

The Welfare Cost of Late-life Depression

Ray Miller* Sayorn Chin[†] Ashish Kumar Sedai[†]

May 11, 2021

Abstract

We quantify the welfare cost of depression among older Americans by estimating a panel VAR model of mental and physical health, labor supply, and consumption using data from the Health and Retirement Study. We use the estimated model and age sixty joint distribution of outcomes to simulate life-cycle paths with and without prevalence of depressive symptoms after age sixty. We estimate that the prevalence of late-life depression costs an average of between 0.85 and 2.1 years in quality-adjusted life expectancy per person. Moreover, depression may result in an average loss of labor supply of up to 1.1 months and lifetime consumption of up to \$16,000. Combining into a single compensating variation welfare metric, we estimate a bound on the average welfare cost of depression of 8-15% of annual consumption after age sixty. On aggregate, this amounts to roughly \$180-360 billion annually. We also project that while the average welfare cost of late-life depression is declining slightly over birth cohorts, the welfare burden is becoming significantly more unequal.

JEL classifications: I14, J14, J11, J26

Keywords: depression, mental health, aging, retirement, consumption, cost-utility analysis

*Assistant Professor, Colorado State University, Department of Economics, Fort Collins, CO, 80523. Email: ray.miller@colostate.edu

[†]PhD Candidate, Colorado State University, Department of Economics, Fort Collins, CO, 80523.

1 Introduction

Depression is one of the leading causes of emotional distress and lower quality of life among older adults (Blazer, 2003; Sivertsen et al., 2015). Depression and depressive symptoms are also highly correlated with other physical and psychiatric conditions in older populations (Moussavi et al., 2007; Vaughan et al., 2015; Soysal et al., 2017; Chu et al., 2019). Increasing depressive symptoms with age have also been shown to be predictive of an increased risk of mortality (Bruce et al., 1994; Chui et al., 2015). Nonetheless, under-treatment of depression remains prevalent in older populations despite the wide availability of effective treatments (Barry et al., 2012; Kok and Reynolds, 2017). Improving our understanding of the comprehensive costs of late-life depression may be a fruitful avenue for expanding uptake of effective antidepressants and treatment therapies.

While preventing and treating late-life depression is of major social importance in its own right, significant spillover benefits are also possible. Empirical studies have found depression to be related to increased risk of frailty, reduced mobility, functional limitations, and progression of chronic diseases (Stuck et al., 1999; Penninx et al., 1999; Ciechanowski et al., 2000; De Groot et al., 2001; Geerlings et al., 2001; Rubio-Guerra et al., 2013; Vaughan et al., 2015; Chiang et al., 2015; Soysal et al., 2017; Penninx, 2017; Lwin et al., 2020). This has led some researchers to hypothesize a causal link from depression to poor physical health in older adults. However, whether the association between depression and physical health is driven by reciprocal influences or common causes remains widely debated (Mayerl et al., 2020).

Many theoretical explanations for why depression would affect physical health have been proposed (Penninx et al., 1999; Bruce, 2001). It could be that depressive symptoms such as sleep disturbance or lost appetite have a direct effect on functional decline and disability. There could also be indirect effects through intermediate behaviors (Bruce, 2001). For example, depressive symptoms could reduce motivation and lead to reduced medical care or poor health behaviors (e.g., smoking, poor nutrition, reduced physical activity). Other proposed mechanisms include antidepressant use (Lakey et al., 2012), increased allostatic load (McEwen, 2003), or other neuronal, hormonal, and/or immunological alterations (Bruce, 2001).

Beyond health effects, late-life depression could also influence an individual’s economic outcomes. For example, depression among older adults increases health service utilization and costs (Luppa et al., 2012). Standard consumer theory suggests this could have a negative contemporaneous effect on consumption expenditures. Consumption could also decrease with depression due to reduced productivity and earnings (Lerner and Henke, 2008) or even decreased utility from goods and services that are complements to good mental health. On the other hand, the life-cycle hypothesis suggests that an unexpected depressive episode could *increase* contemporaneous consumption if there is an associated decline in life expectancy. Moreover, there could be additional dynamic effects that persist over time, for example if depression leads to an early retirement (Doshi et al., 2008; Rice et al., 2011). In the presence of such dynamic

effects, cross-sectional correlations between depression and other health and economic outcomes would only reveal part of the larger story.

In this paper, we adopt a life-cycle approach to better quantify the welfare cost of late-life depression when incorporating persistence and dynamic spillover effects. We extend the panel VAR model proposed by Miller and Bairoliya (2021) to simulate the joint evolution of health and economic outcomes, adapted to include the onset and persistence of late-life depression. We estimate the model using longitudinal data from the Health and Retirement Study (HRS) spanning more than twenty years. Using the observed joint distribution of outcomes at age sixty as initial conditions, we show that model simulations are able to closely match the empirically observed evolution of depressive symptoms, physical health, labor supply, and consumption.

Equipped with our simulation model, we next estimate the welfare cost of late-life depression. As the causal relationships between depressive symptoms and other health and economic outcomes remains unsettled, we take a bounds analysis approach. First, we estimate a lower bound on welfare costs by assuming there are no spillover effects on other model outcomes. More specifically, we leave all expected paths of comorbidities, mortality, and economic outcomes at their baseline levels and only remove the health utility penalty associated with depression at age sixty-two and older. We follow this with an upper bound estimate calculated by running a second set of counterfactual simulations starting from the same initial conditions but removing any possibility of depressive symptoms after age sixty. We consider this an upper bound as it assumes all the statistical relationships estimated in the restricted VAR model are entirely causal.

These analyses provide bounds on the expected costs of late-life depression in terms of quality-adjusted life years (QALYs), labor years, and dollars of consumption. We combine these differing costs using standard expected utility theory by calculating an ex-ante compensating variation (CV) measure of welfare. The welfare concept is akin to asking how much an individual would be willing to pay at age sixty to avoid any possibility of depressive symptoms over their remaining life. As our measure integrates multiple health and economic outcomes, it gives a more comprehensive view of well-being loss than the direct utility cost of depression alone. As it incorporates individual expectations over the entirety of remaining life from age sixty, it also provides a useful single metric of the ex-ante welfare cost of late-life depression.

1.1 Contributions

This study makes several contributions to our understanding of the welfare or utility burden of depression in older adults. First, previous studies have focused on estimating lost quality of life in older populations using cross-sectional observation, clinical settings, and/or limited longitudinal data (Sivertsen et al., 2015). Our estimates capture both contemporaneous and dynamic spill over effects on the evolution of depression, health, and economic outcomes over the entirety of remaining life. This provides a more complete measure of the total welfare burden of depression as it incorporates the cumulative burden of disease over time. We also provide an estimate that combines the

impact of depression on health-related quality-of-life, leisure, and consumption into a single measure grounded in economic and public health theory. Moreover, as our simulations are at the individual level within a larger representative sample, we are able to examine the entire distribution of welfare as opposed to only specific sub-samples or summary aggregates. This approach also allows us to examine how the level and distribution of welfare costs changed over birth cohorts, as opposed to cross-sectional changes over time.

Finally, we also contribute to the literature that has attempted to estimate the economic burden of depression to society. Studies have examined the impact of depression on direct medical costs and indirect workplace costs, including absenteeism from work and presenteeism while at work (Wang et al., 2003; Stewart et al., 2003; Lerner and Henke, 2008; Birnbaum et al., 2010; Luppia et al., 2012). Combined with suicide-related mortality costs, Greenberg et al. (2015) estimate the economic burden of major depression disorders in the U.S. was \$210.5 billion in 2010. While these costs center on the employer or healthcare side, we complement these studies by incorporating economic costs to private individuals. We also focus on older adults and quantify effects from the full range of depressive symptoms as opposed to only major disorders.

2 Data and methods

2.1 Data

The HRS is an ongoing longitudinal survey of U.S. individuals over the age of fifty and their spouses. The survey began in 1992 and data has since been collected every two years with new birth cohorts added periodically. There are currently eight birth cohorts in study—the early HRS cohort (born 1931-36), late HRS cohort (born 1937-41), AHEAD cohort (born before 1924), Children of Depression (born 1924-30), War Babies (born 1942-47), early Baby Boomers (born 1948-53), mid-Baby Boomers (born 1954-59), and late-Baby Boomers (born 1960-65). We use the publicly available RAND HRS Longitudinal File 2016 (V2) to obtain data on depression, health, mortality, and economic outcomes from 1992 to 2016. We also utilize other individual characteristics including age, education, gender, race, birth cohort, region, and occupation.

2.1.1 Depression

Depressive symptoms in the HRS were measured using the eight-item Center for Epidemiologic Studies Depression scale (CESD). The measure ranges from zero (no depressive symptoms) to eight, created by summing the respondent’s number of “yes” answers across eight survey items (with positive items reverse-coded).¹ The CESD is a common measure of depressive symptoms in older adults (Lewinsohn et al., 1997;

¹About 12% of observations in our estimation sample were missing CESD score. These were imputed along with consumption and other missing data as detailed in the online appendix.

Turvey et al., 1999; Steffick, 2000; Karim et al., 2015). The CESD was designed to measure a continuum of psychological distress (symptoms of depression), rather than determining the presence or absence of specific psychiatric disorders. However, a longer form CESD scale has been broadly validated against diagnostic interviews for depression and other anxiety disorders (Fechner-Bates et al., 1994; Lewinsohn et al., 1997). The eight-item CESD has also been shown to be a valid and reliable instrument of depression in a large sample of older Europeans (Karim et al., 2015).

2.1.2 Additional health outcomes

In addition to depression, our model incorporates data on comorbidities. These include eight binary indicators for ever having been diagnosed by a doctor with the following health problems—(1) high blood pressure or hypertension; (2) diabetes or high blood sugar; (3) cancer or a malignant tumor of any kind except skin cancer; (4) chronic lung disease except asthma such as chronic bronchitis or emphysema; (5) heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems; (6) stroke or transient ischemic attack (TIA); (7) emotional, nervous, or psychiatric problems; and (8) arthritis or rheumatism. We also include an indicator for ever reported difficulty with any activity of daily living (ADL). Difficulty with ADLs are a commonly used health metric among the elderly and include activities such as walking across the room, bathing, and getting dressed.

As a final health measure we use self-rated health status reported on a five-point scale from poor (one) to excellent (five). Self-rated health has been shown to be predictive of mortality in the HRS and other datasets, even after controlling for other health conditions, health behavior, and socioeconomic characteristics (Idler and Benyamini, 1997; Stenholm et al., 2014). This may reflect that people have private information about their health over and above disease diagnosis.

2.1.3 Economic outcomes

As our empirical focus is on individuals nearing the end of working life, we limit labor considerations to retirement.² We combine data on weekly hours worked and weeks worked per year to estimate annual hours worked. We define retired individuals as those reporting less than 500 annual hours of work in the current or any previous survey wave.

Consumption data comes from the Consumption and Activities Mail Survey (CAMS). Starting in 2001, the CAMS was sent to a random sub-sample of HRS participants during off years of the core survey. The CAMS collected household spending data on durables, nondurables, transportation, and housing. We use RAND 2017 CAMS

²An extended model may include the intensive margin, partial retirement, and/or reentry into the workforce but this comes with additional model complexity. Moreover, we find relatively small effects of leisure on welfare in our empirical analysis and retirement is likely to be the first-order leisure effect for this age group.

data file (V1), which contains a constructed estimate of total household consumption derived from the available spending data from 2001-2015. We subtract out-of-pocket health spending to create an adjusted measure of household consumption.³ We use this adjusted household consumption divided by the number of household members as our individual consumption measure. As consumption data is only available between the core HRS waves, we merge each CAMS wave with the HRS core data from the previous wave.⁴

A challenge to our analysis is that CAMS data is only available for approximately 20% of HRS respondents for the years 2000-2014. We follow Miller and Bairoliya (2021) and use closely related available data such as wealth and income to address missing consumption data by using the multiple imputation method proposed by Honaker and King (2010) for cross-sectional time-series data (see online appendix for details).

2.2 Simulation model

We extend and estimate the forecasting model proposed by Miller and Bairoliya (2021), adapted to include the onset and persistence of late-life depression. The model follows the structure of a panel vector autoregression (VAR) making it useful for microsimulations. Specifically, we use the model to repeatedly simulate potential outcome paths for each individual with and without the prevalence of late-life depression, given a set of initial (age sixty) conditions. Here we discuss the basic structure of the model and identifying assumptions. The online appendix provides additional details on sample selection, descriptive statistics, and model estimation procedures and results.

The general structure of the simulation model is illustrated in Figure 1. At the beginning of each time period, morbidity status is updated based on (correlated) random shocks, which in turn influence an individual’s self-rated health status. Morbidities and self-rated health then contemporaneously influence an individual’s reported depression status.⁵ We choose this outcome sequence because 1) it is consistent with evidence that general health affects depression (Moussavi et al., 2007; Ambresin et al., 2014); 2) it allows block identification of the system for estimation (details below); and 3) it provides a more conservative estimate of the welfare cost of depression. On the last point, there is quite plausibly some contemporaneous reverse causation between depression and general health (Rothermund and Brandtstädter, 2003; Moussavi et al., 2007). However, our later counterfactual simulations will assume that depression does *not* influence current period general health, yielding the more conservative

³Health spending includes health insurance, medication, health services, and medial supplies. We use the CPI-U to convert all waves to 2010 dollars.

⁴This is the recommended procedure for use of the RAND CAMS data file and is also consistent with the time structure of our simulation model.

⁵Note that we posit each of the morbidity states to contemporaneously influence depression both directly and through changes in self-rated health. For example, a stroke may lower an individual’s self-rated health status which in turn may worsen depression. However, a stroke may also influence depression beyond any changes in self-rated health.

estimate of its total welfare burden. The simulations *will* allow current depression status to influence the evolution of health moving forward through general lagged effects. For example, mild depression today may result in a higher chance of stroke or lower self-rated health the following period. Moreover, higher order lagged effects allow, for example, the recent onset of depression to alter next period self-rated health more than if an individual has been living with depression for an extended period of time.

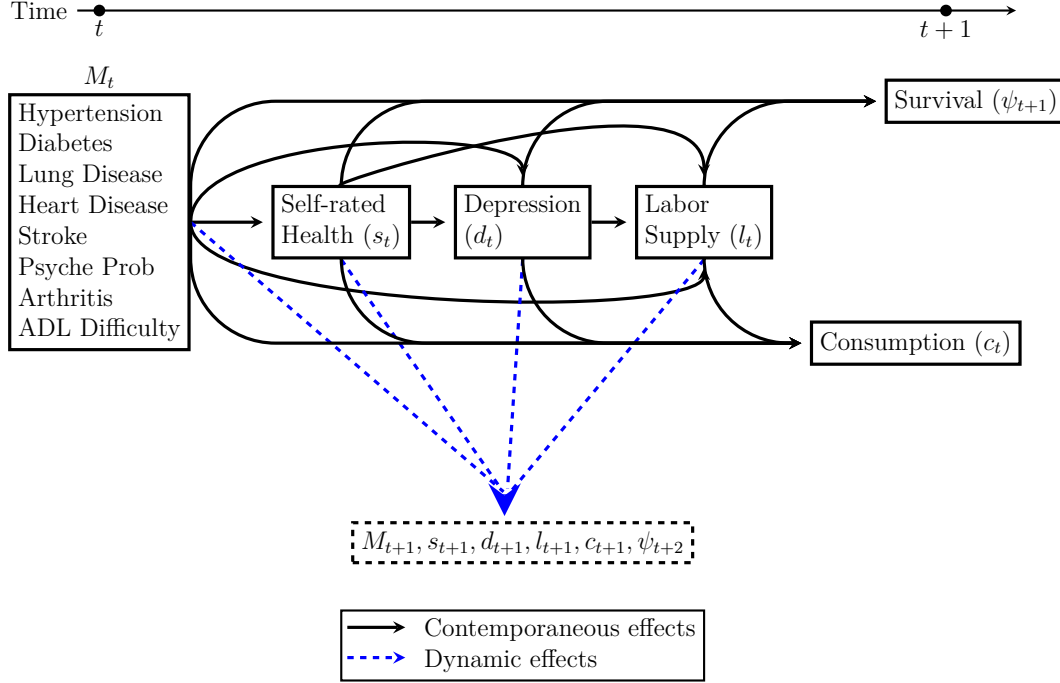


Figure 1: Forecasting model with one period lag

The latter part of the model allows morbidities, self-rated health, and depression to influence labor supply, consumption, and survival to following period of life. The assumed pathway from health to labor supply to consumption is consistent with evidence that health and depression affects the retirement decision (Currie and Madrian, 1999; Doshi et al., 2008; Rice et al., 2011), that consumption declines with retirement (Hall, 2009), and that health impacts economic outcomes, particularly at older ages (Smith, 1999).⁶

2.2.1 Panel VAR representation

While multiple lags are used in estimation of the forecasting model, the following VAR(1) demonstrates the key features of the framework. Let Y_{it} be a vector of outcomes for individual i at time t that includes depression d , log consumption c , retirement

⁶In contrast, the effects of economic status on health appear concentrated during childhood and young adulthood when health trajectories are being established (Smith, 1999).

indicator r , self-rated health s , and our $n = 9$ morbidity states given by $n \times 1$ vector M . Conditional on survival, the outcomes evolve according to the structural VAR(1) model:

$$AY_{it} = BY_{it-1} + \epsilon_{it},$$

where ϵ is a vector of normally distributed shocks with mean zero. The shocks are assumed to be independent and identically distributed (*iid*) across individuals and time and independent across outcomes. The main diagonal terms of matrix A are scaled to one and we assume in our benchmark model that all parameters are homogeneous across individuals and time (e.g. $A_{it} = A \quad \forall i, t$).

We estimate our model in four “blocks” of outcomes—the morbidity block consisting of n outcomes, the self-rated health block (one outcome), the depression block (one outcome), the retirement block (one outcome), and the consumption block (one outcome). The unrestricted model can be written in block matrix form as:

$$\begin{array}{c} n \\ 4 \end{array} \begin{bmatrix} \overbrace{-A_{11} \dots -A_{15}}^n & \overbrace{-A_{12} \dots -A_{15}}^4 \\ \hline -A_{21} & 1 & -a_{23} & -a_{24} & -a_{25} \\ -A_{31} & -a_{32} & 1 & -a_{34} & -a_{35} \\ -A_{41} & -a_{42} & -a_{43} & 1 & -a_{45} \\ -A_{51} & -a_{52} & -a_{53} & -a_{54} & 1 \end{bmatrix} \begin{bmatrix} M_{it} \\ s_{it} \\ d_{it} \\ r_{it} \\ c_{it} \end{bmatrix} = \begin{bmatrix} \overbrace{B_{11} \dots B_{15}}^n & \overbrace{B_{12} \dots B_{15}}^4 \\ \hline B_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ B_{31} & b_{32} & b_{33} & b_{34} & b_{35} \\ B_{41} & b_{42} & b_{43} & b_{44} & b_{45} \\ B_{51} & b_{52} & b_{53} & b_{54} & b_{55} \end{bmatrix} \begin{bmatrix} M_{it-1} \\ s_{it-1} \\ d_{it-1} \\ r_{it-1} \\ c_{it-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,it} \\ \epsilon_{2,it} \\ \epsilon_{3,it} \\ \epsilon_{4,it} \\ \epsilon_{5,it} \end{bmatrix},$$

where $n \times n$ matrix A_{11} has main diagonal terms scaled to one.

As illustrated in Figure 1, the causal pathways we propose suggest a block recursive system. Specifically, we assume that $A_{12} = A_{13} = A_{14} = A_{15} = 0$ in the morbidity block, $a_{23} = a_{24} = a_{25} = 0$ in the self-rated health block, $a_{34} = a_{35} = 0$ in the depression block, and $a_{45} = 0$ in the retirement block. In other words, we assume the contemporaneous causal pathway runs from morbidities to self-rated health to depression to retirement to consumption. However, we allow health and retirement to affect future outcomes through lagged effects.⁷ Block triangulation of the system eliminates simultaneity across blocks and allows for block-by-block estimation.⁸

2.2.2 Exogenous characteristics

We also include a $k \times 1$ vector of exogenous individual characteristics X_{it} as predictors in our model. The VAR(1) model with exogenous regressors takes the following form:

$$AY_{it} = BY_{it-1} + CX_{it} + \epsilon_{it}. \quad (1)$$

⁷Though we assume there is no such feedback from consumption and set $B_{15} = b_{25} = b_{35} = b_{45} = 0$.

⁸Note this produces the same results as the Cholesky decomposition of shocks from a reduced form VAR.

Exogenous characteristics include dummies for age, education, gender, race, census division, census occupation code, and birth cohort. We also include a linear calendar year trend and post-2008 indicator to help account for the influence of the great recession on outcomes.⁹ Lastly, we include a time invariant individual unobserved endowment π in the consumption equation. The endowment π is modeled as a fixed effect with no restriction on the correlation with other model regressors. The unobserved individual effect helps maintain the appropriate variance in consumption across time by effectively acting as a person specific drift in the autoregressive process. We exclude time invariant exogenous regressors (education, gender, race, census division, occupation code, birth cohort) from the consumption equation due to colinearity with the fixed effect. However, we include socioeconomic characteristics instead of additional individual fixed effects in the health and retirement equations because 1) morbidities and retirement are absorbing states and depression and self-rated health are ordinal, each of which poses difficulties in estimating dynamic panel models with fixed effects¹⁰ and 2) the simpler model does well in replicating the dynamics of health and retirement (see online appendix). The resulting exogenous effects then take the following form:

$$CX_{it} = \begin{matrix} n \\ \left\{ \begin{array}{cccccccccc} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} & C_{17} & C_{18} & C_{19} & 0 \\ \hline c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} & c_{27} & c_{28} & c_{29} & 0 \\ c_{31} & c_{32} & c_{33} & c_{34} & c_{35} & c_{36} & c_{37} & c_{38} & c_{39} & 0 \\ c_{41} & c_{42} & c_{43} & c_{44} & c_{45} & c_{46} & c_{47} & c_{48} & c_{49} & 0 \\ c_{51} & 0 & 0 & 0 & 0 & 0 & 0 & c_{58} & c_{59} & c_{510} \end{array} \right. \end{matrix} \underbrace{\begin{bmatrix} Age_{it} \\ Education_i \\ Gender_i \\ Race_i \\ Division_i \\ Occupation_i \\ Cohort_i \\ Year_t \\ Post_t \\ \pi_i \end{bmatrix}}_{k \times 1}.$$

$(n+4) \times k$

Lastly, we normalize $c_{510} = 1$ to allow identification of the unobserved fixed effects in the consumption block.

2.2.3 Morbidities

Block triangulation of the system does not allow direct identification of the structural parameters in the morbidity block as there are $n = 9$ separate outcomes. Instead

⁹Note the inclusion of age, cohort, and calendar year introduces some multicollinearity into the model, so interpreting point estimates on these variables should be done with caution. However, using the estimates for forecasting does not pose an issue (Holford, 1991).

¹⁰For example, as many individuals never enter a given absorbing state in the data, it is not possible to estimate unique unobserved fixed effects for each individual for each absorbing state. As these estimates would be required for our simulations, we do not include additional unobserved fixed effects in absorbing state models.

the morbidity block is estimated as a reduced form VAR. The reduced form system is obtained by premultiplying the structural system block by the inverse of matrix $-A_{11}$:

$$M_{it} = -A_{11}^{-1}B_{11}M_{it-1} - A_{11}^{-1}B_{12}s_{it-1} - A_{11}^{-1}B_{13}d_{it-1} \\ - A_{11}^{-1}B_{14}r_{it-1} - A_{11}^{-1}[C_{11}, \dots, C_{19}]X_{it} - A_{11}^{-1}\epsilon_{1,it}.$$

Denoting $-A_{11}^{-1}B_{1j} = \hat{B}_j$, $-A_{11}^{-1}[C_{11}, \dots, C_{19}] = \hat{C}$ and $-A_{11}^{-1}\epsilon_{1,t} = e_t$ yields the following reduced form system:

$$M_{it} = \hat{B}_1M_{it-1} + \hat{B}_2s_{it-1} + \hat{B}_3d_{it-1} + \hat{B}_4r_{it-1} + \hat{C}X_{it} + e_{it}.$$

In the reduced form VAR all right hand side variables are predetermined at time t and morbidity states do not have direct contemporaneous effect on each other. However, the error terms e_t are composites of morbidity specific structural shocks and thus are potentially correlated across morbidity states (i.e. $cov(e_{it}, e'_{it}) \neq 0$). This allows for contemporaneous correlation in the probability of morbidity states. For example, the onset of heart disease may be correlated with the onset of hypertension or stroke due to correlated contemporaneous shocks. Contemporaneous morbidity shocks are assumed to follow a standard multivariate normal distribution with an $n \times n$ covariance matrix given by Σ .

As morbidity outcomes are binary, forecasting of the measures is not a true linear VAR process. In contrast, we assume a continuous latent variable m^* underlies each observed outcome such that:

$$m_{j,it} = 0 \quad \text{if} \quad m_{j,it}^* \leq 0 \\ m_{j,it} = 1 \quad \text{if} \quad m_{j,it}^* > 0$$

for $j = 1 \dots n$. We then have the following model:

$$\begin{bmatrix} m_{1,it}^* \\ \vdots \\ m_{n,it}^* \end{bmatrix} = \begin{bmatrix} \hat{b}_{11} & \cdots & \hat{b}_{1n} \\ \vdots & \ddots & \vdots \\ \hat{b}_{n1} & \cdots & \hat{b}_{nn} \end{bmatrix} \begin{bmatrix} m_{1,it-1} \\ \vdots \\ m_{n,it-1} \end{bmatrix} + \hat{B}_2s_{it-1} + \hat{B}_3d_{it-1} + \hat{B}_4r_{it-1} + \hat{C}X_{it} + \begin{bmatrix} e_{1,it} \\ \vdots \\ e_{n,it} \end{bmatrix}. \quad (2)$$

Note that each latent morbidity variable is determined by lagged values of the other *observed* self-rated health, depression, and morbidity states. Given the assumed joint normality of the error structure, this morbidity block of equations is in the form of a multivariate probit model.

2.2.4 Self-rated health

As self-rated health is measured on a five point scale, forecasting the measure is again not a linear VAR process. Instead, we assume a continuous latent variable s^* underlies the observed self-rated health state such that the relevant equation given in system (1) can be explicitly written as:

$$s_{it}^* = A_{21}M_{it} + B_{21}M_{it-1} + b_{22}s_{it-1} + b_{23}d_{it-1} + b_{24}r_{it-1} + [c_{21}, \dots, c_{29}]X_{it} + \epsilon_{2,it}, \quad (3)$$

with the observed self-rated health state defined as:

$$s_{it} = \delta \text{ if } \kappa_{\delta-1} < s_{it}^* < \kappa_{\delta} \text{ for } \delta = 1, \dots, 5$$

for cut-points $(\kappa_0, \dots, \kappa_5)$ with $\delta = 1$ representing the worst health state (poor) and $\delta = 5$ the best health state (excellent). Note that latent self-rated health is assumed to depend on the lagged value of the *observed* self-rated health category to incorporate the persistence in general health shocks over the life-course. We assume ϵ_2 is an *iid* shock with standard normal distribution. Thus, the evolution of self-rated health follows an ordered probit structure. Unlike the morbidity block, this equation may be estimated independently of other outcome blocks with all structural parameters identified.

2.2.5 Depression

Similar to other health outcomes, we assume a continuous latent variable d^* underlies the observed depression state such that the forecasting equation given in system (1) can be written:

$$d_{it}^* = A_{31}M_{it} + B_{31}M_{it-1} + a_{32}s_{it} + b_{32}s_{it-1} + b_{33}d_{it-1} + b_{34}r_{it-1} + [c_{31}, \dots, c_{39}]X_{it} + \epsilon_{3,it}, \quad (4)$$

with the observed depression state defined as:

$$d_{it} = \delta \text{ if } \kappa_{\delta} < d_{it}^* < \kappa_{\delta+1} \text{ for } \delta = 0, \dots, 8$$

for cut-points $(\kappa_0, \dots, \kappa_9)$ with $\delta = 0$ representing the no depressive symptoms state and $\delta = 8$ the worst depression state. Again, latent depression is assumed to depend on the lagged value of the observed depression category to incorporate the persistence in depression over time. We assume ϵ_3 is an *iid* shock with standard normal distribution yielding an ordered probit structure for the depression model. Given our block recursive system, this equation may also be estimated independently of other blocks with all structural parameters identified.

2.2.6 Retirement

As retirement is a binary outcome, we assume a continuous latent variable r^* underlies the observed outcome such that:

$$r_{it} = 0 \text{ if } r_{it}^* \leq 0 \\ r_{it} = 1 \text{ if } r_{it}^* > 0.$$

Conditional on working the previous period (and setting $b_{44} = 0$), the retirement model as defined in system (1) is given by:

$$r_{it}^* = A_{41}M_{it} + a_{42}s_{it} + a_{43}d_{it} + B_{41}M_{it-1} + b_{42}s_{it-1} + b_{43}d_{it-1} + [c_{41}, \dots, c_{49}]X_{it} + \epsilon_{4,it}. \quad (5)$$

In addition to exogenous individual characteristics, retirement is influenced by current and lagged values of health (depression, self-rated health, and specific morbidities). We assume ϵ_4 is an *iid* shock with standard normal distribution implying the retirement model has a standard probit structure.

2.2.7 Consumption

The consumption forecasting equation given in system (1) can be explicitly written as:

$$c_{it} = A_{51}M_{it} + a_{52}s_{it} + a_{53}d_{it} + a_{54}r_{it} + B_{51}M_{it-1} + s_{52}d_{it-1} + b_{53}d_{it-1} + b_{54}r_{it-1} + b_{55}c_{it-1} + c_{51}Age_{it} + c_{58}Year_t + c_{59}Post_t + \pi_i + \epsilon_{5,it}. \quad (6)$$

This is a standard linear dynamic panel data model with lagged dependent variable and individual level fixed effects (π). Given our block recursive system, this equation may also be estimated independently of other blocks with all structural parameters identified including the variance of ϵ_5 .

2.2.8 Mortality

The final process to be modeled is survival from one period of life to the next. As all other outcomes are conditional on survival, mortality probabilities are estimated independently of the VAR system above. Conditional on being alive at time $t - 1$, survival to the following period of life is modeled by:

$$\psi_{it} = I \left(\sum_{k=1}^K \left[\gamma_k^M M_{it-k} + \gamma_k^s s_{it-k} + \gamma_k^d d_{it-k} + \gamma_k^r r_{it-k} \right] + \delta X_{it} + u_{it} > 0 \right), \quad (7)$$

where $I(\cdot)$ is an indicator function and $\psi = 1$ indicates survival, X the vector of observed individual characteristics previously defined, and u_{it} an *iid* random shock with standard normal distribution. The specification allows K lags of morbidity states, self-rated health, depression, and retirement to influence survival probability.

2.3 Welfare measure

We use an ex-ante consumption-compensating variation (CV) measure to quantify the welfare costs of late-life depression using simulations from our VAR model. We first define expected lifetime utility at age j for individual i as:

$$E \left[\sum_{a=j}^J \psi_{ia} \beta^{a-j} \phi(h_{ia}) [\bar{u} + \log(c_{ia}) + \nu(l_{ia})] \right]$$

where c is consumption, l leisure, h health, and ψ is a survival indicator. Expectations are taken with respect to the uncertainty in the evolution of consumption, leisure,

health, and mortality after age j . This simple formulation yields an additive decomposition of welfare allowing us to add cumulative corrections for the cost of depression on comorbidities, mortality, leisure, and consumption (see online appendix for derivation). We also check the robustness of our results to more general preferences. The health function $\phi(h) \in [0, 1]$ scales the utility from consumption and leisure such that $\phi(h) = 1$ represents utility in the perfect health state and $\phi(h) = 0$ represents the dead state. At the same time, $\psi\phi(h)$ represents a measure of quality-adjusted life years (QALYs). For instance, $\psi\phi(h) = 1$ represents a year of life with no adverse health conditions; a single QALY.

Let $U_{ij}(1 - \lambda)$ denote the expected lifetime utility at age j from the outcome bundles of individual i if consumption is multiplied by a factor $(1 - \lambda)$ at each age and realization of the world:

$$U_{ij}(1 - \lambda) = E \left[\sum_{a=j}^J \psi_{ia} \beta^{a-j} \phi(h_{ia}) [\bar{u} + \log((1 - \lambda) c_{ia}) + \nu(l_{ia})] \right].$$

The consumption-compensating variation measure of welfare for individual i , λ_{ij} , is derived through the condition:

$$U_{mj}(1 - \lambda_{ij}) = U_{ij}(1), \quad (8)$$

where U_{mj} refers to the expected lifetime utility from the outcome bundles in the absence of any possible depression after age j . In words, λ_{ij} is the proportion of the individual's (depression-free) consumption they would be willing to give up at every age starting from j (in all possible realizations of the world and holding health and leisure fixed) to eliminate all possibility of depression after age j . For example, if person i expects depression to be a serious problem in late life, they may have a welfare measure $\lambda_{ij} = 0.3$. This implies they would be ex-ante willing to give up to 30% of their consumption in every period from age j to avoid any possibility of depression. As this measure is based on potential outcomes over remaining life, it encompasses the likelihood of persistence and emergence of depression over remaining life.

In order to gain a sense of the aggregate cost of depression, we also calculate the product of an individual's expected remaining lifetime consumption at age sixty (ELC) and our EV measure: $\lambda \times ELC$. This is a similar concept (but not the same) as an individual's willingness-to-pay at age sixty to eliminate all possibility of depression after age sixty. Effectively, it is an individual's *expected* willingness-to-pay or the expected value of consumption they are willing to forgo.

2.3.1 Health utility weights

Analysis using our welfare model requires calibration of preference parameters. This includes parameters of the function $\phi(h)$ mapping health states into flow utility. We assume health utility depends linearly on our health state vector: $\phi(h_t) = \gamma h_t$. We use the the Health Utilities Index Mark 3 (HUI3) instrument to obtain our vector of

utility weights γ . The HUI3 was included as a module for a subset of approximately 1,200 of the participants in the 2000 wave of the HRS. We use the HUI multi-attribute utility score (*hui3ou*) which is a health related quality of life measure where death is defined by a score of zero and perfect health by a score of one (Furlong et al., 1998; Feeny et al., 2002; Horsman et al., 2003).

Conceptually, the HUI3 was developed such that $u(h_i) = HUI3_i \times u(h_{best})$ for individual i and general utility function $u(\cdot)$. For example, two years of life for an individual with $HUI3 = 0.5$ is equal (in utility) to one year in the best health state. In the context of our model, we assume that the HUI3 measures the relative utility across health states *holding consumption and leisure fixed*:

$$\gamma h_i [\bar{u} + \log(c_i) + \nu(l_i)] = HUI3 \times h_{best} [\bar{u} + \log(c_i) + \nu(l_i)].$$

This approach is consistent with the HUI3 instrument where the interview script reads: “*when imagining yourself in these health states please remember that where you live, your income, your friends, and family would be the same as now.*” With this assumption, the above equation simplifies to $\gamma h_i = HUI_i$ when $h_{best} = 1$. The utility weights γ can then be estimated by regressing the HUI3 utility score on depression score, self-rated health, and all morbidity indicators.

2.3.2 Calibration of other parameters

Leisure is normalized to one for retired individuals. We assume workers supply 2,000 annual hours to the labor market and set associated leisure to $0.66 = 1 - (2000/5,840)$ where 5,840 is the assumed annual time endowment (16 hours a day \times 365 days). Preferences over leisure are defined by $\nu(l) = -\frac{\theta\epsilon}{1+\epsilon} (1-l)^{\frac{1+\epsilon}{\epsilon}}$, where ϵ is a constant Frisch elasticity of labor supply (the elasticity of labor supply with respect to wage, holding the marginal utility of consumption fixed). We follow Miller and Bairoliya (2021) and set the disutility weight θ such that the marginal cost of leisure equals the marginal benefit for the median individual in our sample. This gives us a benchmark $\theta = 7.82$. We also use a benchmark value of $\epsilon = 1$ while testing the sensitivity of our results to alternate values. Likewise, we choose a discount factor $\beta = 0.98$ corresponding to an annual discounting of one percent. Note that there is additional implicit discounting in the expected utility framework due to mortality risk.

Finally, note that with our benchmark preferences, as long as flow intercept \bar{u} plus log consumption is positive, a retired individual will prefer life to death. After normalizing consumption to thousands of 2010 dollars, we set $\bar{u} = -\log(3)$, implying that \$3,000 of consumption is needed for a retiree to maintain positive flow utility. This is approximately 10% of mean annual consumption in our sample. Although there is not much evidence on this value, Murphy and Topel (2006) argue 10% as a reasonable parameterization. This value also yields a median value of remaining life for sixty year olds in our simulation sample of \$47,000 per QALY.¹¹ In a review of the literature,

¹¹The value of life per QALY at age j is given by $VOL_j/E \left[\sum_{a=j}^J \psi_a \beta^{a-j} \phi(h_a) \right]$ where $VOL_j =$

Ryen and Svensson (2015) estimate mean and median values of life across studies of approximately \$98,000 and \$32,000.¹² Traditional values in the U.S. often range from \$50,000 to \$100,000 (Kaplan and Bush, 1982). In some robustness exercises, we show that using log consumption and a relatively low value of life in our benchmark likely yields conservative estimates of welfare costs.

2.4 Estimation and simulations

Equipped with our simulation model and welfare concept, our empirical analysis involves four steps.

1. We use data from the HRS to estimate the parameters of the simulation model. Here we use data on all individuals aged fifty and older from all available waves of the HRS from 1992-2016 (40,708 unique individuals and 238,091 total individual-year observations). See the online appendix for details on the model estimation sample and procedures.
2. We repeatedly simulate remaining life-cycle paths for all outcomes for a subsample of the HRS respondents using the parameter estimates and age sixty data as initial conditions. This simulation sample includes all individuals with age sixty data and requisite lagged data for simulations. This yields representative results over four birth cohorts—early HRS (EHRS), late HRS (LHRS), War Babies, and (early) Baby Boomers. See the online appendix for details on initial condition descriptives, sampling weights and representativeness, and simulation procedure.
3. We estimate a lower bound of the welfare costs of depression after age sixty for each individual in our simulation sample by assuming there are no spillover effects on other model outcomes. More specifically, we leave all expected paths of morbidities, self-rated health, mortality, labor supply, and consumption at their baseline levels and only remove the health utility penalty associated with depression at age sixty-two and older (i.e., we set all the CESD weights to zero in Table 1).
4. We estimate an upper bound of the welfare costs of depression by running a new set of simulations starting from the same initial conditions but removing any possibility of depressive symptoms after age sixty. We consider this an upper bound as it assumes all the coefficients estimated in the simulation model are

$U_{ij}(1) c_j / \phi(h_j)$.

¹²Ryen and Svensson (2015) document substantial variation across estimates of willingness-to-pay for a QALY, most notably with conversions based on revealed preferences of the value of statistical life (VSL) averaging 5-7 times higher than those based directly on stated preferences. The VSL studies reviewed are by definition measuring value of length of life, while stated preference studies elicited willingness-to-pay for pure quality of life improvements, pure length of life, or a mixture of both.

purely causal. We embed the baseline and counterfactual simulated data within our expected utility framework to construct a measure of the ex-ante welfare cost of future depression at age sixty for each individual in our simulation sample.

3 Results

3.1 Model estimates

We begin with estimation results from our simulation model to demonstrate the association between depression and other outcomes in the data. Select results are provided in Figure 2 while the full set of results are available in the online appendix. The first panel shows the average marginal effects of depressive symptoms on the contemporaneous probability of retirement (controlling for other health outcomes as shown in model (1)). Results indicate that low and mild depression (CESD=0,1,2,3,4) do not have a significant association with the probability of retirement for older adults. However, as the severity of depression increases, a significant association emerges. For example, there is an increase in the probability of retirement of around 2 percentage points (pp) for a CESD score of five, rising to a point estimate of over 5 pp at higher levels of depression. The second panel of Figure 2 shows a small *positive* relationship between the severity of depression and contemporaneous log consumption. For example, a CESD score of five is associated with an increase in consumption of about 0.03 log points. Expectations about the longevity and quality of life may be a plausible explanation for this positive association among older adults. For example, lower expectations about a long and healthy life might push older depressed adults to discount their future utility and hence consume more in the near term. Moreover, there is an additional indirect effect operating on consumption due to any changes in retirement induced by depression.

Panel three of Figure 2 shows a generally positive association between depressive symptoms and mortality. For example, an individual with a CESD score of three has about a 1 pp lower probability of surviving to the next model period compared to if they had a CESD score of zero. Note that at the highest CESD scores the association with mortality is moderately diminished, although results are somewhat noisy given that relatively few individuals have very high levels of depression. The final two panels help illustrate the dynamics of the system by showing the average marginal effects of current depression state on the probability of having a stroke or ADL difficulty the *following* model period. For example, a CESD score of five increases the probability of a stroke the following period by nearly 1 pp and the probability of having difficulty with ADLs by more than 2 pp. Moreover, these relationships continue to propagate dynamically throughout the system influencing the evolution other comorbidities and self-rated health along with future retirement and consumption decisions.

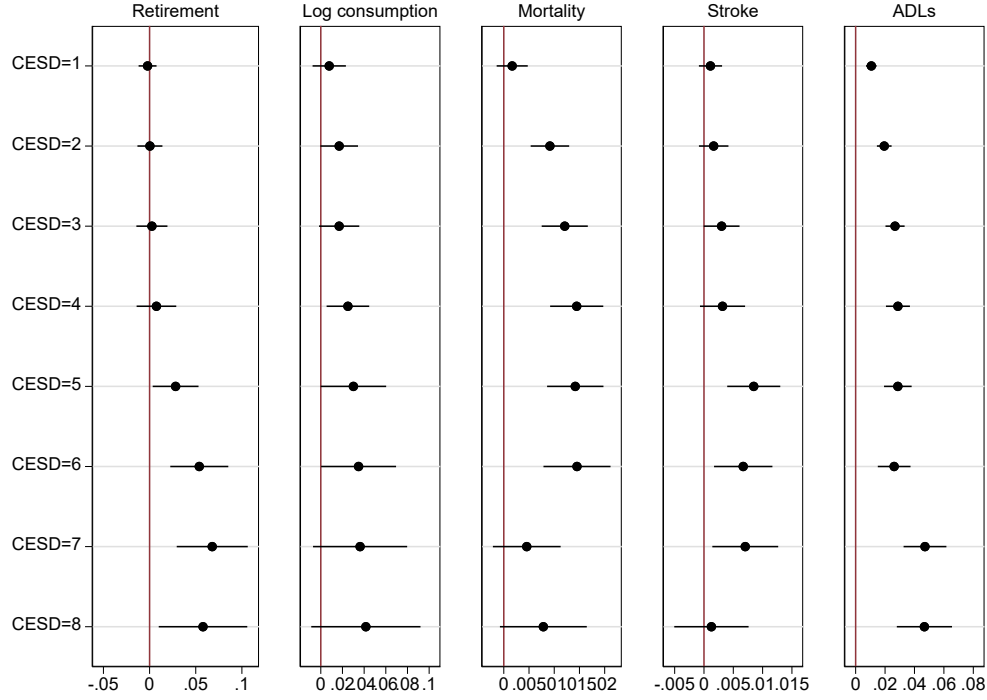


Figure 2: Select estimation results

Notes: Dependent variables across columns. Average marginal effects on the probability of an outcome reported for probit results—retirement, mortality, stroke, and ADLs. Contemporaneous associations reported for retirement and consumption as dependent variables. Lagged associations reported for mortality, stroke, and ADLs. CESD=0 (no depression) is the reference group. Spikes indicate 95% confidence intervals.

3.1.1 Simulation fit

A comparison between mean simulated CESD scores and those based on available data is shown by age and cohort in Figure 3. Additional comparisons for each outcome by cohort are provided in the online appendix. In both the data and simulations, mean CESD score tends to rise with age in the EHRS cohort. For the LHRS cohort the relationship is U shaped, CESD scores decline until the age of 72, then begin to increase. For the younger War Baby and Baby Boomers cohorts there is less available data, but both cohorts have falling CESD scores over the sixties. Note that by construction, the data and simulations are the same at age sixty. However, using only age sixty data and the estimated model parameters, the simulations continue to match the data reasonably well even up to 24 years later (when the EHRS cohort is age 84). Overall, the simulations match the available aggregated data well suggesting our life-cycle dynamics model provides a good approximation of the underlying data generating processes.

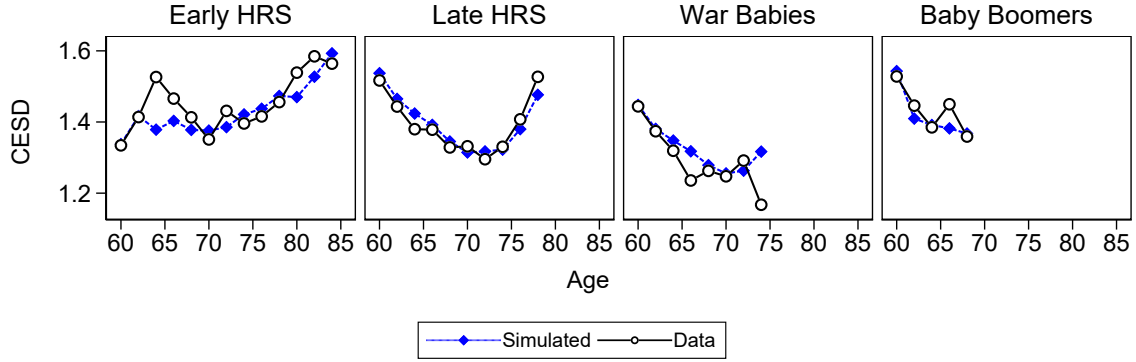


Figure 3: Mean of life-cycle CESD profiles by cohort

Notes: “Data” plots mean of all available data (inclusive of imputed missing values) in HRS by two-year age interval and cohort. “Simulated” plots mean of expected simulated outcome for each observation in the data (i.e. the expected outcome for each person-year observation in the data).

3.1.2 Health utility weights

Table 1 provides our health utility weights γ estimated via a linear regression of HUI3 utility score on health outcomes. Depressive symptoms measured by the CESD scale have a strong and highly significant negative association with utility. For example, moving from no depressive symptoms (the base category) to a CESD score of three lowers flow health utility by 7.9 pp. Moving all the way to a score of eight lowers health utility by 29.3 pp. In addition to depression, self-rated health also has a strong association with health utility. For example, moving from poor health (the base category) to good health improves flow health utility by 25.0 pp. Conditions such as hypertension, diabetes, and cancer have little independent effect on health utility after controlling for their association with self-rated health, depression, and other comorbidities. Other morbidities such as stroke and arthritis have larger (and statistically significant) independent negative effects. The most influential morbidity indicator is difficulty with ADLs, which lowers health utility by an estimated 14.3 pp.

While the eight-point CESD does not map directly into clinical diagnosis of depression disorder, Turvey et al. (1999) propose a CESD score of six or higher to approximate cases of clinical depression. Our weights then imply that a clinically depressed individual in good self-rated health and without other comorbidities would have a health utility score between 0.56-0.68. A clinically depressed individual with poor self-rated health would have a score of 0.31-0.43. In a systematic review, Mohiuddin and Payne (2014) examine results from studies using indirect valuation methods to estimate health utility scores in alternate depressive states. They calculate pooled mean utilities across studies of 0.56 for mild, 0.45 for moderate, and 0.25 for severe depression. By comparison, our results are likely conservative in attributing health utility penalties to depressive states.

Table 1: Estimated health utility weights (γ)

Measure	Weight	SE
Depression		
CESD=1	-0.021	0.015
CESD=2	-0.087	0.018
CESD=3	-0.079	0.023
CESD=4	-0.094	0.028
CESD=5	-0.137	0.030
CESD=6	-0.172	0.039
CESD=7	-0.224	0.046
CESD=8	-0.293	0.056
Hypertension	0.005	0.012
Diabetes	-0.000	0.017
Cancer	0.007	0.017
Lung disease	-0.024	0.021
Heart disease	-0.034	0.015
Stroke	-0.073	0.022
Psych problem	-0.041	0.020
Arthritis	-0.053	0.012
Diff with ADL	-0.143	0.016
Self-rated health		
Fair	0.179	0.026
Good	0.250	0.027
Very good	0.331	0.028
Excellent	0.338	0.032
Constant	0.610	0.030

Notes: Results from regression of HUI3 score on CESD score, self-rated health, and morbidities. SE denotes standard error. $R^2 = 0.48$. $N = 1,089$.

3.2 Cost of depression

We start with a detailed examination of the cost of depression in the EHRS cohort as it is the oldest of the four cohorts and contains the longest panel of available data. Table 2 shows the mean cost of depression after age sixty for the EHRS. The first column provides our lower bound estimate where we simply remove the health utility penalty of depressive symptoms but leave all simulated outcomes at their baseline levels. On average, removing the health utility penalty of depression increases quality-adjusted life expectancy by 0.85 years. The mean associated CV welfare measure is 0.084, implying a willingness-to-pay up to 8.4% of annual consumption over remaining life to avoid any possibility of depression. As shown in the final row, this amounts to an expected loss of \$46,222 of lifetime consumption.

The final four columns of Table 2 provide results from re-simulating outcomes for each individual after removing any possibility of depression after age sixty. Each column cumulatively adjusts welfare for an additional outcome with our “fully adjusted” upper bound provided in the last column. For example, when accounting for the possi-

Table 2: Mean costs of depression after age sixty

	Depression	Cumulative corrections			
		Comorbidities	Mortality	Leisure	Consumption
Expected loss					
QALYs	0.851	1.241	2.074	2.074	2.074
Labor supply (yrs)				0.099	0.099
Consumption (annual)					-0.616
CV (λ)	0.084	0.109	0.159	0.158	0.150
$\lambda \times \text{ELC}$	46.222	63.626	95.459	94.755	91.032

Notes: Estimates use base year respondent analysis weights. ELC denotes expected lifetime consumption. Consumption in \$1000s.

ble spillover effects of depression on comorbidities (column two), eliminating depression increases quality-adjusted life expectancy by 1.24 years. The mean associated CV implies a willingness-to-pay up to 10.9% of annual consumption to avoid depression when accounting for these spillovers, or an expected loss of \$63,626 of lifetime consumption. Further adjusting for the effect of depression on mortality rates yields an increase in quality-adjusted life expectancy of 2.07 years, a willingness-to-pay of 15.9% of annual consumption, or an expected loss of \$95,459 in lifetime consumption.

Moving to leisure, our simulations suggest that eliminating depression after age sixty could increase labor supply only by an average of about 1.1 months (0.099 years). This relatively small impact is likely due to the fact that many individuals in the simulation sample are already retired by age sixty and the direct effects of depression on retirement are minimal except with quite severe symptoms (recall Figure 2). As increased labor supply alone results in a loss in welfare due to less leisure time, the mean CV falls very slightly to 15.8% of annual consumption. Finally, simulations suggest eliminating depression could lower consumption by up to \$616 annually. This is consistent with the positive contemporaneous association between depression and annual consumption shown in Figure 2. The mean associated CV implies the willingness-to-pay falls to 15.0% of annual consumption to avoid depression when accounting for these consumption losses. However, as we demonstrate in the next section, a fall in annual consumption does *not* imply a fall in lifetime consumption.

In order to gain a better sense of how depression influences the dynamics of other outcomes in the system, Figure 4 plots the average percentage change in expected outcomes with the exogenous elimination of all prevalence of depression after age sixty. The first two plots show that the elimination of depression after age sixty leads to a significant decline in psychiatric problems, difficulty with ADLs, and to some extent lung disease. For example, the elimination of depression is associated with nearly a 30% decrease in the probability of diagnosed psychiatric problems by the late-seventies. Associated effects were fairly small for the other morbidities.¹³

¹³Cancer has relatively little association with depression or other morbidities. However, as eliminating depression improves chances of survival even when sick, there is actually a small increase in

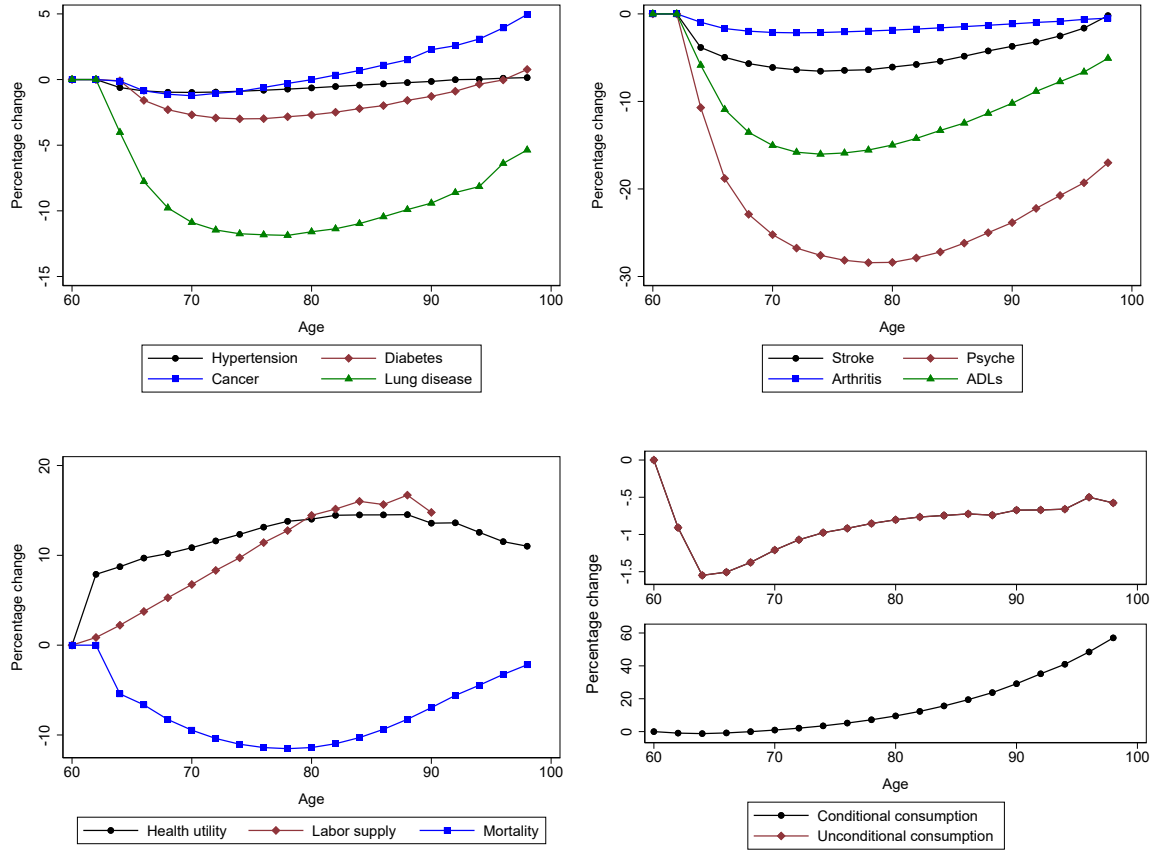


Figure 4: Expected cost of depression by age

Notes: Results plot percentage difference in expected outcomes with the exogenous elimination of all prevalence of depression after age sixty. Sample includes all individuals in the simulation sample. Expected outcomes in first three panels are conditional on survival.

The third plot of Figure 4 shows the upper bound effect of eliminating depression on health utility, labor supply, and mortality. The age-specific mortality rate is estimated to be over 10% lower by the early-seventies and remains more than 5% lower even into the nineties. In contrast, the probability of working is about 15% higher by the late-eighties (though there are few workers left by this age). There is an immediate increase of about 8% in health utility at age sixty-two, which climbs to nearly 15% by age eighty. The final plot shows the response of consumption (unconditional and conditional on survival) to the elimination of depression. There is about a 1-2% decline in annual consumption conditional on survival throughout the remaining life-cycle. In contrast, when examining expected unconditional consumption (i.e. imputing zero consumption for the dead state), there is large rise over time, reaching differences of more than 50% by the early nineties. These plots again highlight the small loss in annual consumption

the prevalence of cancer (conditional on survival) starting around age eighty.

but potential gains in lifetime consumption due to an increase in life expectancy.

3.3 Distribution of cost

Another advantage of our approach is that we have individual level data and simulations so we are able to examine the entire distribution of estimated costs. Figure 5 shows the distribution of estimated QALYs lost due to late-life depression. The “depression only” curve plots our lower bound estimate with a mean of 0.851 as shown in Table 2. The distribution demonstrates substantial inequality in the expected direct utility cost of late-life depression—the worst off individuals expect a loss of nearly three QALYs. When adding the estimated spillover effects on comorbidities, the mean shifts to 1.239 and the distribution flattens. This implies the health utility cost of depression becomes even more unequal when accounting for spillovers on other comorbidities. When further adjusting for increased mortality rates associated with depression, the mean reaches 2.074 and inequality in the distribution continues to rise.

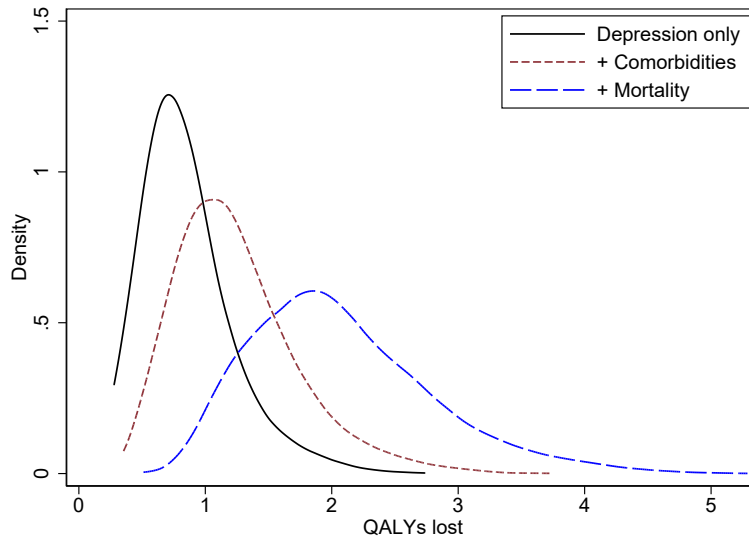


Figure 5: Expected lost quality-adjusted life years (QALYs) after age sixty

Turning to economic outcomes, Figure 6 shows the distribution of the estimated upper bound on the expected loss in consumption and labor supply associated with late-life depression. The first panel shows the change in expected annual consumption (conditional on survival). The negative values again demonstrate the small expected *gain* in annual consumption from late-life depression. For most individuals this gain is less a \$1,000, though a small share expect to gain more than \$3,000 in annual consumption. Despite this rise in annual consumption with depression, the second panel of Figure 6 shows there is an expected fall in total lifetime consumption for all individuals due to decreased life expectancy. The expected loss in lifetime consumption

averages around \$16,000, but ranges from almost zero to over \$100,000. Finally, the third panel shows the distribution of lost labor supply. The potential loss due to early retirement is quite small—less than a year for all individuals.

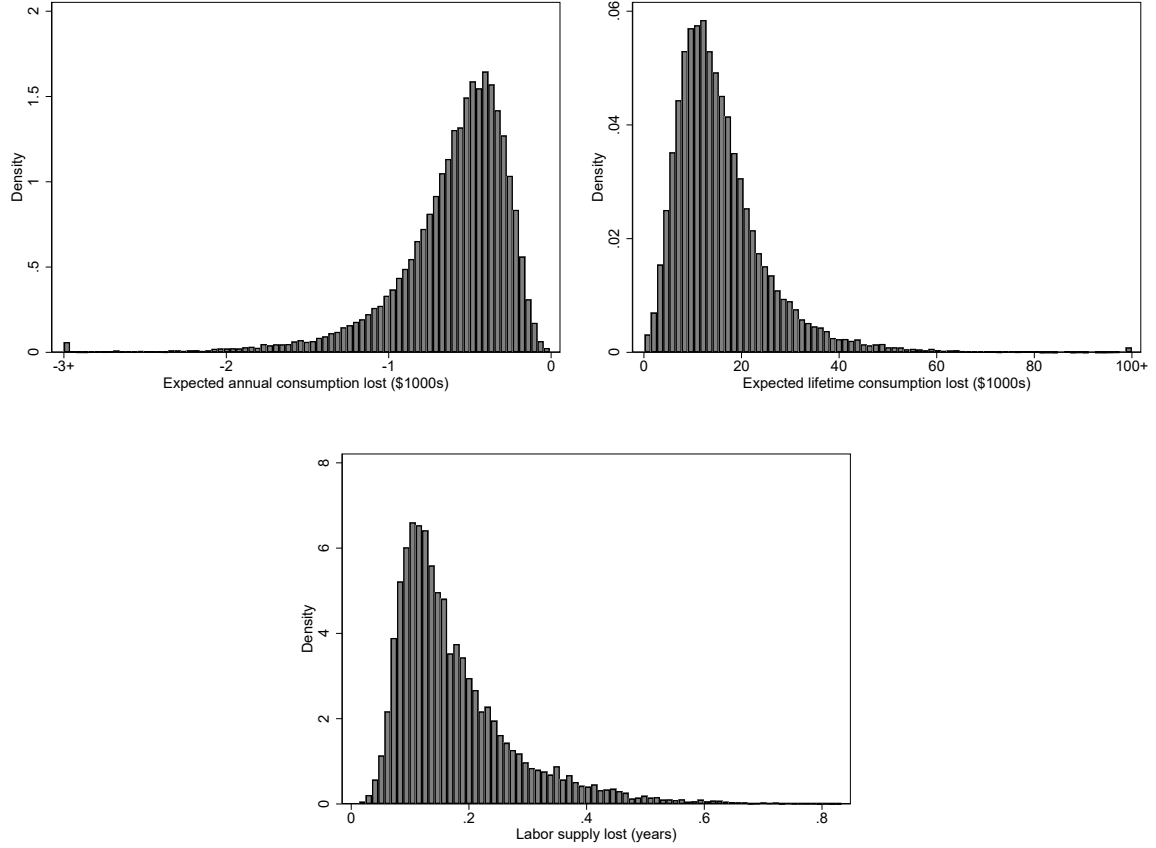


Figure 6: Expected consumption and labor supply loss after age sixty (full model)

Notes: Estimates use base year respondent analysis weights. Labor supply lost conditional on working at age sixty.

Figure 7 shows the distribution of the expected welfare cost of late-life depression (compensating variation) and the approximate expected monetary equivalent. As shown in Table 2, ignoring all spillovers (depression only model) yields a lower bound on the average welfare cost of 8.4% of annual consumption. The first panel of Figure 7 shows that most of the lower bound distribution falls under 10%, though there is a thin right tail suggesting substantially higher costs for a select few. Likewise, the second panel shows this lower bound translates into an expected loss in lifetime consumption of under \$100,000 for most individuals in the sample. When adjusting for potential spillover effects of depression on comorbidities, mortality, leisure, and consumption, there is a substantial increase in the mean and inequality of welfare costs. For example, there is now a substantial portion of the distribution willing to pay over 20% of their annual consumption to avoid the possibility of late-life depression.

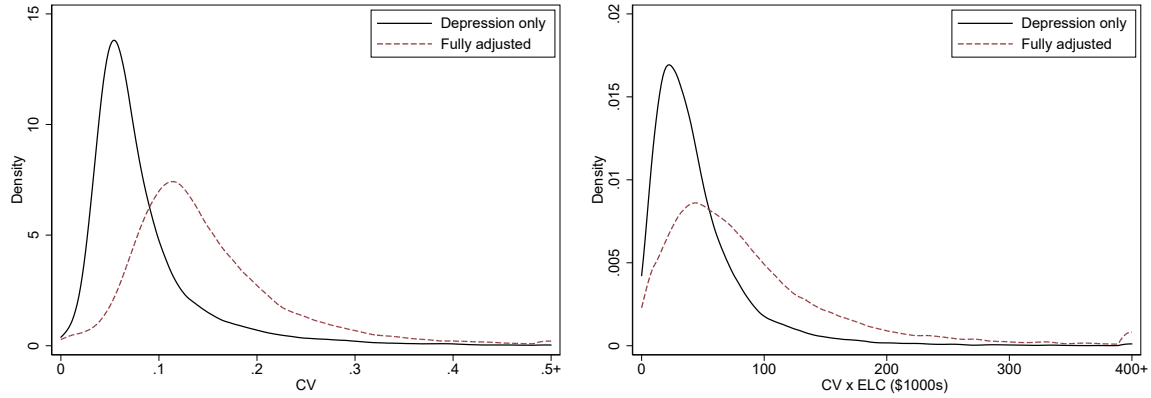


Figure 7: Expected welfare loss after age sixty

Notes: Estimates use base year respondent analysis weights. ELC denotes expected lifetime consumption (\$1000s).

3.4 Cost over birth cohorts

Our analysis so far has focused on results only in the EHRS birth cohort. We now compare our estimated welfare costs across the four cohorts in our simulation sample. We begin by examining the predicted mean CESD depression score by age and cohort (see Figure 8). In contrast to each of the younger cohorts, simulations suggest that the EHRS experienced a rising mean CESD score during their sixties. However, based on currently available data, our model predicts that all cohorts have realized or will realize a rising mean CESD score over much of their seventies and eighties. In general, these trends are consistent with the U-shaped pattern over age found in previous studies (Mirowsky and Ross, 1992; Sutin et al., 2013; Tampubolon and Maharani, 2017; Abrams and Mehta, 2019). After the late-eighties, CESD scores are predicted to fall for all cohorts. In terms of levels, the model predicts that after their mid-sixties, mean depression scores will be lower among War Babies and Baby Boomers than the early or late HRS cohorts.

Table 3 reports the estimated costs of late-life depression by cohort. When only considering the direct health utility penalty of depression (depression only model), the expected loss in QALYs is slightly smaller for the younger cohorts. For example, the average expected QALYs lost falls from 0.85 for the EHRS cohort to 0.80 for Baby Boomers. Likewise, the willingness-to-pay to avoid depression falls from 8.4% to 7.5% of annual consumption. A similar general pattern of falling average costs of depression over cohorts remains when adjusting for spillover effects on other modeled outcomes (full model). For example, the expected QALYs lost falls from 2.07 to 1.93 between the EHRS and Baby Boomer cohorts. Similarly, the expected annual consumption gain from depression falls from \$616 to \$479. There is also a slightly higher gain in labor supply for younger cohorts. In terms of our CV welfare metric, fully-adjusted willingness-to-pay falls from 15.0% to 12.9% of annual consumption. This amounts to

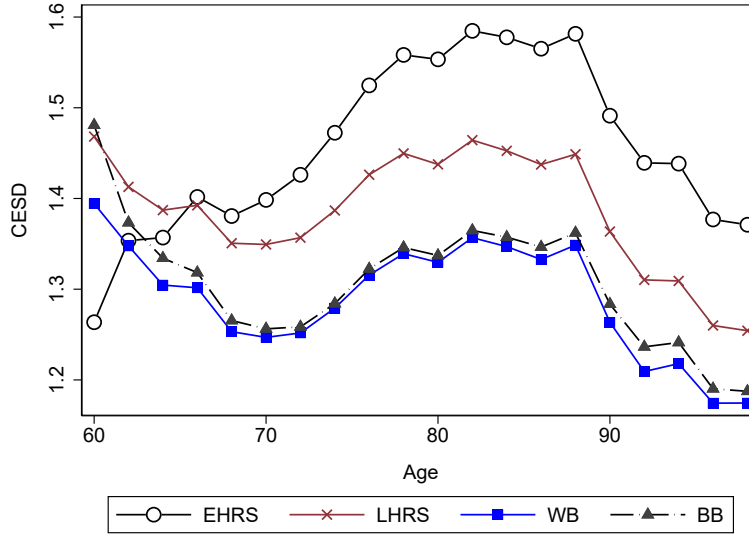


Figure 8: Expected CESD by age and cohort

Notes: Expected CESD score (0-8 scale) conditional on survival.

a fall in the average expected loss of lifetime consumption from \$91,032 in the EHRS cohort to \$77,364 among Baby Boomers.

Table 3: Mean costs of depression after age sixty by birth cohort

	Depression only				Full model			
	EHRS	LHRS	WB	BB	EHRS	LHRS	WB	BB
Expected loss								
QALYs	0.851	0.824	0.789	0.808	2.074	1.985	1.896	1.931
Labor supply (yrs)					0.099	0.101	0.101	0.116
Consumption (annual)					-0.616	-0.597	-0.539	-0.479
CV (λ)	0.084	0.083	0.077	0.075	0.150	0.146	0.136	0.129
$\lambda \times \text{ELC}$	46.222	46.812	44.360	39.756	91.032	91.703	86.544	77.364
CV Gini	0.336	0.360	0.381	0.421	0.278	0.291	0.314	0.347

Notes: Estimates use base year respondent analysis weights. ELC denotes expected lifetime consumption. Consumption in \$1000s.

While we estimate falling *average* costs of depression over cohorts, the final row of Table 3 reveals another important trend. When looking at the distribution within a cohort, we see that the inequality of depression costs is rising. For example, the Gini coefficient on our fully-adjusted CV measure of welfare has increased from 0.27 in the EHRS cohort to 0.34 for Baby Boomers. Thus, while the average cost has decreased slightly across cohorts, the welfare burden of depression has become significantly more unequal *within* cohorts.

3.5 Robustness

Table 4 provides robustness results for our key welfare numbers under several alternate modeling assumption. The online appendix provides additional robustness results for preference parameters $(\beta, \epsilon, \theta)$ and using alternate health utility weights.

Table 4: Select robustness results

	VOL	Depression only				Full model			
		EHRS	LHRS	WB	BB	EHRS	LHRS	WB	BB
Benchmark	47.486	0.084	0.083	0.077	0.075	0.150	0.146	0.136	0.129
$\bar{u} = \ln(1.5)$	67.036	0.119	0.118	0.110	0.110	0.214	0.208	0.195	0.191
$\gamma = 1.5$	73.959	0.140	0.139	0.131	0.122	0.256	0.253	0.237	0.221
$\gamma = 2$	106.412	0.210	0.211	0.197	0.180	0.344	0.342	0.323	0.296
$\gamma = 3$	217.852	0.346	0.353	0.333	0.297	0.460	0.465	0.445	0.404
$\gamma = 3.5$	310.261	0.387	0.397	0.377	0.336	0.485	0.493	0.473	0.430

Notes: Median value of life per QALY (in thousands of dollars) for EHRS cohort denoted by VOL. All other columns report mean CV (λ). Estimates use base year respondent analysis weights. War Babies denoted by WB and Baby Boomers by BB.

3.5.1 Flow intercept

First, we examine the impact of assuming a higher value for the flow intercept \bar{u} . Specifically, we set $\bar{u} = -\log(1.5)$, implying that \$1,500 of consumption is needed for a retiree to prefer life to death. This increases the median value of life to about \$67,000 per QALY in the EHRS cohort. It also increases the estimated mean welfare cost of late-life depression in the EHRS from 8-15% to 11-21% of annual consumption. There are similar increases in welfare estimates for later cohorts and we continue to see a small decline in mean welfare costs across cohorts.

3.5.2 Consumption and leisure utility

We also examine the robustness of results to a more general form of flow utility for consumption and leisure given by:

$$u(c, l, h) = \phi(h) \left[\frac{c^{1-\gamma}}{1-\gamma} \left(1 - (1-\gamma) \frac{\theta\epsilon}{1+\epsilon} (1-l)^{\frac{1+\epsilon}{\epsilon}} \right)^\gamma - \frac{\bar{u}^{1-\gamma}}{1-\gamma} \right] \quad (9)$$

which reduces to our benchmark case with $\gamma = 1$. With $\gamma > 1$ there is more curvature over consumption. These preferences follow those proposed by Trabandt and Uhlig (2011) and Jones and Klenow (2016) which maintain a constant Frisch elasticity of labor supply. Empirical estimates of γ vary substantially and have been argued to be potentially large for retirees, so we examine sensitivity of results to increases in curvature up to $\gamma = 3.5$.

As discussed by Murphy and Topel (2006), as $\frac{1}{\gamma}$ approaches zero, the implied value of life in this framework blows up. In order to gain a sense of this issue in our robustness experiments, the first column in Table 4 shows the median value of life per QALY with higher curvatures. With $\gamma = 2$, the median value of life is \$106,000 per QALY—a high but not completely unreasonable value. The bound on the estimated mean welfare cost of late-life depression in the EHRS rises to 21-34% of annual consumption. When $\gamma = 3.5$, the value of life reaches over \$300,000 per QALY and the welfare bound reaches 38-48%. In a review of 23 studies of the value of life by Ryen and Svensson (2015), only three suggested a mean value above \$150,000. The likely overstated value of life at higher curvatures suggests caution should be taken when interpreting robustness results with high (but empirically plausible) curvature values. Nonetheless, the higher curvature values provide a sense of the robustness of key results and the conservative nature of our benchmark welfare estimates.

4 Conclusion

We estimated a panel VAR model of mental and physical health, labor supply, and consumption using longitudinal data from the Health and Retirement Study. We used the estimated model to repeatedly simulate life-cycle paths for older Americans, with and without the prevalence of late-life depression, given a set of initial age sixty conditions. We estimated an average loss of labor supply of up to 1.1 months, lifetime consumption of up to \$16,000, and quality-adjusted life expectancy of between 0.85 and 2.1 years per person in the EHRS birth cohort. Combining into a single welfare metric, we estimated a bound on the expected welfare loss of depression of 8-15% of annual consumption after age sixty. This amounts to an expected loss in lifetime consumption of approximately \$46,000-\$91,000. In a hypothetical world populated by identical cohorts of size four million at age sixty, this produces a back-of-the-envelope estimate of aggregate welfare loss on the order of \$180-360 billion annually. We also found substantial heterogeneity in the estimated cost with some individuals willing to give up well over 20% of annual consumption to avoid late-life depression. Moreover, while we found a small general decline in average costs over birth cohorts, the welfare burden of depression appears to have become significantly more unequal within cohorts.

This study is not without limitations. Our estimates only include private costs of depression and do not capture public expenses (e.g., Medicare costs) or general equilibrium effects. Moreover, while our statistical model does well in replicating the observed patterns in the data, point estimates cannot be viewed as necessarily causal nor as adhering to any particular unobserved mechanism. This leaves us with only an estimated bound on feasible costs. Our compensating variation measure is also quantitatively sensitive to the choice of curvature in the utility function. Nonetheless, this study's novelty is in estimation of a more comprehensive measure that incorporates life-cycle dynamics to improve our understanding of the welfare costs of late-life depression.

References

- Abrams, L. R. and Mehta, N. K. (2019). Changes in depressive symptoms over age among older americans: differences by gender, race/ethnicity, education, and birth cohort. *SSM-population health*, 7:100399.
- Ambresin, G., Chondros, P., Dowrick, C., Herrman, H., and Gunn, J. M. (2014). Self-rated health and long-term prognosis of depression. *The Annals of Family Medicine*, 12(1):57–65.
- Barry, L. C., Abou, J. J., Simen, A. A., and Gill, T. M. (2012). Under-treatment of depression in older persons. *Journal of affective disorders*, 136(3):789–796.
- Birnbaum, H. G., Kessler, R. C., Kelley, D., Ben-Hamadi, R., Joish, V. N., and Greenberg, P. E. (2010). Employer burden of mild, moderate, and severe major depressive disorder: mental health services utilization and costs, and work performance. *Depression and anxiety*, 27(1):78–89.
- Blazer, D. G. (2003). Depression in late life: review and commentary. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 58(3):M249–M265.
- Bruce, M. L. (2001). Depression and disability in late life: directions for future research. *The American Journal of geriatric psychiatry*, 9(2):102–112.
- Bruce, M. L., Leaf, P. J., Rozal, G. P. M., Florio, L., and Hoff, R. A. (1994). Psychiatric status and 9-year mortality data in the new haven epidemiologic catchment area study. *The American journal of psychiatry*.
- Chiang, H.-H., Guo, H.-R., Livneh, H., Lu, M.-C., Yen, M.-L., and Tsai, T.-Y. (2015). Increased risk of progression to dialysis or death in ckd patients with depressive symptoms: A prospective 3-year follow-up cohort study. *Journal of psychosomatic research*, 79(3):228–232.
- Chu, W., Chang, S.-F., Ho, H.-Y., and Lin, H.-C. (2019). The relationship between depression and frailty in community-dwelling older people: a systematic review and meta-analysis of 84,351 older adults. *Journal of Nursing Scholarship*, 51(5):547–559.
- Chui, H., Gerstorf, D., Hoppmann, C. A., and Luszcz, M. A. (2015). Trajectories of depressive symptoms in old age: Integrating age-, pathology-, and mortality-related changes. *Psychology and Aging*, 30(4):940.
- Ciechanowski, P. S., Katon, W. J., and Russo, J. E. (2000). Depression and Diabetes: Impact of Depressive Symptoms on Adherence, Function, and Costs. *Archives of Internal Medicine*, 160(21):3278–3285.
- Currie, J. and Madrian, B. C. (1999). Health, health insurance and the labor market. *Handbook of Labor Economics*, 3:3309–3416.

- De Groot, M., Anderson, R., Freedland, K. E., Clouse, R. E., and Lustman, P. J. (2001). Association of depression and diabetes complications: a meta-analysis. *Psychosomatic medicine*, 63(4):619–630.
- Doshi, J. A., Cen, L., and Polsky, D. (2008). Depression and retirement in late middle-aged us workers. *Health services research*, 43(2):693–713.
- Fechner-Bates, S., Coyne, J. C., and Schwenk, T. L. (1994). The relationship of self-reported distress to depressive disorders and other psychopathology. *Journal of consulting and clinical psychology*, 62(3):550.
- Feeny, D., Furlong, W., Torrance, G. W., Goldsmith, C. H., Zhu, Z., DePauw, S., Denton, M., and Boyle, M. (2002). Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Medical care*, 40(2):113–128.
- Furlong, W., Feeny, D., Torrance, G., Goldsmith, C., DePauw, S., Zhu, Z., Denton, M., Boyle, M., et al. (1998). Multiplicative multi-attribute utility function for the health utilities index mark 3 (hui3) system: a technical report. Technical report, Centre for Health Economics and Policy Analysis (CHEPA), McMaster University, Hamilton, Canada.
- Geerlings, S., Beekman, A., Deeg, D., Twisk, J., and Van Tilburg, W. (2001). The longitudinal effect of depression on functional limitations and disability in older adults: an eight-wave prospective community-based study. *Psychological Medicine*, 31(8):1361.
- Greenberg, P. E., Fournier, A.-A., Sisitsky, T., Pike, C. T., and Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the united states (2005 and 2010). *The Journal of clinical psychiatry*, 76(2):155–162.
- Hall, R. E. (2009). Reconciling cyclical movements in the marginal value of time and the marginal product of labor. *Journal of Political Economy*, 117(2):281–323.
- Holford, T. R. (1991). Understanding the effects of age, period, and cohort on incidence and mortality rates. *Annual review of public health*, 12(1):425–457.
- Honaker, J. and King, G. (2010). What to do about missing values in time-series cross-section data. *American Journal of Political Science*, 54(2):561–581.
- Horsman, J., Furlong, W., Feeny, D., and Torrance, G. (2003). The health utilities index (hui®): concepts, measurement properties and applications. *Health and Quality of Life Outcomes*, 1(1):54.
- Idler, E. L. and Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior*, pages 21–37.

- Jones, C. I. and Klenow, P. J. (2016). Beyond GDP? Welfare across countries and time. *The American Economic Review*, 106(9):2426–2457.
- Kaplan, R. M. and Bush, J. W. (1982). Health-related quality of life measurement for evaluation research and policy analysis. *Health Psychology*, 1(1):61.
- Karim, J., Weisz, R., Bibi, Z., and ur Rehman, S. (2015). Validation of the eight-item center for epidemiologic studies depression scale (ces-d) among older adults. *Current Psychology*, 34(4):681–692.
- Kok, R. M. and Reynolds, C. F. (2017). Management of depression in older adults: a review. *Jama*, 317(20):2114–2122.
- Lakey, S. L., LaCroix, A. Z., Gray, S. L., Borson, S., Williams, C. D., Calhoun, D., Goveas, J. S., Smoller, J. W., Ockene, J. K., Masaki, K. H., et al. (2012). Antidepressant use, depressive symptoms, and incident frailty in women aged 65 and older from the women’s health initiative observational study. *Journal of the American Geriatrics Society*, 60(5):854–861.
- Lerner, D. and Henke, R. M. (2008). What does research tell us about depression, job performance, and work productivity? *Journal of occupational and environmental medicine*, 50(4):401–410.
- Lewinsohn, P. M., Seeley, J. R., Roberts, R. E., and Allen, N. B. (1997). Center for epidemiologic studies depression scale (ces-d) as a screening instrument for depression among community-residing older adults. *Psychology and aging*, 12(2):277.
- Luppa, M., Sikorski, C., Motzek, T., Konnopka, A., König, H.-H., and G Riedel-Heller, S. (2012). Health service utilization and costs of depressive symptoms in late life—a systematic review. *Current pharmaceutical design*, 18(36):5936–5957.
- Lwin, M. N., Serhal, L., Holroyd, C., and Edwards, C. J. (2020). Rheumatoid arthritis: The impact of mental health on disease: A narrative review. *Rheumatology and therapy*, 7(3):457–471.
- Mayerl, H., Stolz, E., and Freidl, W. (2020). Frailty and depression: Reciprocal influences or common causes? *Social Science & Medicine*, 263:113273.
- McEwen, B. S. (2003). Mood disorders and allostatic load. *Biological psychiatry*, 54(3):200–207.
- Miller, R. and Bairoliya, N. (2021). Health, longevity, and welfare inequality of older americans. (Conditional Accept) *The Review of Economics and Statistics*.
- Mirowsky, J. and Ross, C. E. (1992). Age and depression. *Journal of health and social behavior*, pages 187–205.

- Mohiuddin, S. and Payne, K. (2014). Utility values for adults with unipolar depression: systematic review and meta-analysis. *Medical Decision Making*, 34(5):666–685.
- Moussavi, S., Chatterji, S., Verdes, E., Tandon, A., Patel, V., and Ustun, B. (2007). Depression, chronic diseases, and decrements in health: results from the world health surveys. *The Lancet*, 370(9590):851–858.
- Murphy, K. M. and Topel, R. H. (2006). The value of health and longevity. *Journal of political Economy*, 114(5):871–904.
- Penninx, B. W. (2017). Depression and cardiovascular disease: epidemiological evidence on their linking mechanisms. *Neuroscience & Biobehavioral Reviews*, 74:277–286.
- Penninx, B. W., Leveille, S., Ferrucci, L., Van Eijk, J. T., and Guralnik, J. M. (1999). Exploring the effect of depression on physical disability: longitudinal evidence from the established populations for epidemiologic studies of the elderly. *American journal of public health*, 89(9):1346–1352.
- Rice, N. E., Lang, I. A., Henley, W., and Melzer, D. (2011). Common health predictors of early retirement: findings from the english longitudinal study of ageing. *Age and ageing*, 40(1):54–61.
- Rothermund, K. and Brandtstädter, J. (2003). Depression in later life: cross-sequential patterns and possible determinants. *Psychology and Aging*, 18(1):80.
- Rubio-Guerra, A. F., Rodriguez-Lopez, L., Vargas-Ayala, G., Huerta-Ramirez, S., Serna, D. C., and Lozano-Nuevo, J. J. (2013). Depression increases the risk for uncontrolled hypertension. *Experimental and clinical cardiology*, 18(1):10–2.
- Ryen, L. and Svensson, M. (2015). The willingness to pay for a quality adjusted life year: a review of the empirical literature. *Health Economics*, 24(10):1289–1301.
- Sivertsen, H., Bjørkløf, G. H., Engedal, K., Selbæk, G., and Helvik, A.-S. (2015). Depression and quality of life in older persons: a review. *Dementia and geriatric cognitive disorders*, 40(5-6):311–339.
- Smith, J. P. (1999). Healthy bodies and thick wallets: The dual relation between health and economic status. *The Journal of Economic Perspectives: A Journal of the American Economic Association*, 13(2):144.
- Soysal, P., Veronese, N., Thompson, T., Kahl, K. G., Fernandes, B. S., Prina, A. M., Solmi, M., Schofield, P., Koyanagi, A., Tseng, P.-T., et al. (2017). Relationship between depression and frailty in older adults: A systematic review and meta-analysis. *Ageing research reviews*, 36:78–87.

- Steffick, D. E. (2000). Documentation of affective functioning measures in the health and retirement study. *HRS/AHEAD Documentation Report*, DR-005.
- Stenholm, S., Pentti, J., Kawachi, I., Westerlund, H., Kivimäki, M., and Vahtera, J. (2014). Self-rated health in the last 12 years of life compared to matched surviving controls: the health and retirement study. *PloS one*, 9(9):e107879.
- Stewart, W. F., Ricci, J. A., Chee, E., Hahn, S. R., and Morganstein, D. (2003). Cost of lost productive work time among us workers with depression. *Jama*, 289(23):3135–3144.
- Stuck, A. E., Walthert, J. M., Nikolaus, T., Büla, C. J., Hohmann, C., and Beck, J. C. (1999). Risk factors for functional status decline in community-living elderly people: a systematic literature review. *Social science & medicine*, 48(4):445–469.
- Sutin, A. R., Terracciano, A., Milaneschi, Y., An, Y., Ferrucci, L., and Zonderman, A. B. (2013). The trajectory of depressive symptoms across the adult life span. *JAMA psychiatry*, 70(8):803–811.
- Tampubolon, G. and Maharani, A. (2017). When did old age stop being depressing? depression trajectories of older americans and britons 2002–2012. *The American Journal of Geriatric Psychiatry*, 25(11):1187–1195.
- Trabandt, M. and Uhlig, H. (2011). The Laffer curve revisited. *Journal of Monetary Economics*, 58(4):305–327.
- Turvey, C. L., Wallace, R. B., and Herzog, R. (1999). A revised ces-d measure of depressive symptoms and a dsm-based measure of major depressive episodes in the elderly. *International psychogeriatrics*, 11(2):139–148.
- Vaughan, L., Corbin, A. L., and Goveas, J. S. (2015). Depression and frailty in later life: a systematic review. *Clinical interventions in aging*, 10:1947.
- Wang, P. S., Simon, G., and Kessler, R. C. (2003). The economic burden of depression and the cost-effectiveness of treatment. *International Journal of Methods in Psychiatric Research*, 12(1):22–33.