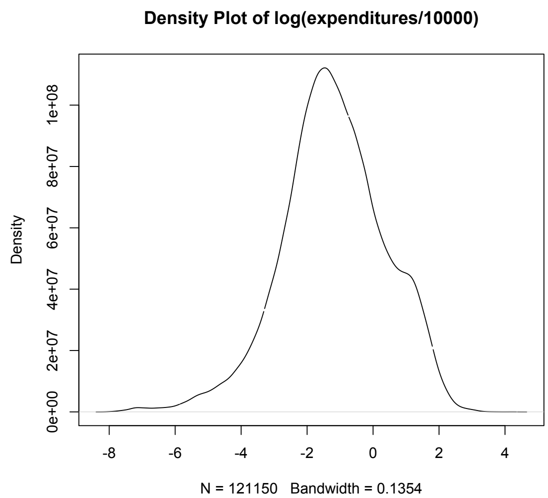

Figure 3: Distribution of the log-transformed expenditure data. The response variable, the expenditure amount for each patient, looks reasonably normal after a log transformation.

4.1 Estimating Age at Acquisition of Disability

The model in Equation 4 depends on p, the age at which disability was acquired, which is unobserved in our data set.

Our strategy is to substitute $\mathrm{E}[\boldsymbol{\gamma}(a,p)]$ for $\boldsymbol{\gamma}(a,p)$ in Equation 4, giving the relation

$$\mathrm{E}[y] = \boldsymbol{\beta}(a,c) + \mathrm{E}[\boldsymbol{\gamma}(a,i)].$$

We then estimate the probability distribution for p. We let $\mathbb{P}_{m,a}(p)$ denote the probability that a person of age a was disabled at age p, where m denotes the person's sex. We estimate

$$\begin{aligned} \mathrm{E}[y] &= \boldsymbol{\beta}(a,c) + \mathrm{E}[\boldsymbol{\gamma}(a,i)] \\ &= \boldsymbol{\beta}(a,c) + \sum_{i=65}^{\mathrm{a}} \mathbb{P}_{m,a}(i)\boldsymbol{\gamma}(a,i). \end{aligned}$$

To estimate $\mathbb{P}_{m,a}(p)$, we used the panel nature of the data as described above. We considered only individuals for whom there were multiple observations and who became disabled in the data set. We computed the proportion of people in this subset who became disabled for each age-sex combination in the data set. This process generated the estimates in Table A.1. We then computed the probability matrix in Table A.2 by normalizing the probabilities across rows so that the probabilities in each row sum to one. This matrix gives the estimates we use for the probability of becoming disabled at a given age for a given sex. We estimate this for males and females separately, and therefore get two sets of estimates.

Accordingly the model for a nondisabled individual is

$$y = \boldsymbol{\beta}(a,c) + \epsilon,$$

and for a disabled individual is

$$y = \boldsymbol{\beta}(a,c) + \sum_{i=65}^{\mathrm{a}} \mathbb{P}_{m,a}(i)\boldsymbol{\gamma}(a,i) + \epsilon.$$

4.2 Generalized Method of Moments Theory

We use the generalized method of moments (GMM) estimation procedure to estimate the model parameters (Hansen 1982). Frequently used by econometricians, this technique does not enjoy the same popularity among statisticians.

GMM is a general estimation procedure. Least squares estimators, method-of-moment estimators for nonlinear models, and generalized least squares estimators, to name a few, can all be derived from a GMM framework and in that sense are special cases of GMM estimation. GMM's generality makes it applicable to a wide range of problems. What makes GMM estimation practical more generally is that its distributional requirements are less stringent than other commonly used methods. For example, GMM estimation easily deals with heteroscedasticity and non-normality (Davidson and MacKinnon 2004).

A convenient and familiar starting point to develop GMM is method of moments (MM) estimation. Consider a model where the observations $Y_1, \ldots, Y_n$ are identically and independently distributed with density $P_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^k$. Equating the first k sample moments to the first k theoretical moments generates a system of k equations which have a unique solution if the moments are in a one-to-one correspondence with the elements of $\boldsymbol{\theta}$.

We apply MM to the simple regression model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad \mathrm{E}[\mathbf{u}\mathbf{u}'] = \boldsymbol{\Sigma}, \mathrm{E}[u_i|\mathbf{x}_i] = 0$$

where $\mathbf{X}$ is $n \times k$, $\mathbf{x}_i$ is the ith row of $\mathbf{X}$, and u_i is the ith element of $\mathbf{u}$. The model implies the theoretical moment conditions $\mathrm{E}[\mathbf{X}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})] = 0$. Corresponding to these theoretical moments are the sample moments $n^{-1}\mathbf{X}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$. Equating these and solving for the parameters $\boldsymbol{\beta}$ yields the familiar least squares estimator.

Now consider using more information than that provided by $\mathbf{X}$ and y to estimate the parameters. Suppose $\mathbf{W}$ is an $n \times l$ matrix, where $l > k$, composed of independent columns that include columns of $\mathbf{X}$ as well as data from additional variables. Assume that $\mathrm{E}[u_i|\mathbf{w}_i'] = 0$, where $\mathbf{w}_i'$ is the row of the $\mathbf{W}$ corresponding to the ith observation. Thus $\mathrm{E}[\mathbf{W}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})] = \mathbf{0}$, a system of l theoretical moments. The problem of solving for $\boldsymbol{\beta}$ now involves selecting k independent linear combinations of sample moments.

To do this, introduce the matrix $\mathbf{J}$, a full rank $l \times k$ matrix. Use the matrix $\mathbf{WJ}$ in place of $\mathbf{W}$ above. This gives $\mathrm{E}[\mathbf{J}'\mathbf{W}'(\mathbf{y} - \mathbf{X}\beta)] = 0$. The MM estimator solves the equations $\mathbf{J}'\mathbf{W}'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})) = 0$. It can be shown that the asymptotic covariance matrix of $\hat{\boldsymbol{\beta}}$ is

$$\left(\operatorname*{plim}_{n\to\infty}\frac{1}{n}\mathbf{J}'\mathbf{W}'\mathbf{X}\right)^{-1}\left(\operatorname*{plim}_{n\to\infty}\frac{1}{n}\mathbf{J}'\mathbf{W}'\boldsymbol{\Sigma}\mathbf{W}\mathbf{J}\right)\left(\operatorname*{plim}_{n\to\infty}\frac{1}{n}\mathbf{J}'\mathbf{W}'\mathbf{X}\right)^{-1},$$

which is not necessarily asymptotically efficient. To find the optimal combination of sample moments, we choose $\mathbf{J}$ to minimize this expression. A suitable $\mathbf{J}$ is

$$\mathbf{J} = (\mathbf{W}'\boldsymbol{\Sigma}\mathbf{W})^{-1}\mathbf{W}'\mathbf{X}.$$

The covariance matrix then becomes

$$\operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\mathbf{X}'\mathbf{W}(\mathbf{W}'\boldsymbol{\Sigma}\mathbf{W})^{-1}\mathbf{W}'\mathbf{X}\right)^{-1}$$

and

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{W}(\mathbf{W}'\boldsymbol{\Sigma}\mathbf{W})^{-1}\mathbf{W}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}(\mathbf{W}'\boldsymbol{\Sigma}\mathbf{W})^{-1}\mathbf{W}'\mathbf{y}.$$

$\mathbf{J}$ effectively chooses a combination of the rows of $\mathbf{W}$ that result in more efficient estimation than in the case where we use $\mathbf{X}$ alone. Notice that if $\mathbf{W} = \mathbf{X}$ and $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$, these reduce to the familiar LS estimators.

For estimation in actual problems, it is convenient to use the fact that the GMM estimator is also the solution to the minimization problem

$$\operatorname*{argmin}_{\boldsymbol{\beta}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{W}(\mathbf{W}'\boldsymbol{\Sigma}\mathbf{W})^{-1}\mathbf{W}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

Computationally, it is often easier to minimize such a criterion function than to solve a system of equations.

Define $f(\mathbf{y}, \mathbf{X}, \boldsymbol{\beta}) = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$. This is one case of an elementary zero function—that is, a function whose expected value is 0 for each observation. GMM works with any such function. Thus $f(\mathbf{y}, \mathbf{X}, \boldsymbol{\beta})$ could be a nonlinear function.

4.3 Elementary Zero Function

Consider the function

$$f(y_i, \mathbf{x}_i, \boldsymbol{\theta}) = \Phi\left(\frac{y_i - g(\mathbf{x}_i, \boldsymbol{\theta})}{s_n}\right) - (1 - q) = U_i \tag{5}$$

where $\Phi(\cdot)$ is the standard normal distribution, s_n is a smoothing parameter fixed at a small number, $g(\mathbf{x}_i, \boldsymbol{\theta})$ is the model, and q is the quantile to be estimated.

We will demonstrate that $\mathrm{E}[U_i] = 0$, and thus that this is a zero function that allows estimation of the parameters at different quantiles of the distribution of the response variable.[1] This is one way of conducting a quantile regression of a nonlinear model.

First we try to provide the intuition behind it through an illustrative example. Suppose we have 100 observations. In the limit as $s_n \to 0$, $\Phi\left(\frac{y_i - g(\mathbf{x}_i, \boldsymbol{\theta})}{s_n}\right)$ returns 0 when the numerator is negative, and 1 when positive. Consider the case where $q = .5$ in equation 5. Estimation of the parameters of this function will find parameter values such that $f(y_i, \mathbf{x}_i, \boldsymbol{\theta}) = .5$ for 50 observations, and $f(y_i, \mathbf{x}_i, \boldsymbol{\theta}) = -.5$ for the remaining observations. That is, the fitted model will generate residuals greater than 0 for exactly half of the observations. In this way it is a median regression. Setting $q = .25$ chooses parameters so that $f(y_i, \mathbf{x}_i, \boldsymbol{\theta}) = .25$ for some observations and $f(y_i, \mathbf{x}_i, \boldsymbol{\theta}) = -.75$ for the remaining observations. Letting m be the number of observations with a positive residual, and solving the equation

$$.25m - .75(100 - m) = 0$$

[1] I thank Jay Bhattacharya and Tom MaCurdy for suggesting this functional form.

gives $m = 75$. This means 75 observations are overestimated and 25 observations are underestimated. That is, it is a quantile regression on the 25th quantile.

The good sense of the choice of a Gaussian (or any) distribution function for the function in Equation 5 is made apparent here. A distribution function of this type differs from using an indicator function in that it is differentiable. That we can choose the arbitrary smoothing parameter s_n ensures numerical differentiability—if the function is too "jagged" to be differentiated, a larger value for s_n will make it smoother. This feature facilitates the optimization problem that is integral to the GMM estimation procedure.

4.4 Demonstrating the Estimation Technique

We conclude this section by empirically demonstrating the suitability of the GMM estimation procedure described. We do so by comparing results from familiar estimation techniques to results from the GMM technique with the specified elementary zero function. Table 3 presents four pairs of estimation results: the first in each pair is from least squares (OLS) or least absolute deviations (LAD) regressions, the other in each pair is from GMM estimation with standard error estimates that are robust to heteroscedasticity. These results come from the estimation of a simpler, linear version of the model of Equation 4. In particular, we use the model

$$y = \beta_0 + \beta_{01}c + \beta_{02}c^2\beta_{10}a + \beta_{20}a^2 + \beta_3 m + \gamma d$$

where c and a are cohort and age, as before. For this simple model, rather than estimating separately for both males and females we close the male indicator variable m.

The first pair of results shows the empirical equivalence of a standard least squares regression and GMM estimation using the elementary zero function

$$f(\mathbf{X}, \boldsymbol{\beta}) = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}.$$

We show in the previous section that, assuming homoscedasticity, this GMM generated identical estimators to least squares regression. These results bear this out.

Additionally, we compare the standard quantile regression (Koenker and Hallock 2001) to the GMM quantile estimation of Equation 4. We present results from the median, the .1 quantile, and the .9 quantile, estimated for males. In each case, results are similar for both estimation methods. Note that at the higher quantiles the precision of the estimates falters. This is unsurprising, as expenditure values are not as compressed in the higher portion of the expenditure distribution of the data.

The similarity (and in the case of ordinary least squares regression, the identity) of the results of the familiar methods and the GMM method serve to establish practical confidence in the GMM estimation procedure we have outlined.

Table 3: Coefficient estimates for comparison of estimation procedures. p-values are in brackets. The OLS column contains ordinary least squares estimates; GMM is generalized method of moments; LAD is least absolute deviations (median regression); Q is quantile regression. The decimal indicates the quantile estimated. Comparing the pairs of results, where pairs are demarcated by the double vertical line, shows that the GMM estimation generates results similar to those from the more familiar estimation techniques. Note at the .9 quantile it is clear that GMM is more efficient than quantile regression.

Variable	OLS	GMM	LAD (Q .5)	GMM .5	Q .1	GMM .1	Q .9	GMM .9
C	-12.1	-12.1	-6.16	-6.75	-36.6	-35.2	-6.40	-3.72
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.017]	[0.000]
COHORT	0.0574	0.0574	0.0464	0.0452	0.0764	0.0782	0.0399	0.0348
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.009]	[0.000]
C2	0.00626	0.00626	0.00832	0.00841	0.0115	0.0102	0.00321	0.00459
	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.353]	[0.000]
A	0.345	0.345	0.203	0.219	0.870	0.834	0.295	0.228
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
A2	-0.00134	-0.00134	-0.000474	-0.000578	-0.00447	-0.00425	-0.00122	-0.000783
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.008]	[0.000]
SEX	0.142	0.142	0.065	0.072	0.349	0.343	0.012	0.011
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.744]	[0.337]
DISAB	0.967	0.967	1.090	1.059	0.878	0.915	0.620	0.689
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]

5 Results and Discussion

5.1 Goodness of Fit of Model

Table 4 gives parameter estimates from the model described in Section 4. That model is:

$$
\begin{aligned}
y =& \beta_{00} + \beta_{01} c \beta_{02} c^2 + \beta_{10} a \beta_{20} a^2 + \beta_{11} c \times a \beta_{12} c^2 \times a + \beta_{21} c \times a^2 \\
&+ \beta_{22} c^2 \times a^2 + \delta \times [\gamma_{10} + \gamma_{11} a] + \Phi(a - p) [\gamma_{20} + \gamma_{21} a] .
\end{aligned}
$$

We provide evidence this is a suitable model with the plot of fitted values versus actual values for a random subset of the data in Figure 4. Nonetheless, several parameteres are insignificant at standard α levels.

Accordingly, we also present estimates from a somewhat pared-down specification, hereafter referred to as the parsimonious model:

$$
y = \beta_{00} + \beta_{01} c + \beta_{02} c^2 + \beta_{11} a + \beta_{12} a^2 + \delta \Big[\gamma_{10} + \Phi(a - p) [\gamma_{20} + \gamma_{21} a] \Big] .
$$

This model has the virtue of more easily interpretable coefficients. Additionally, the estimates are statistically significant at standard α levels, as displayed in Table 5. The positivity of the estimates for the coefficients on cohort and age indicates that expenditures are increasing in those characteristics. The signs of the coefficients on the quadratic terms show that expenditures are rising at an increasing rate for cohort and a decreasing rate in age. Plotting the fitted values for a given disability status and cohort across age makes the results easier to interpret, and we proceed by considering several plots. The plots here are generated by using the parsimonious model, so as to not include explanatory variables that we are not sure are significant determinants of expenditures.

Table 4: Coefficient estimates from the complete model.

	Female		Male	
Variable	Coefficient	p-value	Coefficient	p-value
Constant	20.11	0.060	20.0	0.0130
Cohort	-8.98	0.147	-0.965	0.832
Cohort Squared	-0.114	0.323	-0.423	0.000
Age	-0.398	0.101	-0.404	0.026
Age Squared	0.293	0.034	0.309	0.003
Age × Coh.	1.90	0.192	-0.112	0.916
Age × Coh. Sq.	0.419	0.152	1.22	0.000
Age Sq. × Coh.	-0.909	0.295	0.282	0.648
Age Sq. × Coh. Sq.	-0.330	0.079	-0.823	0.000
γ_{10}	1.45	0.000	1.91	0.000
γ_{20}	6.25	0.000	6.32	0.000
γ_{21}	-0.0815	0.000	-0.090	0.000

Table 5: Coefficient estimates from the parsimonious model.

	Female		Male	
Variable	Coefficient	p-value	Coefficient	p-value
Constant	-9.27	0.000	-4.73	0.000
Cohort	0.676	0.000	0.292	0.000
Cohort Squared	0.00324	0.041	0.0121	0.000
Age	0.280	0.000	0.172	0.000
Age Squared	-0.0972	0.000	-0.0262	0.118
γ_{10}	1.53	0.000	2.09	0.000
γ_{20}	5.52	0.000	4.07	0.000
γ_{21}	-0.0739	0.000	-0.0654	0.000

5.2 Sex Effects

Our results show that males generally spend as much as or more than females, and that life-cycle spending for the two sexes follows a similar pattern. These phenomena are consistent across cohorts and disability statuses, with the exception of the long-term disabled. The expenditure difference is the greatest for earlier cohorts, and reaches a minimum around the 1920 cohort. We show the expected expenditures for the 1900 and 1930 cohort in Figure 5.

Apparent "crossings" of male and female expected expenditures—where female expected expenditures overtake male expected expenditures—generally occur for later cohorts, in areas where we extrapolate outside the data. For example, the 1930 cohort depicted in Figure 5 demonstrates a crossing of the male and female expenditure lines at age 85. However, there are no 85-year-olds in the 1930 cohort in the data set. For observations on such beneficiaries we need to wait until 2015. Similarly, for the 1910 cohort we have no observations before age 80, since our data begin in the year 1992.

There are occasional exceptions that cannot be explained away by appeal to extrapolation outside of the range of observations. Long-term disabled men for all but the earliest cohorts seem to spend less than women. Still, the results generally indicate men spend more than women.

It is notable that the female and male expenditures appear to be very similar. One explanation is that the plots in Figure 5 compare healthy men to healthy women, newly disabled men to newly disabled women, and long-term disabled men to long-term disabled women. If disability is a major determinant of expenditures, or is an indicator of general health status, it is not surprising that healthy 70-year-old men spend comparably to healthy 70-year-old women, or that unhealthy 80-year-old men spend comparably to unhealthy 80-year-old women. To examine this, we estimated a model that controlled only for age and cohort effects. While perhaps slightly more pronounced, the differences in expenditures were similar to those demonstrated in the plots. This suggests that disability-related effects do not drive the differences in expenditures for men and women.

5.3 Cohort Effects

For both male and female Medicare beneficiaries, later cohorts generally spend more than earlier cohorts, as Figure 6 shows. This result suggests that per capita health care costs are rising among the elderly. This is not surprising, given the well-documented fact that U.S. health care costs in general were rising over this period (see, for example, Levit et al. 2003). The rate of increase of expenditures varies from cohort to cohort. For nondisabled and newly disabled beneficiaries, expenditures rise, but less so with each subsequent year of age. For medium- and long-term disabled beneficiaries, health expenditures do not even steadily rise. However, later cohorts still dominate the earlier cohorts.

5.4 Disability Effects

The plots in Figure 7 allow us to compare different disability statuses across age and sex. They consistently show that, at age 65, the longer-term disabled spend more than the newly disabled. These relations persist for each group until a certain age, typically between 75 and 80, at which point the different disability statuses invert their relative ordering, so that now the newly disabled spend more than those who have been disabled for several years, who in turn spend more than those who are long-term disabled.

One interpretation is the following. At early ages, the long-term disabled spend more, but at later ages, the newly disabled spend more. This suggests that the long-term disabled people entering Medicare at age 65 are sicker and therefore need more medical care than the newly disabled people. These relatively sick, long-term disabled die relatively soon. As noted, at a certain age, apparently around age 77, the newly disabled begin to spend more than the remaining long-term disabled. This is presumably because they have suffered some sort of acute event that is expensive to treat and leads to their disablement. A possible story, for example, is that someone breaks a hip, needs surgery, physical therapy, and an initial outlay of funds to acquire a wheelchair and make adjustments to her or his car. If such people survive in a disabled state, their spending remains high, but not as high as it was when they first acquired the disability.

5.5 Quantile Estimations

We also estimated the parsimonious model at the median, .1 and .9 quantiles and present the results in Table 6. The parsimonious model has the benefit of having coefficients that are more easily interpretable than those in the full model.

Again, the signs are as expected for each of the variables. As the coefficients on cohort, age, male, and disabled are easiest to interpret, we restrict our discussion to those estimates. The cohort, age, and male effects vary considerably across the three quantiles, each declining as the quantile increases. The effect of disability, however, increases at higher quantiles. Given the decline in the magnitude of the cohort, age, and male effect, this is unsurprising: at higher expenditures, disability becomes a relatively stronger predictor of expenditure, and other demographic variables become relatively weaker predictors.

This supports the hypothesis that a decline in less severe disability will not lead to a significant decrease in Medicare expenditures if not accompanied by a concurrent decrease in more severe and more costly types of disability. If the number of lower-spending elderly disabled decreases while the population of their higher-spendering peers increases, it is possible for total spending to increase. It is even possible for total spending to increase with declines in both portions of the population if the rate of decline does not outstrip the rate of increase of expenditures. Therefore, a decline

Table 6: Coefficient estimates for the parsimonious model at the .1, .5, and .9 quantiles. p-values are in brackets.

	$q = .1$	$q = .5$	$q = .9$
Constant	-34.3	-5.08	-4.04
	[0.000]	[0.000]	[0.000]
Cohort	0.079	0.041	0.025
	[0.000]	[0.000]	[0.000]
Cohort Sq.	0.0101	0.00929	0.00683
	[.000]	[.000]	[.000]
Age	0.810	0.175	0.237
	[0.000]	[0.000]	[0.000]
Age Sq.	-0.00409	-0.000285	-0.000825
	[0.000]	[0.033]	[0.000]
Male	0.3420	0.0752	0.00982
	[0.000]	[0.000]	[0.396]
Disabled	1.65	1.70	1.73
	[0.000]	[0.000]	[0.000]
γ_{20}	2.03	5.45	4.43
	[0.0980]	[0.000]	[0.000]
γ_{21}	-0.036	-0.076	-0.072
	[0.001]	[0.000]	[0.000]

in the disabled population, if driven by a decline in the category of the less severely disabled, may not lead to lower Medicare expenditures.

Figures 7 and 8 show predicted expenditures for various disability statuses within a given cohort. They include information from Figure 6 as the middle row, but also show predictions derived from .1 quantile and .9 quantile estimations.

The higher and lower quantiles display several features in common with the median results that we have already discussed. For example, the newly disabled typically spend more than their counterparts, and the long-term disabled spend more than the nondisabled. The lower and higher quantiles, however, do not exhibit the same "crossing" that the median predictions did.

It is sensible to assume the lower quantile estimates isolate the less severe—though possibly long-lasting—types of disability and that the higher quantiles isolate more severe disability. Under this assumption, these plots show that the cost of disability, especially new disability, has risen. The increase is especially significant—in levels as as well as percentage points—for more severe types of disability The extent of this increase is obscured by the log scale, so we present the plots for females in untransformed expenditures in Figure 9.

These plots clearly show the increase in expenditures for disabled persons at the higher end of the expenditure distribution. Newly disabled persons demonstrate the highest costs, and these costs have grown with subsequent cohorts. Also, the costs of the medium- and long-term disabled have grown especially fast. Table 7 shows the increase in predicted spending, in dollars, for a disabled 80-year-old male from the 1920 cohort to the 1930 cohort. The increase at the higher quantile is at least 8 times greater for each disability status listed. If the .9 quantile captures the more severe types of disability and the .1 quantile captures the less severe types, the expenditures associated with an increase in the number of severely disabled elderly would require an impressive decrease in the number of less severely disabled elderly.

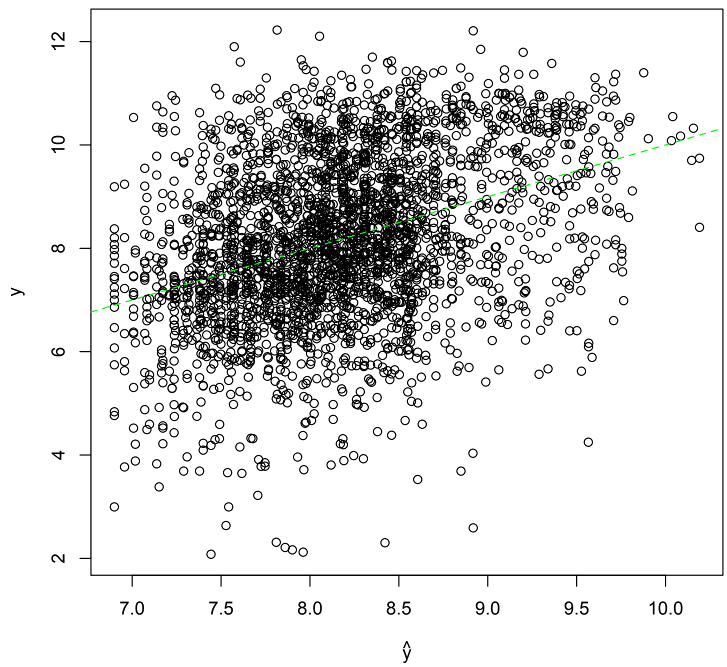

Figure 4: Fitted values versus actual values for the estimation on a random sample of 3,000 males. The overlaid dashed line is the 45 degree line. This gives a sense of the goodness of fit of the model. That is, it shows that the fitted values are nicely correlated with the actual values.

Table 7: Increase in annual expenditures (in dollars) from the 1920 cohort to the 1930 cohort for disability of various durations.

	1 year	3 years	5+ years
0.10 quantile	11,344	5,727	4,498
0.90 quantile	92,577	37,450	27,199

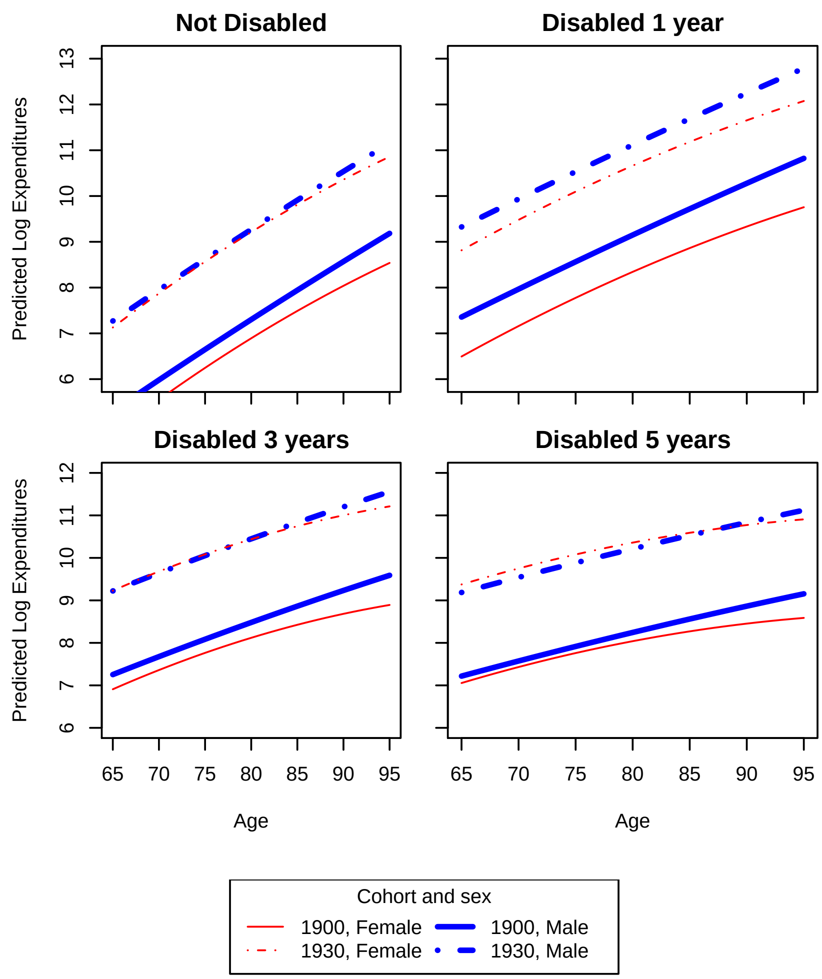

Figure 5: Predicted log expenditures for 1900 and 1930 male and female cohorts, by age, for several disability durations. The intermediary cohorts fall between these two extremes.

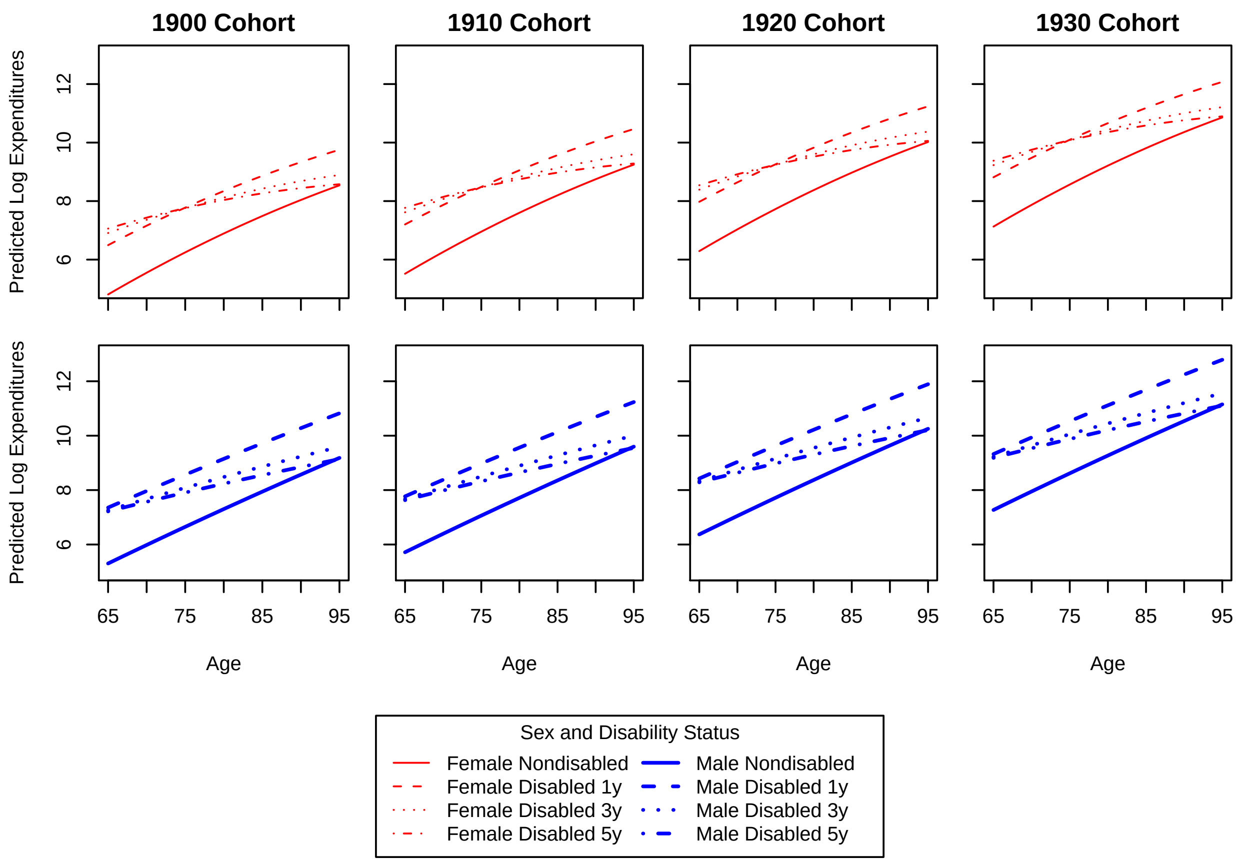

Figure 6: Predicted median log expenditures for various disability statuses, by age, for several different groups.

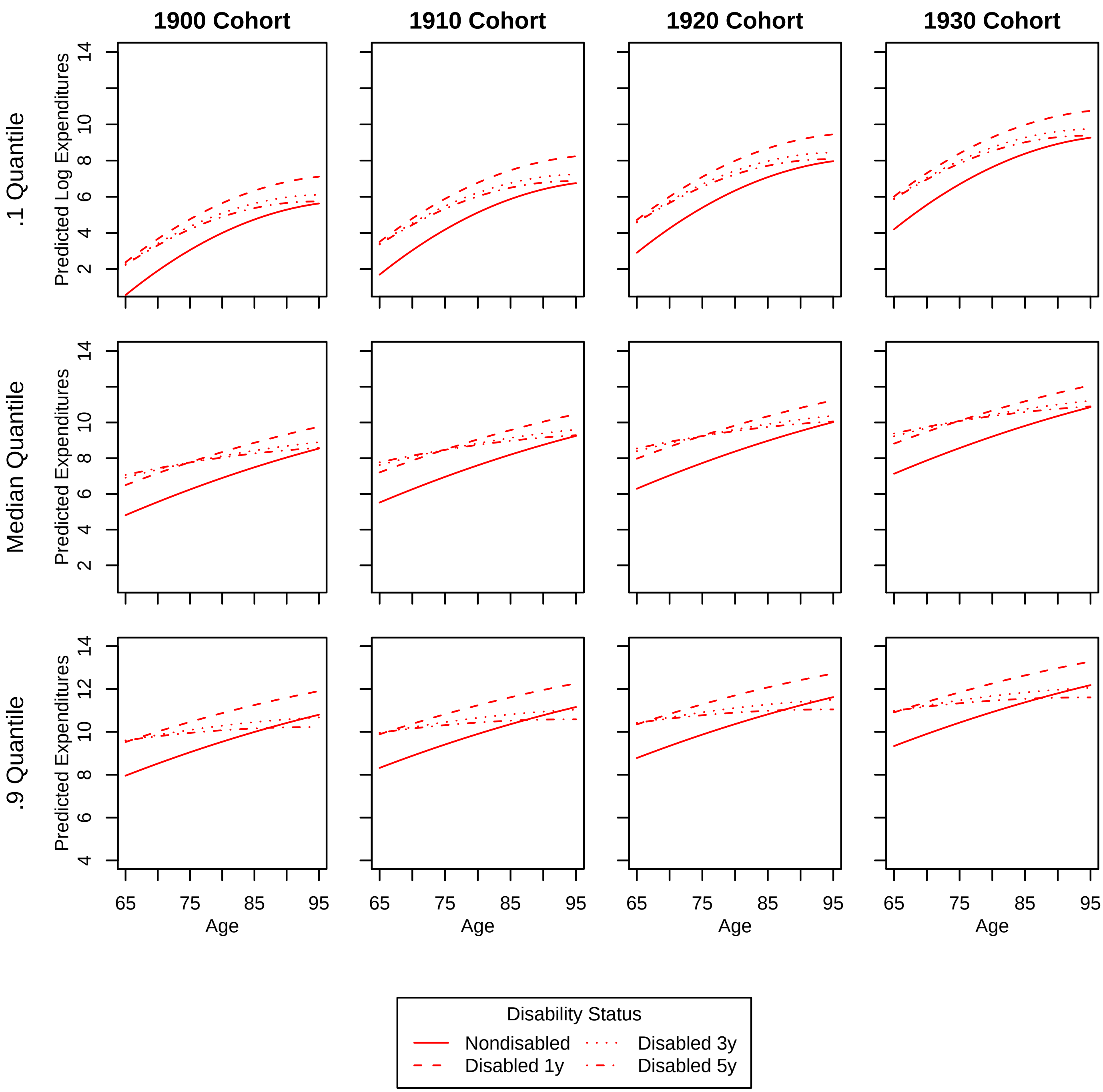

Figure 7: Predicted log expenditures at several quantiles for various female cohorts by age, for several disability statuses. The median plots show that initially, newly-disabled spend less than the long-term disabled. At around age 80, however, the newly disabled begin to spend more.

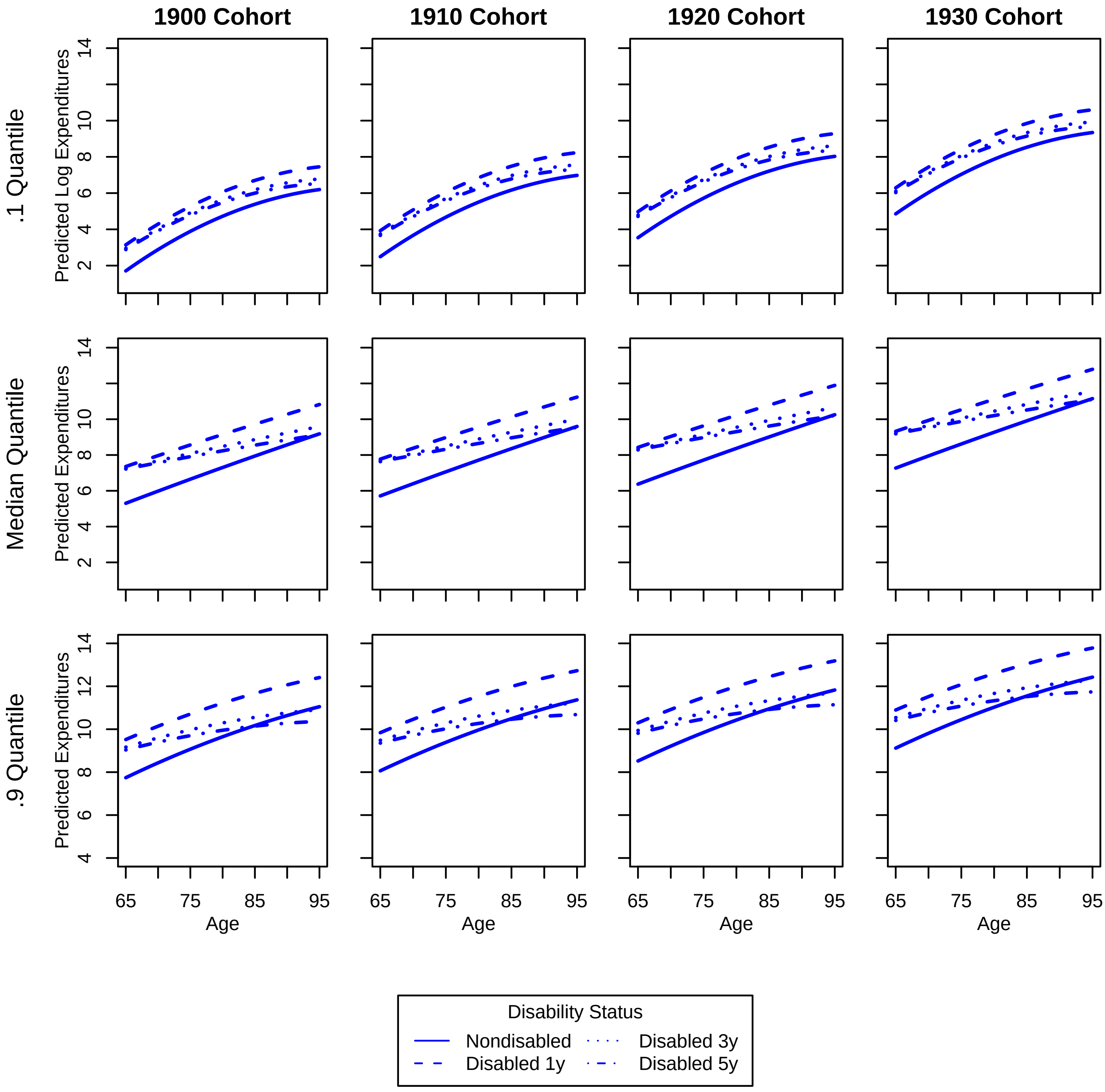

Figure 8: Predicted log expenditures at several quantiles for various male cohorts, by age, for several disability statuses. As for females, the median plots here show that initially, newly-disabled spend less than the long-term disabled. At around age 80, however, the newly disabled begin to spend more.

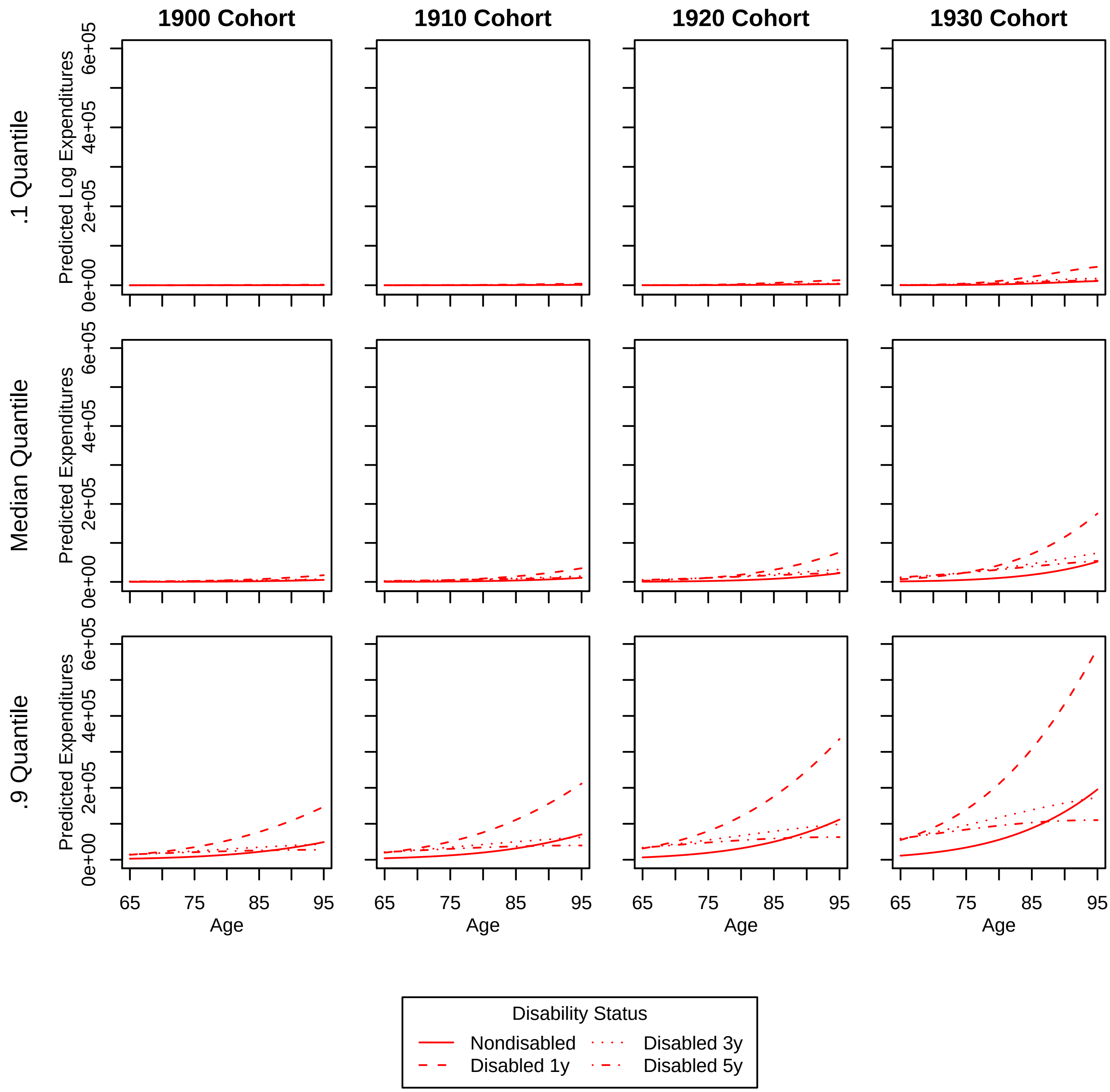

Figure 9: Predicted expenditures for several quantiles by age, for several disability statuses. These plots show the same data as Figure 7, except here expenditures are in dollars, rather than log dollars. They make clear that expenditures have increased tremendously in the higher quantiles of the distribution. Assuming these higher quantiles track more severe types of disability, this suggests that expenditures by the more severely disabled are increasing much faster than by the less severely disabled.

6 Conclusion

Expenditures by the elderly disabled make up a large portion of the Medicare budget. An understanding of the determinants of expenditure of disabled seniors is therefore a matter of significant budgetary importance. Demographic trends indicate there has been a gradual but steady decline in the number of disabled elderly. What this means for future Medicare expenditures is yet unclear.

This paper has advanced understanding of health care expenditures by disabled seniors. Our work is distinguished from other work on the topic as we incorporate a measure of the duration of disability, rather than merely using a binary measure or some measure of the severity of the disability. Using this new approach, we develop a model which we estimate by the generalized method of moments, a general and flexible statistical procedure which we describe.

Using results from the estimated model, we have two especially significant findings. The first is that, beyond a certain age, newly disabled elders spend more than their longer-term disabled counterparts. It appears that disability incurs an initial high period of spending which gradually declines to a still-high level, but not as high as at the onset of the disability.

Second, we find that the increase in expenditures by the disabled is especially high at higher points of the expenditure distribution. This finding encourages us to reconsider the implications of the declining trend in the population size of disabled elderly. As stated in the introduction, some observers have taken this demographic fact to imply that Medicare expenditures on behalf of the disabled elderly will likewise decline. However, our study indicates if the decrease of disabled elderly is amongst those who are on the lower end of the expenditure distribution, the downward trend in the disabled population could fail to decrease total Medicare expenditures on behalf of the disabled elderly. Therefore not only the magnitude, but also the direction, of forecasted Medicare expenditures deserves further study.

BIBLIOGRAPHY

Davidson, R. and MacKinnon, J. G. (2004), *Econometric Theory and Methods*, Oxford University Press.

Elmendorf, D. W. (2009), "Expanding Health Insurance Coverage and Controlling Costs for Health Care before the Senate Committee on the Budget," *Testimony before the Senate*, 1–10.

Freedman, V. A., Crimmins, E., Schoeni, R. F., Spillman, B. C., Aykan, H., Kramarow, E., Land, K., Lubitz, J., Manton, K., Martin, L. G., Shinberg, D., and Waidmann, T. (2004), "Resolving Inconsistencies in Trends in Old-Age Disability: Report from a Technical Working Group," *Demography*, 41, 417–441.

Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 1029–1054.

Koenker, R. and Hallock, K. F. (2001), "Quantile Regression," *Journal of Economic Perspectives*, 15, 143–156.

Levit, K., Smith, C., Cowan, C., Lazenby, H., Sensenig, A., and Catlin, A. (2003), "Resolving Inconsistencies in Trends in Old-Age Disability: Report from a Technical Working Group," *Health Affairs*, 22, 154–164.

Manton, K. G. and Gu, X. (2001), "Changes in the prevalence of chronic disability in the United States black and nonblack population above age 65 from 1982 to 1999," *Proceedings of the National Academy of the Sciences*, 98, 6354–6359.

Manton, K. G., Gu, X., and Lamb, V. L. (2006), "Change in chronic disability from 1982 to 2004/2005 as measured by long-term changes in function and health in the U.S. elderly population," *Proceedings of the National Academy of the Sciences*, 103, 18374–18379.

Mosby (2006), "ADL, IADL," *Mosby's Medical Dictionary.*

National Center for Health Statistics (2009), "Table 26: Life expectancy at birth, at 65 years of age, and at 75 years of age, by race and sex: United States, selected years 1900–2005," Tech. rep.

Orszag, P. (2007), "The Long-term Outlook for Healthcare Spending," *CBO Report*, 1–10.

Spillman, B. C. (2004), "Changes in Elderly Disability Rates and the Implications for Health Care Utilization and Cost," *The Milbank Quarterly*, 82, 157–194.

APPENDIX

A Partial Results from the Estimation of $\mathbb{P}_{m,a}(i)$

Age	Female	Male	Age	Female	Male
65	0.0176	0.0146	81	0.0290	0.0322
66	0.0123	0.0089	82	0.0404	0.0407
67	0.0135	0.0101	83	0.0364	0.0398
68	0.0086	0.0073	84	0.0421	0.0400
69	0.0064	0.0063	85	0.0458	0.0466
70	0.0105	0.0105	86	0.0456	0.0513
71	0.0150	0.0111	87	0.0604	0.0553
72	0.0104	0.0133	88	0.0640	0.0704
73	0.0129	0.0123	89	0.0490	0.0667
74	0.0211	0.0118	90	0.0683	0.0640
75	0.0150	0.0148	91	0.0972	0.0769
76	0.0254	0.0184	92	0.0815	0.0832
77	0.0184	0.0178	93	0.0702	0.0905
78	0.0241	0.0180	94	0.0973	0.1087
79	0.0270	0.0238	95	0.1034	0.1467
80	0.0289	0.0289			

Table A.1: Observed proportion of disabled people with new disability at a given age.

	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
65	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
66	0.21	0.79	0	0	0	0	0	0	0	0	0	0	0	0	0	0
67	0.17	0.62	0.21	0	0	0	0	0	0	0	0	0	0	0	0	0
68	0.13	0.47	0.16	0.23	0	0	0	0	0	0	0	0	0	0	0	0
69	0.1	0.37	0.13	0.18	0.22	0	0	0	0	0	0	0	0	0	0	0
70	0.08	0.28	0.1	0.14	0.17	0.23	0	0	0	0	0	0	0	0	0	0
71	0.06	0.22	0.08	0.11	0.13	0.18	0.21	0	0	0	0	0	0	0	0	0
72	0.05	0.17	0.06	0.09	0.1	0.14	0.16	0.23	0	0	0	0	0	0	0	0
73	0.04	0.15	0.05	0.07	0.09	0.12	0.14	0.2	0.13	0	0	0	0	0	0	0
74	0.03	0.13	0.04	0.06	0.08	0.1	0.12	0.17	0.11	0.15	0	0	0	0	0	0
75	0.03	0.11	0.04	0.06	0.07	0.09	0.11	0.15	0.1	0.14	0.11	0	0	0	0	0
76	0.03	0.1	0.03	0.05	0.06	0.08	0.09	0.13	0.09	0.11	0.09	0.15	0	0	0	0
77	0.02	0.09	0.03	0.04	0.05	0.07	0.08	0.11	0.08	0.1	0.08	0.14	0.11	0	0	0
78	0.02	0.08	0.03	0.04	0.05	0.06	0.07	0.1	0.07	0.09	0.07	0.12	0.09	0.11	0	0
79	0.02	0.07	0.02	0.03	0.04	0.05	0.06	0.09	0.06	0.08	0.06	0.11	0.08	0.1	0.12	0
80	0.02	0.06	0.02	0.03	0.03	0.05	0.05	0.08	0.05	0.07	0.06	0.09	0.07	0.09	0.11	0.12
81	0.01	0.05	0.02	0.02	0.03	0.04	0.05	0.07	0.05	0.06	0.05	0.08	0.06	0.08	0.09	0.11
82	0.01	0.04	0.02	0.02	0.03	0.04	0.04	0.06	0.04	0.05	0.04	0.07	0.05	0.07	0.08	0.09
83	0.01	0.04	0.01	0.02	0.02	0.03	0.04	0.05	0.04	0.05	0.04	0.06	0.05	0.06	0.07	0.08
84	0.01	0.04	0.01	0.02	0.02	0.03	0.03	0.05	0.03	0.04	0.03	0.06	0.04	0.05	0.07	0.08
85	0.01	0.03	0.01	0.02	0.02	0.03	0.03	0.04	0.03	0.04	0.03	0.05	0.04	0.05	0.06	0.07
86	0.01	0.03	0.01	0.01	0.02	0.02	0.03	0.04	0.03	0.03	0.03	0.05	0.04	0.04	0.05	0.06
87	0.01	0.03	0.01	0.01	0.02	0.02	0.02	0.03	0.02	0.03	0.02	0.04	0.03	0.04	0.05	0.05
88	0.01	0.02	0.01	0.01	0.01	0.02	0.02	0.03	0.02	0.03	0.02	0.04	0.03	0.03	0.04	0.05
89	0.01	0.02	0.01	0.01	0.01	0.02	0.02	0.03	0.02	0.03	0.02	0.03	0.03	0.03	0.04	0.04
90	0.01	0.02	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.03	0.03	0.04
91	0	0.02	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.03	0.03	0.04
92	0	0.02	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.02	0.02	0.03	0.02	0.02	0.03	0.03
93	0	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.02	0.01	0.02	0.02	0.02	0.03	0.03
94	0	0.01	0	0.01	0.01	0.01	0.01	0.02	0.01	0.02	0.01	0.02	0.02	0.02	0.02	0.03
95	0	0.01	0	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.02	0.02	0.02

Table A.2: Matrix of probabilities of a disabled male of a given age (column) having acquired his disability at a given age (row).