SUBJECTIVE HEALTH ASSESSMENTS AND ACTIVE LABOR MARKET PARTICIPATION OF OLDER MEN: EVIDENCE FROM A SEMIPARAMETRIC BINARY CHOICE MODEL WITH NONADDITIVE CORRELATED INDIVIDUAL-SPECIFIC EFFECTS

Jürgen Maurer, Roger Klein, and Francis Vella*

Abstract—We use panel data from the U.S. Health and Retirement Study, 1992–2002, to estimate the effect of self-assessed health limitations on the active labor market participation of older men. Self-assessments of health are likely to be endogenous to labor supply due to justification bias and individual-specific heterogeneity in subjective evaluations. We address both concerns. We propose a semiparametric binary choice procedure that incorporates nonadditive correlated individual-specific effects. Our estimation strategy identifies and estimates the average partial effects of health and functioning on labor market participation. The results indicate that poor health plays a major role in labor market exit decisions.

I. Introduction

The ratio of individuals aged 65 years and older to those aged under 65 years in the United States is projected to increase from 0.14 in 2005 to approximately 0.26 in 2050.1 This increased dependency ratio is likely to produce severe budgetary pressure as the economy endeavors to cover the cost of an aging population, which has a substantially lower propensity to be in formal employment. Figure 1 presents the age profile of active labor market participation rates, defined as currently employed, and age-specific prevalence rates of major and minor health conditions for male respondents from the U.S. Health and Retirement Study (HRS), 1992–2002.2 Although approximately 85% of those 51 to 55 years old actively participate in the labor market, this rate declines sharply as individuals progress through their 50s and 60s, before falling below 40% for those aged 66 to 70 years. The fraction reporting to have been diagnosed with a major health condition increases from 13% to 53% over the age range 51 to 70 years, while prevalence rates for minor health conditions increase from 50% to over 80%.

Empirical evidence has established the role of financial incentives associated with the Social Security system in the labor market exit decisions of middle-aged workers (Stock & Wise, 1990; Rust & Phelan, 1997; French, 2005). The decline in health and functioning is also likely to have implications for the labor supply, although evaluating its impact is not straightforward. This is partially due to the difficulty in measuring individual health and functioning. While a comprehensive set of biomarkers combined with objective physical assessments is ideal, their collection is costly and difficult. Accordingly, most empirical work employs survey data and self-reported health measures. In such a setting, both the endogeneity and the self-assessed nature of the responses are potential problems due to justification bias and the subjective nature of the response scales.

Justification bias arises when an individual’s propensity to report a work-limiting health condition depends on his labor market status. For example, healthy individuals may self-report work disability to justify their labor market inactivity. This bias can be reduced through the use of more objective self-reported assessments of health and functioning on specific domains (Dwyer & Mitchell, 1999). These measures include self-reports on mobility or large muscle activities, such as self-reported difficulties walking several blocks or getting up from a chair (Fonda & Herzog, 2004). Their use, although somewhat controversial (Baker, Stable, & Deri, 2004), has become common in the absence of biomarkers, objective medical records or reliable health assessments by health care professionals (see Currie & Madrian, 1999).

An additional concern is the individual-specific response heterogeneity in self-assessments, which results in subjective response scales (Kerkhofs & Lindeboom, 1995; Lindeboom & van Doorslaer, 2004; Jürges, 2007). For example, respondents with varying levels of mobility may report similar levels of difficulty walking due to their interpretations of “difficulty.” One strategy to restore comparability is the use of vignettes (Tandon, Saloman, & King, 2001, or Kapteyn, Smith, & van Soest, 2007) which portray the same state of health of hypothetical persons to all respondents, who are then asked to rate specific aspects of these hypothetical persons’ health. These assessments can be used to translate subjective health assessments to a common scale. The administration of vignettes, however, is costly, and they are often not available. An alternative approach applicable in panel data is to account for the subjectivity of the responses to a condition on a control function, computed as some appropriate function of the explanatory variables, which summarizes the within-unit distribution of responses for each individual. Health effects are then identified by within-unit variation in responses over time.

This paper evaluates the impact of health on labor force participation while accounting for justification bias and subjective response scales. We employ self-reported domain-specific measures of health and functioning and model subjective response scales as nonseparable multidimensional

Received for publication December 20, 2007. Revision accepted for publication March 4, 2010.

* Maurer: Institute of Health Economics and Management and Faculty of Business and Economics, University of Lausanne. Email: jürgen.maurer@unil.ch; Klein: Rutgers University; Vella: Georgetown University.

1 See U.S. Census Bureau (2004).

2 Following Smith (2003), the prevalence of any major condition is measured by a binary variable indicating the presence of lung disease, cancer (but not skin cancer), heart attack, or stroke. Having a minor condition indicates the prevalence of hypertension, diabetes, psychological problems, or arthritis.

individual-specific effects. We exploit repeated observations for each individual to construct a control function, which accounts for the dependence between the self-reported health assessments and any subjective responses scales by capturing the individual-specific variability in self-reported health and functioning measures. We define structural effects of health on labor market participation when self-reported health assessments feature individual-specific response heterogeneity. We state assumptions under which these structural effects can be identified and estimated using panel data and provide estimates based on the HRS. The following section describes our econometric strategy. Section III introduces the data and provides the model specification. Section IV presents the estimation results, and section V concludes.

II. Econometric Strategy

A. Model

We focus on the extensive margin of the employment decision of individual \( i \) at time \( t \). We consider data on individual \( i \) for \( T \) periods (\( T \) fixed) and model labor market participation as

\[
E[Y_{it} | X_{it}, \eta_{it}] = g(X_{it} \theta, \eta_{it}),
\]

where \( Y_{it} \) is an indicator function denoting whether the individual is employed; \( X_{it} \equiv (D_{it}, H_{it}) \) denotes the explanatory variables consisting of demographic variables \( D_{it} \) and self-reported health controls \( H_{it} \); and \( \eta_{it} \equiv (\alpha_i, \varepsilon_{it}) \) represents a potentially multidimensional error term consisting of a vector of time-invariant unobservable individual-specific effects \( \alpha_i \) and a time-varying unobservable error term \( \varepsilon_{it} \). This error term may display some serial correlation, but it is assumed to be stationary with finite variance, ruling out unit roots in the error structure. Similarly, all distributions are assumed to be stationary, such that corresponding expectations do not change over time. The time-invariant unknown function \( g(\cdot, \cdot) \) maps the respondent’s observable and unobservable characteristics \( (X_{it}, \eta_{it}) \) into observed employment status \( Y_{it} \), assuming a linear index structure with unknown parameter \( \theta \) for the \( X_{it} \). The model allows time-invariant subjective response scales in the measurement of \( H_{it} \) by including individual-specific effects \( \alpha_i \). These individual-specific effects need not be separable from the effects operating through \( X_{it} \).

Our main interest is how individual health assessments \( H_{it} \) affect active labor market participation \( Y_{it} \). We model \( Y_{it} \) as a binary choice model that features contemporaneous health assessments \( H_{it} \) and other demographic variables \( D_{it} \) as explanatory variables. The model thus not only allows labor market exits, but also potential (re-)entries into the labor force, which have been shown to be quantitatively important (Maestas, 2007). The model, however, does not incorporate a lagged dependent variable and treats labor market exits and entries symmetrically. While these assumptions are
restrictive, and thus a potential limitation of our approach, the evidence regarding the symmetry of health effects on retirement appears reassuring (Disney, Emmerson, & Wakefield, 2006). Also, our model does not address potential reverse causation from work status to health, as most recent papers find little, if any, evidence of retirement effects on health (see Neumann, 2008; Coe & Lindeboom, 2008).

An important issue that our model addresses is the dependence between self-reported health \(H_i\) and the response scales \(\alpha_i\), and the possibility each enters the model in a nonseparable way. Specifically, individuals with a higher propensity to report health limitations will, ceteris paribus, report worse health, resulting in a mechanical dependence between \(H_i\) and \(\alpha_i\). Moreover, \(H_i\) and \(\alpha_i\) might interact in explaining labor market participation. For example, the impact of self-reported poor health may be smaller for individuals who report health limitations more easily. As a consequence, the model should incorporate the dependence between \(X_i\) and \(\eta_i\) and be nonseparable in structure.

Altonji and Matzkin (2005; hereafter AM) study a class of panel data models like equation (1), and our approach to identification and estimation of our empirical model features similarities to theirs. Like AM, we invoke a conditional independence assumption based on a control function to estimate the parameters of interest.\(^5\) Our main assumption is:

**Assumption. Control Function:** There exists a control function \(Z_i \gamma\), such that \(X_i \equiv (X_{i1}, \ldots, X_{iT})\) and \(\eta_i\) are conditionally independent given \(Z_i \gamma\), that is,

\[
(X_i \perp \eta_i) | Z_i \gamma,
\]

(2)

with \(Z_i\) a vector of known transformations of \(X_i\) and \(\gamma\) an unknown parameter.

This assumption postulates conditional independence between the \(X_i\) and \(\eta_i\) given an appropriately specified one-dimensional control function \(Z_i \gamma\). Here, \(Z_i \gamma\) is assumed to control for the dependence between the time-invariant response scales \(\alpha_i\) and the corresponding self-assessments of health and functioning \(H_i\).

Our choice of control function is guided by the idea of exchangeability forwarded in AM, which states that the conditional distribution of the composite error term \(\eta_i|X_{i1}, \ldots, X_{iT}\) is independent of the order of the \(X_{ij}\). While this assumption excludes time-varying response heterogeneity, it holds if response scales are modeled as time-invariant individual effects. Importantly, exchangeability restricts the control function to consist of symmetric functions of \(X_{i1}, \ldots, X_{iT}\), such as the within-unit means or variances of \(X_{ij}\).\(^4\) In our application, \(Z_i\) will consist of within-unit means as simple measures of location and spread, as most of the self-assessed health measures we use are binary. Similar strategies, in parametric settings, are Mundlak (1978) and Chamberlain (1984).

This control function approach complements the vignettes methodology described above, as the two approaches rely on different assumptions about reporting behavior and feature different data needs. Vignettes can be used with cross-sectional data sets, provided there are suitable vignette data for each of the relevant domains of self-assessed health. Our approach does not need external data but requires repeated observations on each respondent’s self-reported health measures. The control function approach assumes that response scales can be treated as time-invariant individual effects, while vignettes assume that individuals use the same response scales for evaluating the health of hypothetical people as they use for evaluating their own health.

### B. Object of Interest

Our focus is on the structural dependence between health and employment while accounting for potentially differential response behavior. Using equation (2), write the conditional probability of \(Y_{it}\) given \(X_i\), as

\[
\Pr(Y_{it} = 1|X_i) = E[Y_{it}|X_i] = E[g(X_{it}, \eta_{it})|X_i]
\]

\[
= \int g(X_{it}, \eta_{it}) dF_{\eta_{it}|X_i}
\]

\[
= \int g(X_{it}, \eta_{it}) dF_{\eta_{it}|Z_i \gamma}
\]

\[
\equiv H^*(X_{it}, Z_i \gamma).
\]

(3)

where \(H^*(\cdot, \cdot)\) denotes an unknown function of the indices \(X_{it}\) and \(Z_i \gamma\), and \(F_{\eta_{it}|Z_i \gamma}\) denotes the distribution of \(\eta_{it}\) conditional on \(Z_i \gamma\). Equation (3) implies that the relationship between the index \(X_{it}\) and the dependent variable \(Y_{it}\) depends on the control function \(Z_i \gamma\). Thus, the effects of a self-assessed functional limitation, for example, will generally depend on the response scales. We define the effect of a self-assessed limitation as its expected impact on labor force participation for an individual randomly drawn from the population. This definition is related to existing approaches to evaluating partial effects in the presence of unobserved heterogeneity (see, for example, Chamberlain, 1984) and is similar to the average structural function (ASF) of Blundell and Powell (2003, 2004). Accordingly, we employ the term average structural function, which is also based on integration with respect to the marginal distribution of the control function, to describe the object we report here.

\(^{5}\) Other examples of control function estimation in semi- and nonparametric settings are Newey, Powell, and Vella (1999), Imbens and Newey (2007), and Blundell and Powell (2003, 2004).

\(^{4}\) AM conjecture that “in actual panel data applications with exchangeability, conditioning on one or two \(z\) functions to capturing the location of \(x_i\) (such as the average of the elements of \(x_i\)) and the dispersion of the elements of \(x_i\) (such as the variance) will be sufficient to eliminate most of the relationship between the unobservable terms and \(x_{it}\)” (p. 1079).
We define the ASF for a specific realization of the demographic and health controls \( x^0 \) as

\[
\mu(x^0) = \int g(x^0, \eta_i) dF_{\eta_i} = \int \int g(x^0, \eta_i) dF_{\eta_i} dF_{\gamma} = \int H^*(x^0, \gamma) dF_{\gamma},
\]

where the \( F \) denote cumulative distribution functions. The ASF corresponds to the expected probability of observing \( \{ Y_i = 1 \} \) given \( x^0 \), but replacing the conditional distribution of the error term \( \eta_i \) given \( x^0 \) by the marginal distribution of \( \eta_i \) when taking expectations. The ASF thus summarizes the average structural relationship between \( x \) and \( Y \), with the average taken with respect to the (marginal) population distribution of response scales. Differences in the ASF for the treated, whereas changes in the ASF resemble average treatment effects.

While our modeling approach is similar to AM, the ASF is conceptually distinct from their local average response (LAR). The LAR represents the average derivative of the dependent variable with respect to the explanatory variables \( \tau \) (that is, \( \mu(x') - \mu(x) \)), can therefore be interpreted as the expected effect on \( Y \) of changing \( X \) from \( x' \) to \( x'' \) for a randomly selected respondent.

The ASF of interest here as it represents the average effect of self-assessed limitations of health and functioning on labor market participation for the entire population rather than for a specific subpopulation defined by actually reporting a particular realization \( x\). In addition, the ASF is convenient in that it requires integration only with respect to the marginal distribution of the control function. It thus avoids the more demanding estimation of the conditional density of the control function, given the high dimension of the regressors, which would be required to obtain the LAR.

### C. Estimation

Evaluation of the ASF requires estimates of \( \theta, \gamma \) and \( H^*(\cdot, \cdot) \). Given the double index binary choice structure in equation (3), we employ the semiparametric maximum likelihood procedure proposed in Klein and Vella (2009). Defining

\[
P_\mu(\theta, \gamma) \equiv \Pr(Y_i = 1|X_i) \equiv H^*(X_i, \theta, \gamma),
\]

the likelihood function can be written

\[
L(\theta, \gamma) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} I_i(\theta, \gamma),
\]

where

\[
I_i(\theta, \gamma) \equiv \tau_i(Y_iLn[P_\mu(\theta, \gamma)] + [1 - Y_i]Ln[1 - P_\mu(\theta, \gamma)]),
\]

and \( \tau \) denotes a trimming function, which is defined below.

Following Klein and Vella (2009), we represent \( P_\mu(\theta, \gamma) \) as

\[
P_\mu(\theta, \gamma) \equiv H^*(X_i\theta, \gamma) = \frac{f_1(X_i\theta, \gamma)}{(f_0(X_i\theta, \gamma) + f_1(X_i\theta, \gamma))},
\]

where \( f_1(X_i\theta, \gamma) \) is defined as \( f_1(X_i\theta, \gamma) = P_q \cdot d_q(X_i\theta, \gamma) \), with \( P_q \) denoting the unconditional probability of \( Y_i = q, q \in \{0, 1\} \) and \( d_q(X_i\theta, \gamma) \) denoting the joint density of \( (X_i\theta, \gamma) \) conditional on \( Y_i = q, q \in \{0, 1\} \), respectively. A quasi-likelihood function can be constructed by replacing the true \( f_1(X_i\theta, \gamma) \) in equation (6) by corresponding estimates. The index parameters \( \theta, \gamma \) are estimated by maximizing the following quasi-likelihood,

\[
\tilde{\theta}, \tilde{\gamma} = \arg \max_{\theta, \gamma} \tilde{L}(\theta, \gamma) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{I}_i(\theta, \gamma),
\]

with

\[
\tilde{I}_i(\theta, \gamma) = \tilde{\tau}_i(Y_iLn[\tilde{P}_\mu(\theta, \gamma)] + [1 - Y_i]Ln[1 - \tilde{P}_\mu(\theta, \gamma)]),
\]

and

\[
\tilde{P}_\mu(\theta, \gamma) = \frac{\tilde{H}^*(X_i\theta, \gamma)}{(\tilde{f}_0(X_i\theta, \gamma) + \tilde{f}_1(X_i\theta, \gamma))},
\]

where the hats denote estimates. Details regarding the implementation of the estimator are provided in the appendix to this paper and Klein and Vella (2009).

With estimated index coefficients, we compute the ASF at a particular value \( x^0 \) as

\[
\tilde{ASF}(x^0) = \frac{1}{N} \sum_{i=1}^{N} \tilde{H}^*(x^0\theta, \gamma),
\]

where the average is taken with respect to the marginal distribution of the estimated control function \( \gamma \).

\(^5\text{The asymptotic distribution of the resulting ASF estimate is stated in the appendix.}\)
III. Data and Model Specification

Our empirical analysis employs the 1992–2002 waves of the HRS as compiled by the RAND Corporation (RAND, 2004). The HRS consists of a nationally representative sample of around 7,600 households (12,654 individuals) with at least one household member born in the years 1931 to 1941. The data collection began in 1992 and comprises information from biannual follow-ups. It contains extensive demographic information and measures of health, financial position, and labor market status. Because we focus on the employment decisions of older men, we select all men from the original sample of reference persons. We use only age-eligible individuals with no missing observations on any variable included in the model over all six waves, except those for which the RAND files already provided imputations. This sample selection produces a balanced panel consisting of 1,809 men, each observed six times.

A. Outcome of Interest

While the concept of active labor market participation appears straightforward, different institutional definitions and arrangements make its practical measurement more difficult. Distinct definitions of labor market exit or retirement are often not compatible with less institutional definitions of labor supply, such as hours worked or direct questions focusing on participation. We follow some earlier literature (Disney et al., 2006) and concentrate on whether the respondent is currently working for pay. This measure is not contaminated by specific institutional arrangements, such as the claiming of Social Security benefits, and is easily understood. Moreover, its use has the additional advantage that our model can incorporate “unretirement.” This may be important, as recent evidence for the United States indicates that reentry into the labor force is quantitatively important for workers of older ages (see, for example, Maestas, 2007). This choice of outcome also circumvents other definitional issues that arise from different institutional arrangements, such as a detailed distinction between disability, retirement, and unemployment, which often reflects differences in program eligibility rather than effective labor supply decisions (Bound & Burkhauser, 1999).

B. Explanatory Variables

We construct the structural index with contemporaneous variables capturing the individual’s current health status and other demographic characteristics. The control index consists of within-unit location measures for the subset of subjective assessments that feature in the structural index.

Structural index. The explanatory variables of primary interest here are the contemporaneous measures of health status and functional limitations. We use multiple quasi-objective health measures to account for the multidimensional character of personal health. That is, we use the prevalence of doctor-diagnosed major and minor conditions as a first set of quasi-objective health measures. To directly capture potential heterogeneity in the functional associations of these health conditions, we use three additional indicator variables representing the presence of any mobility limitation (walking one block, walking several blocks, walking across a room, climbing one flight of stairs, and climbing several flights of stairs), any limitation in large muscle functioning (sitting for two hours, getting up from a chair, stooping, kneeling or crouching, and pushing or pulling large objects) and any limitation in activities of daily living (ADL) (bathing, dressing, eating, getting out of the bed, or walking across a room), respectively. We also include an indicator for the current prevalence of depression based on the respondent’s scoring on the Centre for Epidemiological Studies Depression Scale (CES-D). Specifically, we classify a respondent as suffering from depression if he scores four or more items on the abbreviated eight-item CES-D administered in the HRS.

The choice of health controls, and the questions on which they are based, are guided by four considerations: (a) no reference to the labor market, (b) quasi-objectivity, (c) differentiation of severity, and (d) inclusion of both physical and mental health. Accordingly, our measures focus on well-defined domains of health and functioning rather than labor supply. This is likely to restrict the presence of justification bias. Our measures of health and functioning also provide a reasonable degree of differentiation in severity. Finally, we incorporate mental health as a potential determinant of labor force withdrawal (Conti, Berndt, & Frank, 2006).

Our depression measure illustrates our approach to modeling functioning and mental health. We measure disability due to depression using the respondent’s scores on the abbreviated CES-D rather than a question directly asking whether any form of depression limits the amount or type of work that the respondent can do. By avoiding a direct reference to the respondent’s labor market status in the administration of the eight CES-D items, we reduce justification bias. However, even responses on rather subjective survey items like those constituting the CES-D renders our measure vulnerable to subjective response scales. We thus use the overall individual-specific average CES-D score over the entire sample period in constructing the control function to account for individual-specific propensities to score on the CES-D items.

To capture economic or family-related incentives for retirement, we include age, race, and marital status controls in our model. We also include measures of educational attainment and occupational status. Finally, the model contains a measure of the labor supply and self-rated health of the spouse. The age controls capture the potentially confounding Social Security incentives associated with labor force withdrawal at ages 62 and 65 years. At those ages, the number of labor force exits tends to feature peaks beyond flexibly specified age trends, which seem largely due to inherent incentives

---

6 See Gustman and Steinmeier (2000) for a more detailed exploration of the notion of retirement based on the HRS.
in the Social Security system, such as pension eligibility age, the tax treatment of pensions, discontinuities in pension accrual, actuarial considerations, or the illiquidity of pension wealth. Similar to our controls for marital status, spouse’s labor supply, and self-rated health should capture confounding (nonmonetary) incentives for joint retirement or differences in need of care of the partner. The last is thereby proxied by the overall self-rated health of the partner measured on a (potentially subjective) scale ranging from 1 (excellent) to 5 (poor).

Table 1 provides summary statistics for the dependent variable and the explanatory variables in the structural index for all six waves. Our sample covers the age range between 50 and 71 years. Blacks and Hispanics comprise 13% and 7%, respectively, of our sample. Furthermore, there is substantial variation in the level of attained education, and approximately half of the sample report their longest occupation as manual. The table also indicates substantial prevalence rates of various health conditions and functional limitations among middle-aged men. Around two-thirds of the respondents have experienced a minor health condition, and more than a quarter report a major health event. Approximately a quarter of the respondents report at least one mobility limitation, 41% some functional limitations regarding large muscle activities, and 6% report being severely disabled, that is, limited in at least one ADL. Finally, around 5% of the respondents appear to suffer from serious depressive symptoms, as captured by the abbreviated CES-D.

Table 1.—Summary Statistics for the Pooled HRS Data, 1992–2002

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>s.d.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currently working for pay (binary)</td>
<td>0.66</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Age</td>
<td>60.30</td>
<td>4.62</td>
<td>50</td>
<td>71</td>
</tr>
<tr>
<td>Age ≥ 62 (binary)</td>
<td>0.40</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Age ≥ 65 (binary)</td>
<td>0.20</td>
<td>0.40</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Black (binary)</td>
<td>0.13</td>
<td>0.33</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic (binary)</td>
<td>0.07</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Years of education</td>
<td>12.75</td>
<td>3.18</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>Manual occupation (binary)</td>
<td>0.50</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Any major health condition (binary)</td>
<td>0.26</td>
<td>0.44</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Any minor health condition (binary)</td>
<td>0.66</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Any mobility limitation (binary)</td>
<td>0.25</td>
<td>0.43</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Any large muscle activity limitation (binary)</td>
<td>0.41</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Any ADL limitation (binary)</td>
<td>0.06</td>
<td>0.24</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Depression: CESD score &gt; 4 (binary)</td>
<td>0.05</td>
<td>0.22</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Respondent married or partnered (binary)</td>
<td>0.84</td>
<td>0.37</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Respondent’s partner working (binary)</td>
<td>0.47</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Respondent’s partner’s self-rated health (ordinal)</td>
<td>2.06</td>
<td>1.33</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Number of respondents</td>
<td>1,809</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of time periods</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Control function. The role of the control function is to account for any potentially confounding effects of individual-specific response scales. Accordingly, we include the individual-specific means of whether the respondent is suffering any limitation in mobility, large muscle functioning, or ADL, as well as whether the respondent suffers from depressive symptoms as indicated by a CES-D score higher than 4. Finally, we include the within-unit average of the spouse’s self-rated health. The use of within-unit means of the controls can be motivated by the binary nature of most of the explanatory variables.

IV. Results

Table 2 presents the coefficient estimates and their respective standard errors, with each index identified up to location and scale. For normalization, we excluded intercepts from the two indices and set the coefficients of age/10 (structural index) and average self-rated health of the spouse (control function) equal to 1. Because the function that maps the estimated indices into the outcome probabilities is not parametrically specified, we cannot infer the size, or even direction, of any estimated effects by simple inspection of the parameter estimates. Focusing on statistical significance, we find some of the sociodemographic controls, notably age, marital status, and labor market status of the spouse (if

Table 2.—Parameter Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural index</td>
<td>1.00</td>
<td>—</td>
</tr>
<tr>
<td>Age/10</td>
<td>−0.06***</td>
<td>0.011</td>
</tr>
<tr>
<td>Age ≥ 62 squared</td>
<td>0.09**</td>
<td>0.038</td>
</tr>
<tr>
<td>Age ≥ 65</td>
<td>0.02</td>
<td>0.025</td>
</tr>
<tr>
<td>Black</td>
<td>0.02</td>
<td>0.022</td>
</tr>
<tr>
<td>Hispanic</td>
<td>−0.08</td>
<td>0.049</td>
</tr>
<tr>
<td>Years of education</td>
<td>0.01</td>
<td>0.148</td>
</tr>
<tr>
<td>Years of education2</td>
<td>−0.04</td>
<td>0.057</td>
</tr>
<tr>
<td>Longest occupation manual</td>
<td>0.03</td>
<td>0.021</td>
</tr>
<tr>
<td>Any major health condition</td>
<td>0.10**</td>
<td>0.043</td>
</tr>
<tr>
<td>Any minor health condition</td>
<td>0.03</td>
<td>0.017</td>
</tr>
<tr>
<td>Any mobility limitation</td>
<td>0.08**</td>
<td>0.042</td>
</tr>
<tr>
<td>Any large muscle activity limitation</td>
<td>0.03</td>
<td>0.021</td>
</tr>
<tr>
<td>Any ADL limitation</td>
<td>0.03</td>
<td>0.045</td>
</tr>
<tr>
<td>Depression: CESD score &gt; 4</td>
<td>0.14**</td>
<td>0.071</td>
</tr>
<tr>
<td>Respondent married or partnered</td>
<td>0.11**</td>
<td>0.050</td>
</tr>
<tr>
<td>Respondent’s partner working</td>
<td>−0.22**</td>
<td>0.089</td>
</tr>
<tr>
<td>Respondent’s partner’s self-rated health</td>
<td>0.00</td>
<td>0.008</td>
</tr>
</tbody>
</table>

Control function

Average: Respondent’s partner’s self-rated health | 1.00 | — |
Average: Any mobility limitation                | −3.42***          | 0.427 |
Average: Any large muscle activity limitation   | −0.07             | 0.463 |
Average: Any ADL limitation                     | −3.60***          | 0.562 |
Average: Depression: CES-D score > 4            | 1.14**            | 0.534 |

*, **, and *** indicate variables that appear statistically significant at the 10%, 5%, and 1% level, respectively.
present), enter the model statistically significantly. In addition, there are statistically significant effects for some of the health controls: the presence of any major health condition, any self-assessed mobility limitation, and the presence of self-reported depressive symptoms as measured by the eight-item CES-D scales. Several components of the control function also enter the model statistically significantly. We find statistically significant coefficients for the within-unit averages of self-assessed mobility, ADL limitations, and the indicator based on self-reported depressive symptoms from the CES-D.

Figure 2 presents the mapping between the two estimated indices and the probability of active labor market participation. Given the age normalization and the estimated positive index coefficients of our health limitation variables for the structural index, we would expect the probability of active labor market participation to generally decline with increasing structural index values. This is consistent with figure 2. Overall, the estimated probabilities display substantial variation over the support of the data, ranging from values larger than .9 to values smaller than .3. The estimated control function does not appear to have a large effect on the slope of the estimated probability surface with respect to the structural index for most of its support but highlights some heterogeneity in its upper tail. Conditional on the structural index, the control function also has a nontrivial effect on the level of the probability surface, with lower values of the control function generally leading to lower levels of active labor market participation. These findings imply that respondents who persistently report mobility or ADL limitations, or both, typically feature a lower probability of working for pay ceteris paribus, whereas the reverse is true for people more persistently showing signs of depression. In addition, the slope of the probability of working for pay in the structural index seems somewhat steeper for smaller control function values, implying that changes in the structural index have a larger effect on labor force participation among respondents who persistently report mobility or ADL limitations, or both, and a smaller effect for persons with persistent signs of depression. These findings are consistent with multiple explanations, including more persistent reporting of mobility or ADL limitations among persons with more serious forms of these limitations, and lower reporting thresholds for depressive feelings and behavior among persons who report those feelings and behaviors more frequently.9

Figure 3 presents estimates for the ASF and its 95% pointwise confidence interval. Similar to the double index probability plots, the ASF is downward sloping in the structural index. The ASF estimates decrease from a value of around .9 for a structural index value of 3.5 to just over .3 for a structural index of 4.5. Moreover, the ASF appears to flatten at higher values of the structural index, which is consistent with multiple exit routes from labor market participation. Finally, the ASF estimates seem reasonably precise for most structural index values, as the pointwise confidence intervals are fairly tight except in the tails.

Table 3 presents estimates of the ASF for different configurations of the explanatory variables to illustrate the average structural dependence between active labor market participation and selected individual explanatory variables in the structural index. We focus on different configurations of the health controls and different ages and living arrangements, as these characteristics entered the model statistically significantly. The remaining variables of the structural index are held fixed at their respective sample means. The table summarizes the structural dependence of active labor market participation, living arrangements, and health for respondents ages 60, 63, and 66, respectively. With regard to health, we first consider the impact of different physical health states in the absence of depressive symptoms. We thereby label physical health as “perfect” if the respondent does not suffer from any health condition or functional limitation, as “good” if he suffers from a minor health condition only, as “fair” if he has both a major and a minor health condition, and as “poor” if he suffers from a major and a minor health condition, as well as functional limitations with respect to both mobility and large muscle activities. We also consider the case of mental health problems for respondents who are in perfect physical health and suffer from depression only, as well as for respondents who are in “poor physical health” who suffer from depression as an additional comorbidity.

The table indicates that the effects of health on active labor market participation are economically significant. An exogenous change from perfect to poor physical health reduces the probability of active labor market participation by 11 to 19

---

9 As a robustness check, we constructed the control function using only observations from waves 2 to 5. The results were quantitatively similar to those presented here.
### Table 3: Average Structural Function for Selected Health States and Demographic Characteristics

<table>
<thead>
<tr>
<th>Partnership Status</th>
<th>Physical Health</th>
<th></th>
<th>Mental Health Depression</th>
<th>Poor Physical Health and Depression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Perfect</td>
<td>Good</td>
<td>Fair</td>
</tr>
<tr>
<td>Age 60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>0.74</td>
<td>0.72</td>
<td>0.65</td>
<td>0.56</td>
</tr>
<tr>
<td>Partner not working</td>
<td>0.80</td>
<td>0.79</td>
<td>0.73</td>
<td>0.64</td>
</tr>
<tr>
<td>Partner working</td>
<td>0.66</td>
<td>0.64</td>
<td>0.56</td>
<td>0.47</td>
</tr>
<tr>
<td>Partner not working and in poor health</td>
<td>0.66</td>
<td>0.64</td>
<td>0.56</td>
<td>0.47</td>
</tr>
<tr>
<td>Age 63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>0.61</td>
<td>0.59</td>
<td>0.51</td>
<td>0.43</td>
</tr>
<tr>
<td>Partner not working</td>
<td>0.70</td>
<td>0.68</td>
<td>0.60</td>
<td>0.51</td>
</tr>
<tr>
<td>Partner working</td>
<td>0.53</td>
<td>0.51</td>
<td>0.44</td>
<td>0.37</td>
</tr>
<tr>
<td>Partner not working and in poor health</td>
<td>0.52</td>
<td>0.50</td>
<td>0.43</td>
<td>0.37</td>
</tr>
<tr>
<td>Age 66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>0.53</td>
<td>0.51</td>
<td>0.44</td>
<td>0.38</td>
</tr>
<tr>
<td>Partner not working</td>
<td>0.62</td>
<td>0.60</td>
<td>0.52</td>
<td>0.44</td>
</tr>
<tr>
<td>Partner working</td>
<td>0.45</td>
<td>0.44</td>
<td>0.38</td>
<td>0.34</td>
</tr>
<tr>
<td>Partner not working and in poor health</td>
<td>0.45</td>
<td>0.43</td>
<td>0.38</td>
<td>0.34</td>
</tr>
</tbody>
</table>

ASF point estimates with standard errors in parentheses. The first four columns (Physical Health) present ASF estimates for different physical health states in the absence any mental health limitation (depression). “Perfect” physical health indicates the absence of any physical health condition or functional limitation. “Good” physical health refers to the presence of a minor health condition only, whereas “fair” physical health indicates the presence of both a minor and a major health condition. “Poor” physical health refers to the presence of a major and a minor health condition, as well as a mobility and a large muscle activity limitation. Column 5 Mental Health refers to a health state characterized by the presence of depression but an absence of any physical health conditions (“perfect” physical health). Column 6 Poor Physical Health and Depression refers to a state of poor physical health coupled with the presence of depression. Education, occupation, and race are evaluated at their respective sample means for all estimates presented in this table.
of labor market exits among older American men. Thus, while financial incentives to delay retirement may be very effective among those in good health, we would still expect early labor market exits by individuals for whom early retirement represents a constraint rather than a choice.

REFERENCES


Bound, J., and R. Burkhauser, “Economic Analysis of Transfer Programs for People with Disability” (pp. 3417–3528), in D. Card and O. Ashenfelter (Eds.), Handbook of Labor Economics, vol. 3c (Amsterdam: North-Holland, 1999).


APPENDIX

This appendix describes the implementation of the estimator and its asymptotic distribution. The discussion follows Klein and Vella (2009), which also contains the relevant proofs.10

1. Construction of the Semiparametric Quasi-Likelihood

The form of the likelihood function is straightforward, and the complications are due to the bias-reducing method that we employ. We follow Klein and Vella (2009) and use local smoothing. The following discussion is related to the various stages of estimation in the presence of local smoothing.

1.1. Density Estimators under Local Smoothing. Let $L$ be a symmetric, smooth univariate kernel function satisfying condition C8 in Klein and Spady (1993, p. 394). We use a normal kernel that satisfies this condition. Let $W$ be a matrix such that $W_{ij} = \Sigma_{ij}$, the inverse of the sample covariance matrix of $X_i$ and $X_j$, $i \neq j$, for observations with $Y_i = q$ for $q = 0$ and $q = 1$, respectively. Partitioning $W_{ij} = \begin{bmatrix} W_{ii} & W_{ij} \\ W_{ji} & W_{jj} \end{bmatrix}$ conformally with $(X_i, Z_i, Y_i)$, we define

$$k_{ij}^2(v, h, \lambda) = \frac{\text{det}(\Sigma_{y_i})^{-1}}{(\lambda h)^3} K\left(\frac{W_{ii}(v - Z_i)}{\lambda h}\right) K\left(\frac{W_{ii}(v - Y_i)}{\lambda h}\right).$$

We then define estimators for $\hat{f}(v) = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{1 - Y_{ij}} k_{ij}^2(v, h, \lambda)$, $\hat{f}_k(v) = \frac{1}{N} \sum_{i=1}^{N} k_{ij}^2(v, h, \lambda)$.

1.2. Smooth trimming functions. Define a smooth trimming function as

$$\tau(r, \alpha) = \left[1 + \exp(NT^\alpha |r|)\right]^{-1}.$$

1.3. Estimated local smoothing parameters. Referring to section 1.1, denote $\hat{\eta}_0$ as the geometric mean of $\hat{f}(v, h, \lambda)$ with $q = 0$ and $q = 1$, respectively, and let $\hat{\eta}_q = \frac{\hat{f}(v, h, \lambda)}{\hat{\eta}_0}$. Define local smoothing parameters as

$$\hat{\eta}_q = \frac{\hat{f}(v, h, \lambda)}{\hat{\eta}_0} = \left[1 + \frac{(\hat{f}_k(v) - 1)/\ln (NT) - \hat{\eta}_q \cdot \hat{\eta}_q}{\hat{\eta}_q \cdot \hat{\eta}_q} \right]^{-1/2},$$

where the smoothed indicator $\hat{d}_q$ is given by

$$\hat{d}_q = \tau\left(\frac{1}{\ln (NT)} - \hat{\eta}_q \cdot \hat{\eta}_q, 0.01\right),$$

where the parameter $\alpha$ in section 1.2 is set to 0.1.

1.4. Three-stage local smoothing. Employing section 1.3, the estimator for $f(v)$ under three-stage local smoothing is defined as

$$\hat{f}_q(v) = \frac{1}{N^2} \sum_{i=1}^{N} Y_{ij} k_{ij}^2(v, h, \lambda),$$

$$\hat{f}_q(v, h, \lambda) = \left[1 + \frac{(\hat{f}_k(v) - 1)/\ln (NT) - \hat{\eta}_q \cdot \hat{\eta}_q}{\hat{\eta}_q \cdot \hat{\eta}_q} \right]^{-1/2},$$

where $1$ denotes a vector of ones and $h_k, k \in \{1, 2, 3\}$ bandwidth components with $h = O(N^{-\epsilon})$, setting $r_3 = \frac{1}{\lambda}$ and $0 < \delta < \frac{1}{2}$ and then setting $r_2 = (r_3 - \delta/2)/2$ and $r_1 = (r_3 - \delta/4)$.

1.5. Semiparametric probability function. Define

$$\hat{P}(v, \gamma) = \frac{\hat{f}_q(v)}{\hat{f}_q(v) + \hat{d}_q} = \frac{\hat{f}_q(v) + \Delta_{\alpha \lambda} \gamma}{\hat{f}_q(v) + \Delta_{\alpha \lambda} \gamma + \Delta_{\alpha \lambda}},$$

with $\hat{\gamma}(v) = \frac{\hat{f}(v) + \hat{d}(v)}{\hat{f}(v) + \hat{d}(v) + \Delta_{\alpha \lambda}}$ representing the unconditional density of $v$. The $\Delta_{\alpha \lambda} \gamma$ are smooth adjustment factors defined as

$$\Delta_{\alpha \lambda} \gamma = \hat{\gamma}(v) \gamma [1 + \exp((\hat{f}_q(v) - (\hat{\eta}_q \cdot \hat{\eta}_q))]^{-1},$$

where $\alpha = 0.02$, $\gamma = 0.02$, $\hat{\gamma} = O(1)$, small $\epsilon > 0$, and $\Delta_{\alpha \lambda} = \hat{\Delta}_{\alpha \lambda} + \Delta_{\alpha \lambda}.$

1.6. Pilot estimation. Let $p_0$ be the lower $\alpha$ quantile of $Y_i$ (we used $\alpha = 0.2$) for the continuous variable $X^2$, and let $\hat{p}_0$ be the upper $(1-\alpha)$-th sample quantile. For all continuous variables $X^2$ (which in this example is age only), define the indicators: $\tilde{p}_0 \equiv I(X^2 < x_0)$. In the notation of section 1.1, define a pilot probability estimator as

$$\hat{P}_0(\phi) = \frac{\hat{f}_q(v, 1/11, 1)}{\hat{f}_q(v, 1/11, 1) + \hat{f}_q(v, 1/11, 1)},$$

with $\phi = (0, \gamma)$. With the initial (non-smooth) trimming function $\hat{\eta}_q = \Pi_{j=0}^{\phi}$, the pilot estimator for $\phi = (0, \gamma)$ is defined as

$$\hat{\phi}_0 = \arg \max \hat{L}_0(\phi),$$

with

$$\hat{L}_0(\phi) = \sum_{i=1}^{N} \log(\hat{P}_0(\phi)) + (1 - Y_i) \log(1 - \hat{P}_0(\phi)).$$
1.7. Final estimator. With \( \hat{\phi}_0 = (\hat{\theta}_0, \hat{\gamma}_0) \) defined in section 1.6, let \( \hat{V}_0 = (\hat{V}_{10}, \hat{V}_{20}) \equiv (\hat{X}_0 \hat{\theta}_0, \hat{Z}_0 \hat{\gamma}_0) \) denote the pilot estimates of structural index and control function, respectively. Denote \( V_j \) as the lower \( j \)th sample quantile (we used \( \beta = 0.01 \)) of \( V_{10} \), \( k \in \{1, 2\} \), and let \( \tau \) be the corresponding \((1 - \beta)\)th sample quantile. Construct the smooth index trimming function:

\[
\hat{\tau}_k = \hat{\tau}_{10} \cdot \hat{\tau}_{20},
\]

with

\[
\hat{\tau}_{10} = \tau \left( \hat{V}_{10} - V_1, \frac{1}{12} \right) \cdot \tau \left( \hat{V}_{10} - V_2, \frac{1}{12} \right)
\]

for \( k \in \{1, 2\} \) with \( \tau \) being the smooth trimming function defined in section 1.2. With semiparametric probabilities defined as in section 1.5, a final estimator for the index coefficients \((\hat{\theta}, \hat{\gamma})\) can be defined as

\[
\hat{\phi} = \arg \max_\theta \hat{L}(\hat{\phi})
\]

with

\[
\hat{L}(\phi) \equiv \sum_{i, j} \hat{\rho}_i(\phi) \equiv \sum_{i, j} \hat{\tau}_k \left[ \hat{Y}_i \hat{L}_n(\hat{P}_i(\phi)) + (1 - \hat{Y}_i) \hat{L}_n(1 - \hat{P}_i(\phi)) \right].
\]

2. Asymptotic Distribution of the Index Parameters

The panel data estimator leads to within-unit serial correlation in the scores \( \hat{\rho}_i(\phi) = \hat{V}_0 \hat{\rho}_i(\phi) \) of the (estimated) quasi-likelihood function. We thus require a robust variance estimator for asymptotic inference. As in standard parametric models, we can account for the serial correlation in the scores of the estimated quasi-likelihood function by including cross-products in the relevant components of the asymptotic variance. This adjustment is valid since our estimated quasi-likelihood function can be treated as if it were a known likelihood function for inference (Klein & Spady, 1993). The estimates for the two index parameters therefore converge at \( \sqrt{N} \) and are asymptotically normal distributed. Therefore:

\[
\sqrt{N}(\hat{\phi} - \phi_0) \xrightarrow{d} N(0, \Omega),
\]

where a robust estimator of the asymptotic variance \( \Omega \) is given by

\[
\Omega(\hat{\phi}) = A(\hat{\phi})^{-1} B(\hat{\phi}, \hat{\gamma}) A(\hat{\phi})^{-1},
\]

with

\[
A(\hat{\theta}, \hat{\gamma}) = N^{-1} \sum_{i=1}^N \sum_{t=1}^T \hat{\gamma}_i(\hat{\phi}) \hat{\gamma}_i(\hat{\phi})'
\]

\[
B(\hat{\theta}, \hat{\gamma}) = N^{-1} \sum_{i=1}^N \sum_{t=1}^T \hat{\gamma}_i(\hat{\phi}) \hat{\gamma}_i(\hat{\phi})'
\]

and \( B(\hat{\phi}) \) accounting for the serial correlation in the scores (Wooldridge, 2001).

3. Computation of the Average Structural Function

With an estimate of \( \hat{P}_i(\hat{\theta}, \hat{\gamma}) \equiv \hat{H}_i^*(x_i \hat{\theta}, \hat{Z}_i \hat{\gamma}) \), we estimate the ASF at \( \hat{x}_0 \) by averaging \( \hat{H}_i^*(x_i \hat{\theta}, \hat{Z}_i \hat{\gamma}) \) over the estimated marginal density of \( \hat{Z}_i \).

Define this estimated marginal as

\[
\hat{f}_Z(t_2) \equiv \frac{1}{NTH} K[(t_2 - \hat{Z}_i \hat{\gamma})/h].
\]

The estimate for the ASF for given demographic and health characteristics \( x_0 \) in the structural index can then be obtained as

\[
\hat{ASF}(x_0) \equiv \int \hat{H}_i^*(x_0 \hat{\theta} - \hat{Z}_i \hat{\gamma}) d\hat{f}_Z(t_2).
\]

4. Asymptotic Distribution of the Average Structural Function

The asymptotic distribution of the estimator for the ASF is

\[
\sqrt{hNT}(\hat{ASF}(x_0) - ASF(x_0)) \xrightarrow{d} N(0, \sigma^2).
\]

A corresponding estimator of the variance \( \sigma^2 \) can be computed as

\[
\hat{\sigma}^2 = \sum_{i=1}^N \left[ Y_i \hat{H}_i^*(x_i \hat{\theta}, \hat{Z}_i \hat{\gamma}) \right]^2 \frac{1}{h} K[(x_i \hat{\theta} - x_i \hat{\theta})/h] \hat{w}_i^2,
\]

with

\[
\hat{w}_i = \left[ \hat{f}_Z(\hat{Z}_i \hat{\gamma}) / \hat{f}(x_i \hat{\theta}, \hat{Z}_i \hat{\gamma}) \right]
\]

and

\[
\hat{f}(t_1, t_2) \equiv \frac{1}{NTH} K[(t_1 - x_i \hat{\theta})/h] K[(t_2 - \hat{Z}_i \hat{\gamma})/h].
\]

11 The proof of these results is available from the authors on request.