Beyond Income: Health, Wealth, and Racial Welfare Gaps Among Older Americans

Sayorn Chin*
Colorado State University

Ray Miller
Colorado State University

May 24, 2022

Abstract

We estimate racial disparities in well-being among the older U.S. population using an expected utility framework that incorporates differences in consumption, leisure, health, mortality, and wealth. We find large racial disparities in expected welfare later in life. Moreover, disparity measures based on cross-sectional consumption substantially underestimate racial welfare gaps by ignoring disparities in expected elderly health, wealth, and mortality. Our decomposition exercises show that a majority of the estimated welfare gaps are determined by age sixty initial conditions as opposed to racial differences in dynamic processes after age sixty. This suggests that policies aimed at closing racial gaps in late-life may be more successful and efficient if targeted earlier in the life-cycle. In other words, outside of direct wealth transfers, it may largely be too late to target such interventions directly at older populations.

JEL classifications: I14, J14, J11, J26
Keywords: race, inequality, health, wealth, aging, consumption, mortality

*Corresponding Author: PhD Candidate, Department of Economics, Colorado State University, Fort Collins, CO, 80523. Email: Sayorn.Chin@ColoState.edu.

1. Introduction

Racial inequality remains large and persistent in many social and economic domains (e.g., Darity Jr and Myers Jr, 1998; Pager and Shepherd, 2008; Margo, 2016). Income and consumption have traditionally been the chosen metrics for examining racial economic disparities in the United States. However, additional factors have been more closely examined in recent years. For example, a persistent wealth gap has been identified between White, Black, and Hispanic Americans (Smith et al., 1997; Shapiro and Kenty-Drane, 2005; Aliprantis et al., 2019; Ashman and Neumuller, 2020; Conley, 2000; Bhutta et al., 2020). Importantly, these alternate metrics provide somewhat different pictures of racial inequities. For instance, studies have found that income inequality across racial groups is usually lower than wealth inequality, implying some underestimation of the broader racial well-being gap when only income is considered (Bhutta et al., 2020).

When alternate metrics are broken down by age cohort, the differences in captured inequality are even greater. In particular, research has indicated that wealth inequality may be a significantly better measure than income when examining welfare disparities at older ages (Smith et al., 1997; Bhutta et al., 2020; Ozawa and Tseng, 2000). Other studies have cited inequality in lifespan, health outcomes, and even leisure as major underlying factors of welfare disparity among older populations (Benhabib et al., 2017; Manton, 1987; Lynch, 2008; Adams et al., 2011; Steptoe et al., 2015; Adams et al., 2011; Hribernik and Mussap, 2010; Han and Patterson, 2007; Pollack et al., 2007; Shea et al., 1996; Smith and Egger, 1993). That health disparities matter a great deal at older ages is perhaps unsurprising given that most population level health differences are concentrated in late-life (Deaton and Paxson, 1998; Minkler et al., 2006). The question then remains around the appropriate use of a single metric such as income, wealth, or life expectancy to analyze welfare gaps across racial lines. While each such variable individually contributes to the gaps in racial well-being, it remains unclear if the adoption of such narrowly defined metrics can adequately capture the true welfare inequality between racial groups (e.g., Patton et al., 2016; Lepinteur, 2019; Strife and Downey, 2009). Accounting for the underlying factors contributing to welfare may reveal patterns of inequality that conflict with well-established estimates.

The use of a multidimensional approach to measuring welfare has been adopted by some social scientist when measuring inequality (Maasoumi and Nickesburg, 1983; Rohde and Guest, 2013; Maasoumi, 1986; Maasoumi and Nickelsburg, 1988; Goetz, 1991). Similar to other money metrics of inequality, multidimensional measures create an index based on aggregating attributes of welfare using a social welfare function. This composite measure of welfare combines indicators in their original form that are weighted based on their contribution to overall welfare (Maasoumi, 1986; Manduca, 2018). Individual utility functions are used when creating the aggregate inequality index and the decomposition of these aggregate measures allows for the estimation of the relative contribution of each measure to total welfare inequality.

Aggregate inequality measures have been found to be more informative than the unitary analysis, and more successfully reflect the distribution changes within and between demographic groups in the Unites States (Maasoumi and Nickelsburg, 1988; Rohde and Guest, 2013). These measures, however, fail to account for dynamic spillovers across indicators, which would not be captured with the ad hoc aggregation of individual welfare indicators. Furthermore, the choice of weights applied to each indicator is subjective to the researcher and is required to be sample specific. That is, it is difficult to unambiguously determine how important one indicator is relative to another and how much a surplus on one criterion should be used to compensate for a shortfall in another.

The aim of this paper is to estimate racial welfare inequality among the older U.S. population using an expected utility framework that incorporates differences in consumption, leisure, health, wealth, and mortality. We take a life-cycle approach to better quantify aggregate inequality by incorporating contemporaneous and dynamic spillovers across all modeled outcomes at the individual level. This is an important departure from estimates derived using aggregate models as they may fail to capture the inter-linkages among these factors. For example, if economic and health outcomes are strongly correlated, racial disparity measures based on cross-sectional income or consumption might underestimate the aggregate racial welfare inequality and would only be presenting a part of the bigger story. Furthermore, the share of Americans over age 65 is projected to reach 20% by 2030 and continue to rise thereafter (Vespa et al., 2018). This highlights the importance of understanding the underlying factors of inequality among older Americans. Our measure of inequality is constructed using a similar framework as Miller and Bairoliya (2021). Specifically, we propose a panel vector autoregressive (VAR) model to approximate the joint late-life evolution of consumption, leisure, health, mortality, and wealth (valued as bequests at death). Throughout the paper, we will use the terms: wealth and bequest interchangeably, but they convey the same meaning. We estimate parameters of the model using longitudinal data from the Health and Retirement Study (HRS) supplemented with data from the Consumption and Activities Mail Survey (CAMS). Together, these provide a long and rich panel (1992-2014) for our analysis. We then use the estimated system to simulate potential outcome paths by race for a sub-sample of HRS respondents starting from age sixty. Finally, these paths are embedded in a simple expected utility framework to compute a forward-looking ex-ante metric of welfare (measured in consumption equivalents) for each individual in our sample at age sixty. As our measure incorporates individual expectations about outcomes over the entirety of remaining life, it provides a useful single metric of ex-ante well-being at older ages.

Based on the data available in the HRS, we estimate welfare gaps among study participants who self-reported as non-Hispanic Black (hereafter, Black), Hispanic, and non-Hispanic White (hereafter, White). Our main findings can be summarized as follows:

1. Ex-ante age sixty welfare was significantly higher among White HRS respondents. Mean welfare for Black respondents was 38% that of White respondents (Black-White welfare ratio of 0.38). The analogous estimate for Hispanic compared to White respondents was 34% (Hispanic-White welfare ratio of 0.34).

2. Expected annual consumption gaps over remaining life explain the

largest share of the welfare gaps between races, accounting for roughly 60-70% of the overall gaps. The mean Black-White welfare ratio based only on consumption was estimated to be 0.62 (or 62%). The analogous estimate for the Hispanic-White ratio was 0.51 (or 51%).

3. Black and Hispanic respondents retired earlier than White respondents overall, but these differences had only small effects on our aggregate measure of racial welfare gaps.

4. Health and longevity (life expectancy) were important for overall welfare gaps. Accounting for longevity differences was more important for Black participants, decreasing the estimated mean Black-White welfare ratio by 12 percentage points (pp). In contrast, the welfare cost of living in poor health was more important for Hispanic participants, decreasing the estimated Hispanic-White welfare ratio by 7 pp.

5. Smaller financial bequests (or wealth at death) are nearly as important to estimated welfare gaps as health and longevity. Adjusting for bequests lowers the Black-White welfare ratio an additional 10 pp and the Hispanic-White ratio an additional 9 pp.

Further simulations in which the most racially dispersed health risk factors (hypertension and diabetes) are counterfactually eliminated in late-life only marginally closes overall welfare gaps. Moreover, decomposition exercises show that a majority of the estimated welfare gaps are determined by age sixty initial conditions as opposed to racial differences in dynamic processes after age sixty. This suggests that policies aimed at closing racial gaps in late-life may be more successful and efficient if targeted earlier in the life-cycle. In other words, outside of direct wealth transfers, it may largely be too late to target such interventions directly at older populations.

This study makes several contributions to the existing literature on measuring racial inequality. First, most previous studies carried out estimation in a cross-sectional or clinical setting (Aliprantis et al., 2019; Rohde and Guest, 2013; Maasoumi, 1986; Maasoumi and Nickelsburg, 1988). Our study employs a longitudinal panel that captures both contemporaneous and dynamic spillover effects across several economic and health outcomes. This allows for a more comprehensive measure that incorporates the cumulative contribution of each factor to welfare. Our use of microsimulations from a model of life-cycle dynamics also allows us to construct a measure at the individual level within a larger representative sample, so we can examine the entire distribution of welfare. Our forward-looking framework also incorporates differences in the uncertain evolution of outcomes over remaining life, providing a more complete measure of racial welfare inequality when compared to other multidimensional measures (Maasoumi, 1986; Maasoumi and Nickelsburg, 1988; Rohde and Guest, 2013). We also use a broader indicator of health, incorporating several morbidities and physical limitations, in addition to self-reported health.

Finally, we contribute to the literature that more specifically focuses on racial inequality among older populations. Existing studies in this area have generally focused on a single metric like wealth (Smith et al., 1997; Ozawa and Tseng, 2000; Williams et al., 2001; Martin and Soldo, 1997). We add to this line of research by examining racial inequality among older Americans using a dynamic and multi-dimensional metric. Our simulations also shed light on how successful early versus late-life interventions may be in impacting racial welfare gaps at older ages.

2. Data and Methods

2.1 Data

We used data from the Health and Retirement Study (HRS), a nationally representative, biennial longitudinal survey of multiple cohorts that tracks individuals age 50 and above in the U.S. The HRS data consists of seven birth cohorts: the initial HRS cohorts, born between 1931 and 1941; the Study of Assets and Health Dynamics Among the Oldest Old (AHEAD) cohort, born before 1924; Children of Depression (CODA) cohort, born between 1924 and

1930; War Baby (WB) cohort, born between 1942 and 1947; and Early, Mid, and Late Baby Boomer, born after 1947. Our primary data source was the publicly available RAND HRS Longitudinal File. We obtained data on race, health, mortality, economic outcomes, and other individual characteristics including age, education, gender, birth cohort, region, and occupation. Next, we discuss in more detail the variables used in the analysis.

2.1.1 Race/Ethnicity Variables

Race/ethnicity was solicited in the HRS by asking individuals "Do you consider yourself Hispanic or Latino?" and "Do you consider yourself primarily White or Caucasian, Black or African American, American Indian or Asian, or something else?" In our analysis, we classified race/ethnicity into three categories based on the response provided: White, non-Hispanic; Black, non-Hispanic; and Hispanic. We excluded American Indian or Alaskan Native, Asian or Pacific Islander, and Unknown from the analysis since these categories are not representative.

2.1.2 Health Outcomes

In addition to race, our model further incorporates data on co-morbidities. Data on co-morbidities includes binary indicators for doctor's diagnosis of eight (8) specific health problems as well as reported difficulties with activities of daily life (ADLs). The health problems incorporated are (1) high blood pressure and hypertension; (2) diabetes; (3) cancer or any kind of malignant tumor, excluding melanoma; (4) chronic lung disease excluding asthma, chronic bronchitis or emphysema; (5) heart attack, coronary heart disease, angina, congestive heart failure or other heart related problems; (6) stroke or transient ischemic attack; (7) emotional, nervous or psychiatric problems; and (8) arthritis or rheumatism. Difficulty with ADLs include activities such as bathing, getting dressed and toileting. These health metrics are mostly observed and are somewhat objective measures of health outcomes. However, self-rated health outcomes, where individuals are asked to rank their health on some scale, have been shown to be a good predictor of mortality, after controlling for other health conditions, health behavior and socioeconomic characteristics (Idler and Benyamini, 1997). As such, we included an indicator of self-rated health status on a five-point scale from poor (one) to excellent (five). This is important for testing if people have significant private information about their health above diagnosis given by a doctor or other observable indicators of health.

2.1.3 Economic Outcomes

Annual hours worked was estimated using self-reported data on weekly hours and number of weeks worked. For the purposes of this study, retired individuals are defined as those with less than 500 hours of work per year. To estimate individual consumption, we use consumption data provided by the Consumption Activities Mail Survey (CAMS). A constructed estimate of total household consumption derived from household spending data on durables, non-durables, transportation and housing collected from 2001-2015 was untilized. From this estimate found in RAND 2017 CAMS data file, out-of-pocket health spending was subtracted and this adjusted household consumption spending was divided by the number of household members to derive our individual consumption measure. We merge consumption data from CAMS with data from the previous core HRS wave. CAMS data is available for about 20% of HRS respondents from 2000-2014. To address missing consumption data, we follow Miller and Bairoliya (2021) and apply the multiple imputation method, proposed by Honaker and King (2010) for cross-sectional time-series data, which relies on closely related available data such as wealth and income (see Online Appendix for details). Finally, we estimate expected bequests using estimated household asset wealth from the RAND HRS data file. These assets include financial, housing, and other durable wealth (e.g. vehicles, jewelry, etc).

2.2 Simulation Model and Estimation

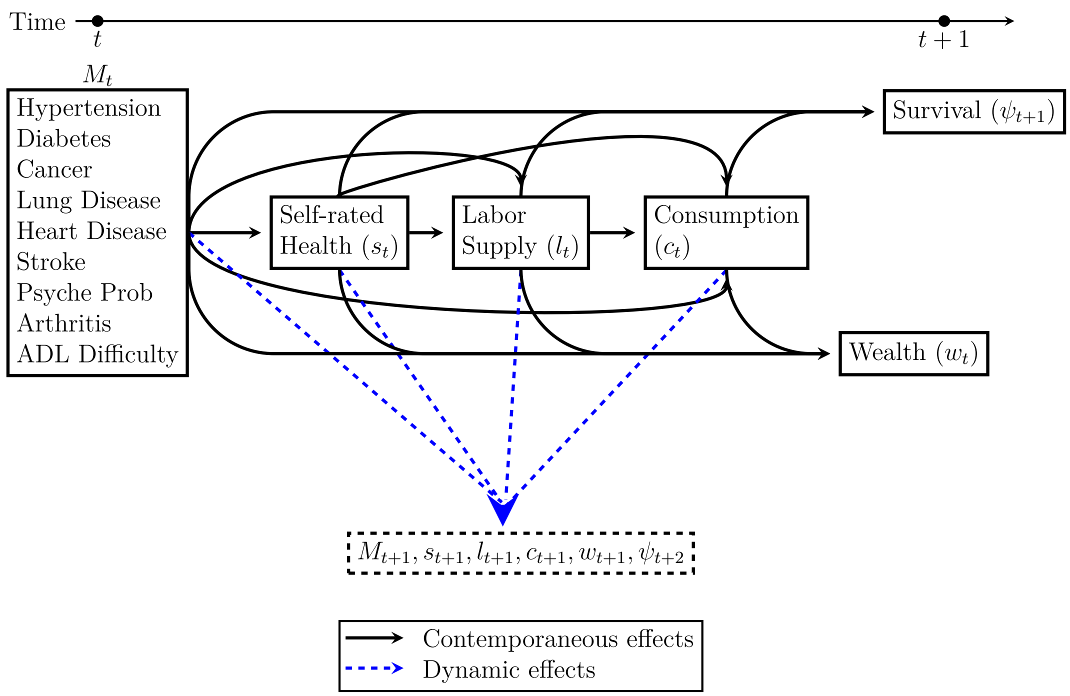
We propose a panel vector autoregressive (VAR) model to approximate the joint late-life evolution of consumption, leisure, health, mortality, and wealth (valued as bequests at death) across racial groups. This allows us to (1) better quantify the racial welfare gaps in a given population and (2) examine how much of the racial welfare gaps can potentially be closed through a number of counterfactual experiments. Our basic model is illustrated in Figure 1. At the beginning of each time period, morbidity status is updated based on random shocks and exogenous characteristics of an individual. The individual then updates their self-rated health. Self-rated health then affects current period labor supply (i.e. the retirement decision), which further affects consumption, wealth, and the probability of survival to the next time period. These are contemporaneous effects. It is important to note that it is assumed that morbidities can contemporaneously affect consumption and labor supply directly. Additionally, our model allows current period morbidity, self-rated health and labor supply, to dynamically affect future outcomes by including lagged effects of these variables. An important aspect of including lagged effects is that it allows for more recent diagnoses of a morbidity to have a greater impact on health changes than long-standing diagnoses.

The life-cycle dynamics are modeled as a statistical process and will be estimated directly from data. Though modeling the explicit maximization of lifetime utility would allow better policy analysis, this approach requires solving an intertemporal structural model of endogenous savings, labor supply and multiple morbidity or health outcomes and investments, which is very challenging. Given that the goal of this paper is to develop a welfare measure that reflects actual data, we believe that a statistical process more appropriate in this regard.

2.2.1 Panel VAR Representation

The following VAR(1) demonstrates the key features of the framework. Y_{it} is a vector of outcomes for individual i at time t. This includes log consumption

Figure 1: Simulation Model With One Period Lag



c, retirement indicator r, self-rated health s, cube root of wealth w, and our $n = 9$ morbidity states given by n × 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. $A_{it} = A \quad \forall i, t$).

Our model is estimated in "five blocks" of outcomes: the morbidity block, the self-rated health block, the retirement block, the consumption block, and the wealth block. This is presented in the following matrix form:

$$
\begin{array}{cc}
 & \begin{array}{c|cccc} \overbrace{\qquad}^{n} & & \overbrace{\qquad\qquad\qquad\qquad}^{4} & & \end{array} \\
\begin{array}{c} n \\ \\ 4 \\ \\ \\ \end{array} & \left[\begin{array}{c|cccc} -A_{11} & -A_{12} & -A_{13} & -A_{14} & -A_{15} \\ \hline -A_{21} & 1 & -a_{23} & -a_{24} & -a_{25} \\ -A_{31} & -a_{32} & 1 & -a_{34} & -a_{35} \\ -A_{41} & -a_{42} & -a_{43} & 1 & -a_{45} \\ -A_{51} & -a_{52} & -a_{53} & -a_{54} & 1 \end{array}\right]
\end{array}
\left[\begin{array}{c} M_{it} \\ \hline s_{it} \\ r_{it} \\ c_{it} \\ w_{it} \end{array}\right]
=
\begin{array}{c}
\begin{array}{c|cccc} \overbrace{\qquad}^{n} & & \overbrace{\qquad\qquad\qquad\qquad}^{4} & & \end{array} \\
\left[\begin{array}{c|cccc} B_{11} & B_{12} & B_{13} & B_{14} & B_{15} \\ \hline B_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ B_{31} & b_{32} & b_{33} & b_{34} & b_{35} \\ B_{41} & b_{42} & b_{43} & b_{44} & b_{45} \\ B_{51} & b_{52} & b_{53} & b_{54} & b_{55} \end{array}\right]
\end{array}
\left[\begin{array}{c} M_{it-1} \\ \hline s_{it-1} \\ r_{it-1} \\ c_{it-1} \\ w_{it-1} \end{array}\right]
+
\left[\begin{array}{c} \epsilon_{1,it} \\ \hline \epsilon_{2,it} \\ \epsilon_{3,it} \\ \epsilon_{4,it} \\ \epsilon_{5,it} \end{array}\right],
$$

where $n \times n$ matrix A_{11} has main diagonal terms scaled to one. The causal pathways we propose suggest a block recursive system and this is illustrated in Figure 1. 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 retirement block, and $a_{45} = 0$ in the consumption block. This assumes that the contemporaneous causal pathway runs from morbidities to self-rated health to retirement to consumption to wealth. Health and retirement are allowed to affect future outcomes through lagged effects. Lagged consumption is allowed to impact future wealth, but consumption and wealth are otherwise assumed not have lagged effects.[1] By applying such block triangulation of the system, we eliminate simultaneity across blocks and allow for block-by-block estimation.

2.2.2 Exogeneous Characteristics

We further 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:

$$AY_{it} = BY_{it-1} + CX_{it} + \epsilon_{it}. \tag{1}$$

Exogenous characteristics include dummies for age, education, gender, census division, census occupation code, and birth cohort. We also include a linear trend for calendar year and a post-2008 indicator to help controlling for the impact of the great recession on outcomes. Lastly, we include a time invariant individual unobserved endowment in the consumption equation (π^c) and in the wealth equation (π^w). The endowments π are modeled as fixed effects with

[1] i.e. $B_{14} = B_{15} = b_{24} = b_{25} = b_{34} = b_{35} = b_{45} = 0$

no restriction on the correlation with other model regressors. This requires the exclusion of the other time invariant regressors–education, gender, census division, occupation code and birth cohort from the consumption and wealth equations. The unobserved individual effect helps maintain the appropriate variance in consumption and wealth across time by acting as a person specific drift in the autoregressive process. The socioeconomic characteristics listed above are included instead of additional individual fixed effects in the health and retirement equations because 1) difficulties in estimating dynamic panel models with fixed effects due to the fact that morbidities and retirement are absorbing states and self-rated health is ordinal and 2) the simpler model does well in replicating the dynamics of health and retirement (see Online Appendix for more details). The resulting exogenous effects then take the following form:

$$CX_{it} = \begin{array}{r} n\left\{\vphantom{C_{11}}\right. \\ \\ \\ \\ \\ \end{array} \underbrace{\left[\begin{array}{ccccccccccc} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} & C_{17} & C_{18} & C_{19} & 0 & 0 \\ \hline c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} & c_{27} & c_{28} & c_{29} & 0 & 0 \\ c_{31} & c_{32} & c_{33} & c_{34} & c_{35} & c_{36} & c_{37} & c_{38} & c_{39} & 0 & 0 \\ c_{41} & 0 & 0 & 0 & 0 & 0 & 0 & c_{48} & c_{49} & c_{410} & 0 \\ c_{51} & 0 & 0 & 0 & 0 & 0 & 0 & c_{58} & c_{59} & 0 & c_{511} \end{array}\right]}_{(n+4)\times k} \underbrace{\begin{bmatrix} Age_{it} \\ Education_i \\ Gender_i \\ Race_i \\ Division_i \\ Occupation_i \\ Cohort_i \\ Year_t \\ Post_t \\ \pi_i^c \\ \pi_i^w \end{bmatrix}}_{k\times 1}.$$

Lastly, we normalize c_{410} and c_{511} to one to allow identification of the unobserved fixed effects in the consumption and wealth blocks.

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

$$M^*_{it} = -A_{11}^{-1}B_{11}M_{it-1} - A_{11}^{-1}B_{12}s_{it-1} - A_{11}^{-1}B_{13}r_{it-1} - A_{11}^{-1}\left[C_{11},\ldots,C_{19}\right]X_{it} - A_{11}^{-1}\epsilon_{1,it}.$$

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

$$M^*_{it} = \hat{B}_1 M_{it-1} + \hat{B}_2 s_{it-1} + \hat{B}_3 r_{it-1} + \hat{C}X_{it} + e_{it}.$$

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

Morbidity outcomes are binary resulting in forecasting of the measures not being a true linear VAR process. As such, we assume a continuous latent variable $m^\star$ underlies each observed outcome such that:

$$\begin{aligned} m_{j,it} &= 0 \quad if \quad m^\star_{j,it} \leq 0 \\ m_{j,it} &= 1 \quad if \quad m^\star_{j,it} > 0 \end{aligned}$$

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

$$\begin{bmatrix} m^\star_{1,it} \\ \vdots \\ m^\star_{n,it} \end{bmatrix} = \begin{bmatrix} \hat{b}_{11} & \cdots & \hat{b}_{1n} \\ \vdots & \ddots & \vdots \\ \hat{b}_{n1} & \cdots & \hat{b}_{nn} \end{bmatrix} \begin{bmatrix} m_{1,it-1} \\ \vdots \\ m_{n,it-1} \end{bmatrix} + \hat{B}_2 s_{it-1} + \hat{B}_3 r_{it-1} + \hat{C}X_t + \begin{bmatrix} e_{1,it} \\ \vdots \\ e_{n,it} \end{bmatrix}. \tag{2}$$

It is important to note that each latent morbidity variable is determined by lagged values of the other observed self-rated health and morbidity states.

This morbidity block of equations is in the form of a multivariate probit model.

2.2.4 Self–Rated Health

Self-rated health is measured on a five point scale, so, similar to morbidity outcomes, forecasting the measure is not a linear VAR process. We again assume a continuous latent variable $s^{\star}$ underlies the observed self-rated health state such that the relevant equation given in system (1) can be explicitly written as:

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

with the observed health state defined as:

$$s_{it} = \delta \;\; if \;\kappa_{\delta-1} < s_{it}^{\star} < \kappa_{\delta} \;\; for \;\delta = 1, \ldots, 5$$

for cut-points $(\kappa_0, ..., \kappa_5)$ with δ = 1 representing the worst health state (poor) and $\delta = 5$ the best health state (excellent). To incorporate the persistence in general health shocks over the life-course, we assume that latent self-rated health depends on the lagged value of the observed self-rated health category. 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 Retirement

As retirement is a binary outcome, we again assume a continuous latent variable $r^{\star}$ underlies the observed outcome such that:

$$r_{it} = 0 \quad if \quad r_{it}^{\star} \leq 0$$
$$r_{it} = 1 \quad if \quad r_{it}^{\star} > 0.$$

Conditional on working the previous period (and setting $b_{33} = 0$), the retirement model as defined in system (1) is given by:

$$r_{it}^{\star} \quad = \quad A_{31}M_{it} + a_{32}s_{it} + B_{31}M_{it-1} + b_{32}s_{it-1} + [c_{31}, \ldots, c_{39}]\,X_{it} + \epsilon_{3,it}. \quad (4)$$

Retirement is influenced by current and lagged values of self–rated health and specific morbidities and exogenous individual characteristics. We assume ϵ_3 is an iid shock with standard normal distribution implying the retirement model has a standard probit structure.

2.2.6 Consumption and Wealth

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

$$\begin{aligned} c_{it} = {} & A_{41}M_{it} + a_{42}s_{it} + a_{43}r_{it} + B_{41}M_{it-1} + b_{42}s_{it-1} + b_{43}r_{it-1} \\ & + b_{44}c_{it-1} + c_{41}Age_{it} + c_{48}Year_t + c_{49}Post_t + \pi_i^c + \epsilon_{4,it}. \end{aligned} \quad (5)$$

Similarly, the equation for wealth is given by:

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

These are both standard linear dynamic panel data models with lagged dependent variable and individual level fixed effects (π). These equations may also be estimated independently of other blocks with all structural parameters identified including the variance of ϵ_4 and ϵ_5.

2.2.7 Mortality

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

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

where $I(.)$ 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, and retirement to influence survival probability.

2.2.8 Simulations

Equipped with our forecasting model, our empirical analysis involves three steps.

1. We use data from the HRS to estimate the parameters of the forecasting model. Here we use data on all individuals aged fifty and older from all available waves of the HRS from 1992-2014 (35,889 unique individuals and 216,626 total individual-year observations). See Online Appendix for details on model estimation procedures and results.

2. Using the parameter estimates and age sixty data as initial conditions, we repeatedly simulate remaining life-cycle paths for mortality, health, consumption, and leisure for a sub-sample of the HRS respondents. This simulation sample includes all individuals in the initial HRS cohort with age sixty data and requisite lagged data for simulations. See Online Appendix for details on initial condition descriptives, sampling weights and representativeness, and simulation procedure.

3. We embed the simulated data within our expected utility framework (detailed in the following section) to construct a measure of ex-ante welfare at age sixty for each individual in our simulation sample.

3. Welfare Measure

We extend and modify the measure proposed by Miller and Bairoliya (2021) to include the potential gains in welfare from leaving bequests. We begin by defining expected (remaining) lifetime utility at age j for individual i as:

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

where c is consumption (in thousands of dollars), l leisure, h health, b bequests, and Ψ is a survival indicator. We assume log utility over consumption and additive separability with leisure. This allows for a simply decomposition of results. We also report some robustness checks where we relax these assumptions. Health measure h is a vector of indicators for each modelled morbidity and self-rated health. Utility from consumption and leisure is assumed to be scaled by health function $\phi(h) \in [0, 1]$. Note that $\phi(h) = 1$ represents utility for a person in perfect health and $\phi(h) = 0$ represents utility for a person who is dead. Combining the survival indicator with the health function gives a measure of quality-adjusted life years (QALYs). For example, $\psi\phi(h) = 1$ represents a year of life with no adverse health conditions. Furthermore, we consider the potential gains in welfare from leaving bequests as it could be quantitatively important in driving inequalities across racial groups as bequests can be large and are likely correlated with health and consumption.

We use a consumption-equivalent variation measure of welfare. Specifically, welfare for individual i at age j is defined to satisfy the condition:

$$U_{ij} = E\left[\sum_{a=j}^{J} \psi_{ma}\beta^{a-j}\phi(h_{ma})[\bar{u} + log(\lambda_{ij}) + v(l_{ma})] + (1-\psi_{ma})\beta^{a-j}\zeta(b_{ma})\right]$$

where ψ_m, h_m, l_m, and b_m denote a chosen reference level of survival, health, leisure, and bequests. Welfare λ_{ij} is thus defined as the fixed annual consumption that—when combined with the reference health, leisure, survival, and bequest profiles—yields the same expected lifetime utility as the outcome profiles of individual i. For example, if $\lambda_{ij} = 20$, this implies the individual would be ex-ante indifferent between receiving their own stochastic outcome profiles moving forward or receiving $\$20,000$ in annual consumption along with the reference profiles for health, leisure, bequests, and survival.

The welfare condition may also be rearranged to yield the following additive decomposition:

$$log\left(\lambda_{ij}\right) = \tilde{\psi} \sum_{a=j}^{J} \beta^{a-j} \left[E\left[\psi_{ma}\phi\left(h_{ma}\right)\right] E_{\psi}\left[log\left(c_{ia}\right)\right] + \Phi\right] \tag{8}$$

$$+ \tilde{\psi} \sum_{a=j}^{J} \beta^{a-j} E\left[\psi_{ma}\phi\left(h_{ma}\right)\right] \left(E_{\psi}\left[\nu\left(l_{ia}\right)\right] - E_{\psi}\left[\nu\left(l_{ma}\right)\right]\right) \tag{9}$$

$$+ \tilde{\psi} \sum_{a=j}^{J} \beta^{a-j} \left(E\left[\psi_{ia}\right] - E\left[\psi_{ma}\right]\right) E_{\psi}\left[\phi\left(h_{ma}\right)\right] E_{\psi}\left[u_{ia}\right] \tag{10}$$

$$+ \tilde{\psi} \sum_{a=j}^{J} \beta^{a-j} \left(E_{\psi}\left[\phi\left(h_{ia}\right)\right] - E_{\psi}\left[\phi\left(h_{ma}\right)\right]\right) E\left[\psi_{ia}\right] E_{\psi}\left[u_{ia}\right] \tag{11}$$

$$+ \tilde{\psi} \sum_{a=j}^{J} \beta^{a-j} E\left[(1-\psi_{ia})\zeta(b_{ia}) - (1-\psi_{ma})\zeta(b_{ma})\right] \tag{12}$$

where

$$\begin{aligned}\Phi =& \left(E\left[\psi_{ia}\phi\left(h_{ia}\right)u_{ia}\right] - E\left[\psi_{ia}\phi\left(h_{ia}\right)\right] E_{\psi}\left[u_{ia}\right]\right) \\ &- \left(E\left[\psi_{ma}\phi\left(h_{ma}\right)\nu\left(l_{ma}\right)\right] - E\left[\psi_{ma}\phi\left(h_{ma}\right)\right] E_{\psi}\left[\nu\left(l_{ma}\right)\right]\right).\end{aligned}$$

The first term in (8) is expected utility from consumption weighted by the reference quality-adjusted life expectancy. The Φ term is an adjustment for uncertainty over the life-cycle. Combined, these provide an individual's consumption-equivalent welfare before adjusting for expected leisure, life expectancy, health, or bequests. The term (9) is the welfare adjustment for leisure—–the expected utility difference in leisure weighted by the reference quality-adjusted life expectancy. The correction term (10) is the difference in life expectancy weighted by how much a life year is worth—–the expected flow utility from outcome bundles of individual i. Term (11) corrects for expected health differences between individual i and the reference over remaining life. Finally, (12) adjusts welfare for differences in expected bequests.

3.1 Calibration

For calibrating preference parameters, we assume health utility depends linearly on our health state vector: $\phi(h_t) = \gamma h_t$. We then follow Miller and Bairoliya (2021) in using the Health Utilities Index Mark 3 (HUI3) instrument to calibrate the vector of utility weights γ (see Online Appendix for details). The HUI3 has been extensively used in the health utility literature (Furlong et al., 1998; Feeny et al., 2002; Horsman et al., 2003) and was collected for a subset of HRS respondents in year 2000.

We normalize leisure to one for retired individuals and set leisure for workers to 0.66.[2] Preferences over leisure are given by $v(l) = -\frac{\phi_\epsilon}{1+\epsilon}(1-l)^{\frac{1+\epsilon}{\epsilon}}$, where ϵ is a constant Frisch elasticity of labor supply. We follow Jones and Klenow (2016) and set $\epsilon = 1$ and derive a benchmark disutility weight $\theta = 7.82$ such that the marginal cost of leisure is equated to the marginal benefit for the median individual in our sample. We also choose a discount factor $\beta = 0.98$, which corresponds to an annual discount rate of one percent. We follow De Nardi (2004) and define preferences for bequests by $\zeta(b) = \Phi_1\left(1 + \frac{b}{\Phi_2}\right)^{1-\sigma}$. Here, Φ_1 reflects the strength of the bequest motive and Φ_2 measures the extent to which bequests are a luxury good. We follow De Nardi (2004) and set $\Phi_1 = -9.5$, $\Phi_2 = 11.6$, and $\sigma = 1.5$ for our benchmark calibration.

With preferences as defined above, as long as flow intercept $\bar{u}$ plus log consumption is positive, a retired individual will prefer life to death. We set $\bar{u} = -log(2)$, implying that \$2,000 of consumption is needed for a retiree to maintain positive flow utility. This is approximately 10% of mean annual consumption in our sample which has been argued to be a reasonable parameterization of the flow intercept (Murphy and Topel, 2006). This value of $\bar{u}$ also yields a

[2] Assume an endowment of $5,840$ hours per year (16 hours a day $\times$ 365 days) and workers supply $2,000$ of these hours to labor. Leisure is set to the remaining proportion of time $1-(2000/5,840) = 0.66$.

median value of remaining life for sixty year olds of about $60,000 per QALY in our simulation sample. This median value is in the range of typical values reported in the literature (Ryen and Svensson, 2015; Kaplan and Bush, 1982).

3.2 Reference Outcomes

We must specify our reference profiles for welfare calculations. The same reference profiles are used to calculate welfare for all individuals. We choose retirement by age sixty as our reference for leisure (i.e. full leisure from age sixty onward). The standard approach for calculating health-adjusted welfare equivalents is to use a notion of "normal" or "good" health as the reference (Fleurbaey, 2005, 2009; Fleurbaey and Gaulier, 2009; Schokkaert et al., 2013; Fleurbaey et al., 2013; Samson et al., 2018). The logic is that when two individuals are in good health, we can compare them based only on consumption differences. In this spirit, we follow Miller and Bairoliya (2021) and choose a constant reference health level of $\phi(h_{ma}) = 0.8$ and a reference sixty year-old life expectancy of 24 years. We also conduct a robustness check using a longer reference life expectancy. Finally, we choose a reference bequest level of $\$500,000$. So to summarize, we are are assuming that the welfare of age 60 retirees knowing they will live to age 84 in "good" health leaving a bequest of $\$500,000$ can be compared solely in terms of expected consumption profiles.

4. Welfare Results

Our simulation sample for welfare analysis across racial groups includes (1) sample age sixty descriptive statistics, (2) model estimates, (3) mean outcomes and welfare measure, (4) decomposition exercise, (5) impulse responses, and (6) robustness and sensitivity. We discuss these results using the EHRS cohort as our benchmark group as it is the earliest of the seven and contains the longest panel of available data.

4.1 Descriptive Statistics

Table 1 provides a summary of initial (age sixty) conditions in the simulation sample by racial group. Black and Hispanic respondents reported higher prevalence of hypertension, diabetes, stroke, arthritis, and difficulty with ADLs than White respondents (with stroke and arthritis being the exception for Hispanic respondents). For example, reported diabetes averaged 23% for Black respondents and 19% for Hispanic respondents compared to only 10% for White respondents—a 2.3-fold and 1.9-fold difference, respectively. For self-reported health, 14% of Black respondents and 19% of Hispanic respondents reported poor health status but only 6% of White respondents. Black and Hispanic respondents retired earlier than White respondents overall—average of 56% for Black respondents and 60% for Hispanic respondents compared to 50% for White respondents. Moreover, cross-sectional consumption at age sixty averaged \$18,350 for Black and \$14,770 for Hispanic respondents compared to \$30,190 for White respondents—these differences are 1.6-fold and 2-fold, respectively. Finally, 48% of Black and 71% of Hispanic respondents reported less than a high school education at age sixty with a corresponding number among White respondents of only 24%.

4.2 Model Estimates

This section provides select results from our simulation model to gain a broader sense of the correlation between race and other outcomes in the data. Specifically, Figure 2 plots the estimated average marginal effects of race on health and retirement indicators. Relative to White respondents, Black and Hispanic respondents are associated with an increased probability of hypertension, diabetes, stroke, difficulty with ADLs, self-rated poor health, and early retirement (with stroke being the exception for Black respondents). For example, relative to White respondents, Black and Hispanic respondents are associated with a marginal increase in the probability of hypertension of about 2.8 and 1.8 percentage points (pp), respectively. Moreover, we find that race, conditional on morbidities, is associated with self-rated health. For instance, the average

Table 1: Simulation Sample Age Sixty Descriptive Statistics by Race

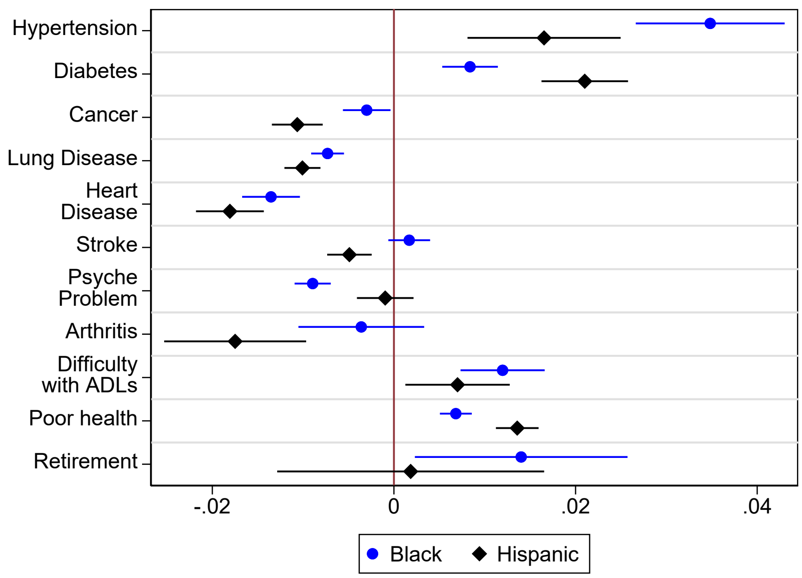
	White	Black	Hispanic
Individuals	2,339	536	235
Hypertension (%)	35.27	59.85	37.76
Diabetes (%)	10.00	22.79	19.30
Cancer (%)	7.05	5.48	5.07
Lung disease (%)	7.66	4.84	3.63
Heart disease (%)	14.04	13.03	10.23
Stroke (%)	2.66	5.53	1.63
Psyche problem (%)	7.18	6.36	12.24
Arthritis (%)	44.81	47.19	41.30
Difficulty with ADLs (%)	9.85	19.89	25.26
Self-rated health (%)			
Poor	5.75	13.51	19.11
Fair	12.96	25.71	31.64
Good	28.00	29.83	29.34
Very good	34.42	20.17	13.12
Excellent	18.87	10.78	6.78
Retired (%)	50.25	55.55	60.24
Annual consumption ($1000s, mean)	30.19	18.35	14.77
Male (%)	47.09	45.08	37.49
Education (%)			
<HS	23.92	48.03	70.99
HS	36.01	27.05	16.51
Some College	20.45	16.37	7.25
College	19.61	8.54	5.25

Notes: Respondents from the initial HRS cohort. Estimates using base year respondent analysis weights. Consumption is reported in real 2010 dollars. Source: HRS.

marginal increase in the probability of reporting poor health is about 0.8 pp for

Black respondents and 1.7 pp for Hispanic respondents.

Figure 2: Average Marginal Effect of Race on Health and Retirement Probabilities



Notes: Dependent variables across rows. White non-Hispanics are the reference group. Spikes indicate 95% confidence intervals.

4.3 Welfare Gaps in EHRS Cohort

In this analysis, we examine mean outcomes and the distribution of our welfare measure across racial groups of sixty year olds from the EHRS cohort presented in Table 2. Panel A shows mean consumption, retirement, life expectancy, QALE, and expected bequests at age sixty. Panel B shows the cumulative contribution of each factor to our welfare measure. In addition, we present the mean Black-White and Hispanic-White outcome and welfare ratios.

Examining mean outcomes in Panel A, age sixty annual consumption for Black respondents is about 61% that of White respondents (Black-White ratio of 0.61). The analogous estimate for Hispanic compared to White respondents is about 49% (Hispanic-White ratio of 0.49). Black and Hispanic

Table 2: Outcomes and Welfare by Race

Measure	**Mean**				
	White	Black	Hispanic	Black-White-Ratio	Hispanic-White-Ratio
Panel A: Outcomes					
Consumption	30.190	18.346	14.773	0.608	0.489
Retired	0.502	0.555	0.602	1.106	1.199
Life Exp.	21.540	18.556	21.137	0.861	0.981
QALE	16.866	13.572	14.296	0.805	0.848
Bequests	396.459	101.327	95.810	0.256	0.242
Panel B: Welfare					
Consumption	23.817	14.752	12.120	0.619	0.509
Leisure	22.046	13.853	11.426	0.628	0.518
Life Exp.	22.677	11.476	11.399	0.506	0.503
Health	20.178	9.583	8.636	0.475	0.428
Bequests	18.322	6.946	6.340	0.379	0.346

Notes: Estimates use base year respondent analysis weights. Consumption and welfare reported in $1000s. Life expectancy and QALE reported in years. Retired is an indicator. Panel B presents cumulatively adjusted welfare estimates.

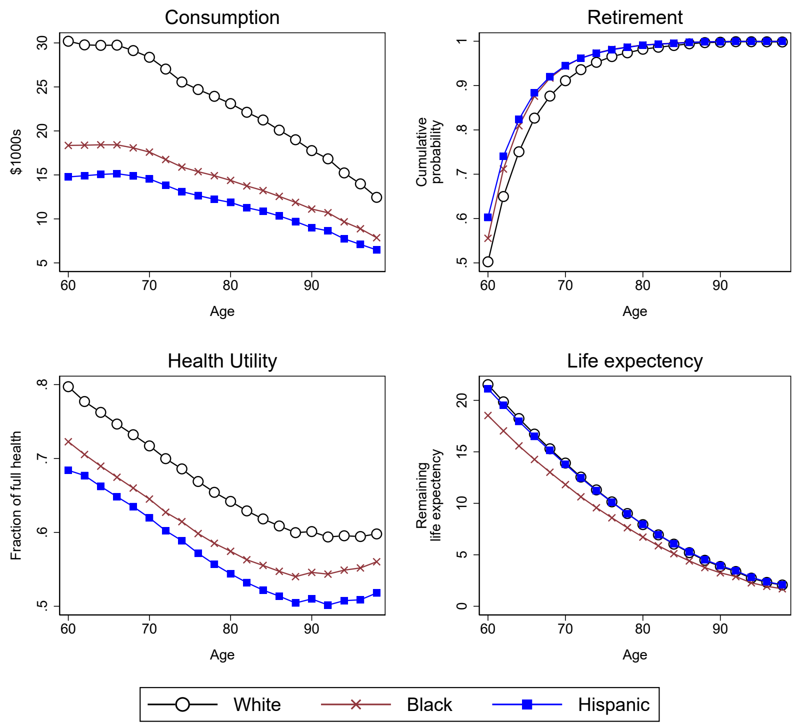
respondents retired earlier than White respondents overall, with the Black-White and Hispanic-White ratio of about 1.11 (or 110%) and 1.20 (or 120%), respectively. Moreover, White respondents reported higher life expectancy, QALE, and financial bequests. For example, life expectancy averaged about 18.6 years for Black respondents compared to 22 years for White respondents. On the other hand, the difference in life expectancy between Hispanic and White respondents is less than a year. However, the difference is bigger in QALE with about 14 years for Hispanic respondents compared to 17 years for White respondents. Furthermore, expected financial bequests for Black and Hispanic respondents is about 26% and 24% that of White respondents, respectively.

Ignoring differences in expected leisure, life expectancy, health, and financial bequests in Panel B still results in substantial overall welfare gap between races—the "consumption" Black-White welfare ratio is about 0.62 (or 62%) and the Hispanic-White welfare ratio is 0.51 (or 51%). Moreover, average expected consumption for Black and Hispanic respondents of the welfare distribution is only $14,752 and $12,120, respectively, compared to an average of $23,817 for White respondents. Adjusting welfare for lost leisure due to working past age sixty slightly decreases the welfare gap—increases the Black-White welfare ratio by an additional 1 pp and the Hispanic-White welfare ratio by 1 pp. This is because Black and Hispanic respondents expected to retire earlier than White respondents overall, but these differences had only small effects on our fully-adjusted measure of racial welfare gaps. In other words, adjusting welfare for later retirement lowers average welfare by $900 $(14,753 - 13,853)$ for Black and $694 $(12,120 - 11,426)$ for Hispanic respondents. This implies that Black and Hispanic respondents would be willing to give up to an average of $900 and $694 in expected annual consumption to retire at age sixty, respectively.

Health and life expectancy are important for overall welfare gaps. Further adjusting for life expectancy differences is more important for Black respondents, decreasing the estimated mean Black-White welfare ratio by 12 pp. In contrast, the welfare cost of living in poor health is more important for Hispanic respondents, decreasing the estimated Hispanic-White welfare ratio by 7 pp. The last row of Panel B shows adjustments for leaving financial bequest, yielding our fully-adjusted welfare measure. Smaller financial bequests are nearly as important to estimated welfare gaps as health and longevity. Adjusting for bequests lowers the Black-White welfare ratio an additional 10 pp and the Hispanic-White ratio an additional 8 pp.

Figure 3 plots average expected life-cycle profiles to gain a sense of the differences across racial groups. Mean racial gaps in consumption, retirement, and life expectancy are largest at age sixty and gradually decline as individuals age (although substantial consumption gaps extend into the nineties). In contrast, health gaps are also big at age sixty, but largely persist over re-

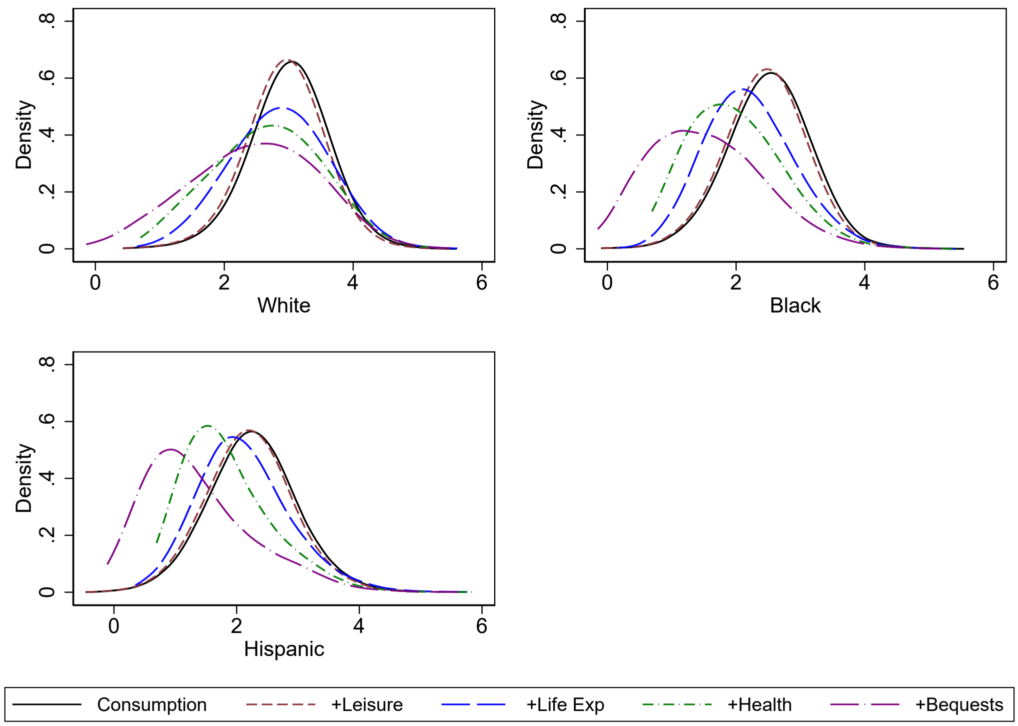
Figure 3: Average Life Cycle Profiles by Race



maining life. Overall, our welfare results for the EHRS cohort suggest that consumption, health, life expectancy, and bequests all play important roles in driving racial welfare inequality, while leisure through earlier retirement plays a comparatively minor role.

Figure 4 provides the cumulative change in the distribution of log welfare at age sixty across racial groups for closer examination. Adjusting for leisure, life expectancy, health, and bequests each have a larger negative impact on the welfare distribution of Black and Hispanic respondents than White respondents. It is also worth noting that inequality *within* the Black and Hispanic respondent population increases more than the White population with each adjustment (i.e., the left tail of the welfare distribution become fatter). This is broadly consistent with existing evidence on inequality. For example, the relative income disparity between the top and bottom 10 percent has been

Figure 4: Cumulative Change in Distribution of Log Welfare by Race



shown to be particularly acute for Black Americans. In 2016, the 90th percentile of Black households earned nearly ten times as much as the 10th percentile (Pew Research Center, 2018).

4.4 Decomposition

The increase in welfare inequality across racial groups in our estimates must either be driven by (1) differences in the distribution of age sixty initial conditions across races and/or (2) differences in the stochastic processes faced by each racial group after age sixty. We aim to answer the following key question in this analysis: how much of the racial welfare gaps is explained by initial (age sixty) conditions versus differences in outcome dynamics after age sixty? Towards this goal, we conduct a number of experiments to estimate the impact of initial differences at age sixty as well as the differential evolution of outcomes across racial groups after age sixty. In all our experiments, we eliminate disparities by assigning initial conditions or late-life transitions of White participants to Black and Hispanic participants. Our main decomposition results are presented in Table 3. We report the Black-White and Hispanic-White ratios for quality-adjusted life year (QALE), expected lifetime consumption (ELC), and our fully-adjusted welfare measure at age sixty.

Table 3: Decomposition

	QALE ratio		ELC ratio		Welfare ratio	
Experiment	Black-White	Hispanic-White	Black-White	Hispanic-White	Black-White	Hispanic-White
Baseline	0.809	0.846	0.559	0.491	0.399	0.355
Transitions	0.847	0.806	0.578	0.459	0.432	0.337
Initial conditions	0.953	1.039	0.965	1.053	0.901	1.075

Notes: Estimates use base year respondent analysis weights.

In our first set of experiments, we assign the transition probabilities of White participants after age sixty to Black and Hispanic groups to understand how the evolution of outcomes after sixty influences gaps in QALE, ELC, and welfare. As shown in the second row of Table 3, differences in the evolution of outcomes explain very little of the racial welfare gaps. For example, assigning White transition probabilities to Black participants only increases the QALE ratio by 3.8 pp, ELC ratio by 1.9 pp, and fully-adjusted welfare by 3.3 pp. Outcomes for Hispanic respondents actually become slightly *worse* when receiving White transition probabilities, with the QALE ratio declining by 4 pp, ELC ratio by 3.2 pp, and the fully-adjusted welfare ratio by 1.8 pp.

We next turn to the role of age sixty differences in explaining the estimated racial welfare gaps. In our previous experiment we only changed the evolution of outcomes after age sixty while keeping the initial distribution of outcomes the same for each racial group. As shown in the last row of Table 3, if we instead assign the initial conditions of White respondents to all groups, the estimated

Black-White and Hispanic-White ratios in QALE, ELC, and the fully-adjusted welfare measure increase significantly. For example, equating initial conditions increases the Black-White welfare ratio by 50 pp and the Hispanic-White ratio by 70 pp. Our decomposition exercises show that a majority of the estimated welfare gaps are determined by age sixty initial conditions as opposed to racial differences in dynamic processes after age sixty.

4.5 Morbidity Counterfactuals

This section tries to further understand how morbidities influence outcomes and welfare across racial groups. Table 4 shows the gain in QALE and ELC, the loss in bequests, and Black-White/Hispanic-White welfare ratios with the exogenous elimination of all hypertension or diabetes after age sixty. We chose hypertension and diabetes because our model estimates (see Figure 2) show that, conditional on other health measures and exogenous characteristics, racial gaps are largest for these two morbidities. Moreover, hypertension and diabetes are established risk factors for downstream morbidities such as stroke, ischemic heart disease, renal dysfunction, kidney failure, and other medical problems (e.g., Lewington, 2002; Rapsomaniki et al., 2014; Huang et al., 2014; Kokubo and Iwashima, 2015; Raghavan et al., 2019).

As shown in Table 4, eliminating hypertension saw Black and Hispanic respondents gain slightly larger QALE than White respondents at age sixty. Specifically, Black (Hispanic) respondents gain about 1.5 (1.4) years compared to 1.2 years for White respondents. However, White respondents saw a gain of $32,152 in ELC compared to $26,594 for Black and $19,484 for Hispanic respondents. This is driven by the larger annual consumption among White respondents. On the flip side, this larger gain in lifetime consumption is partially offset by larger declines in bequests for White respondents—$13,947 compared to only $5,390 for Black and $5,116 for Hispanic respondents. Eliminating late-life diabetes showed smaller effects but similar overall patterns as hypertension—smaller health gains but larger consumption gains and bequest losses for White respondents. The exception is that Black respondents gained more lifetime consumption than White respondents in the diabetes experiment.

Table 4: Eliminating Late-life Hypertension and Diabetes by Race

	Hypertension			Diabetes		
Outcomes and Welfare	White	Black	Hispanic	White	Black	Hispanic
QALE gain	1.241	1.462	1.409	0.699	1.040	1.251
ELC gain	32.152	26.594	19.484	17.926	19.788	17.016
Bequest loss	13.947	5.390	5.116	5.499	2.385	3.460
Welfare ratio	–	0.384	0.341	–	0.391	0.356
Baseline ratio	–	0.379	0.346	–	0.379	0.346

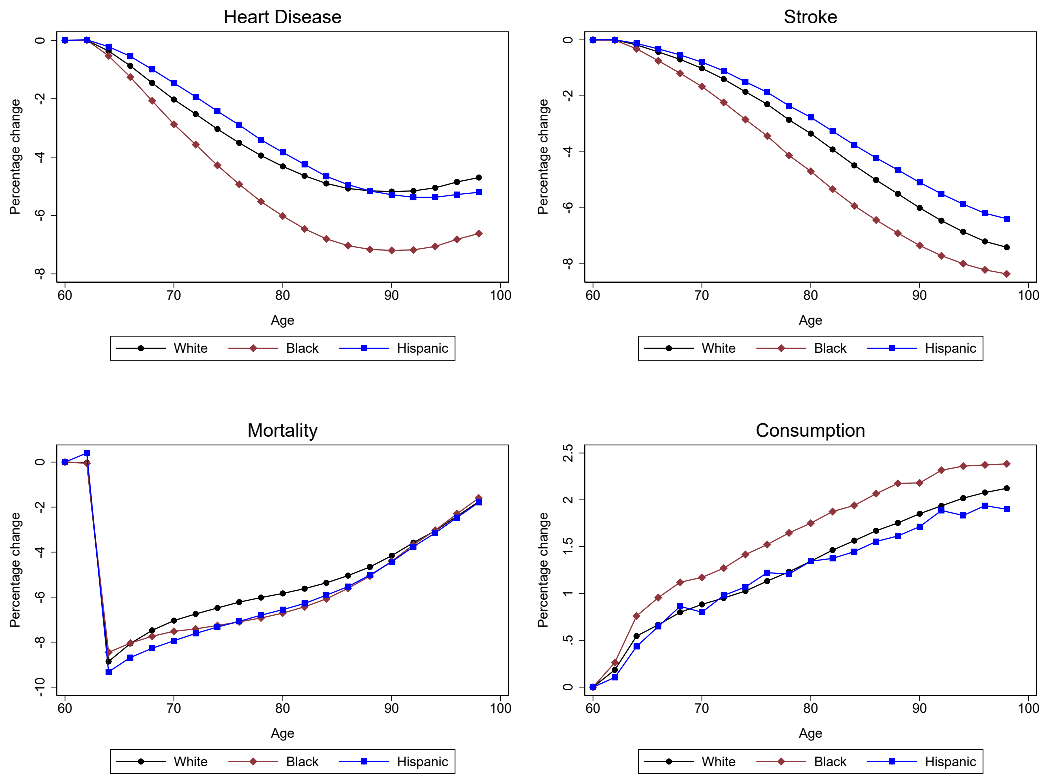
Notes: Estimates use base year respondent analysis weights. Consumption and bequests reported in \$1000s. QALE reported in years. Welfare ratio is measured as Black-White and Hispanic-White.

Unlike with hypertension, we also see higher QALE gains (and bequest losses) for Hispanic respondents than Black respondents.

Comparing the counterfactual welfare ratio to the baseline suggests that eliminating late-life hypertension and diabetes only marginally closes overall welfare gaps. Specifically, eliminating hypertension after age sixty increases the Black-White welfare ratio by 0.005 pp and actually *lowers* the Hispanic-White ratio by 0.005 pp. Eliminating diabetes saw slightly larger improvements, with the Black-White welfare ratio increasing by 0.012 pp and Hispanic-White welfare ratio by 0.01 pp.

In order to gain a better sense of how morbidities influence the dynamics of other outcomes in the system across racial groups, Figures 5 and 6 plot the average percentage change in several expected outcomes with the exogenous elimination of hypertension and diabetes after age sixty. Eliminating hypertension after age sixty lowers the average probability of developing heart disease and stroke for all races, though changes were strongest for Black respondents. For example, Black respondents saw a decreased probability of heart disease of about 6% by age eighty compared to about 4% for White and Hispanic respondents. Likewise, the probability of stroke by age eighty decreased by about 5% for Black respondents compared to 3% for White and Hispanic respondents. Interestingly, despite the smaller morbidity gains, we see similar mortality gains for Hispanic compared to Black respondents. Although Black respondents also see the largest gains in consumption.

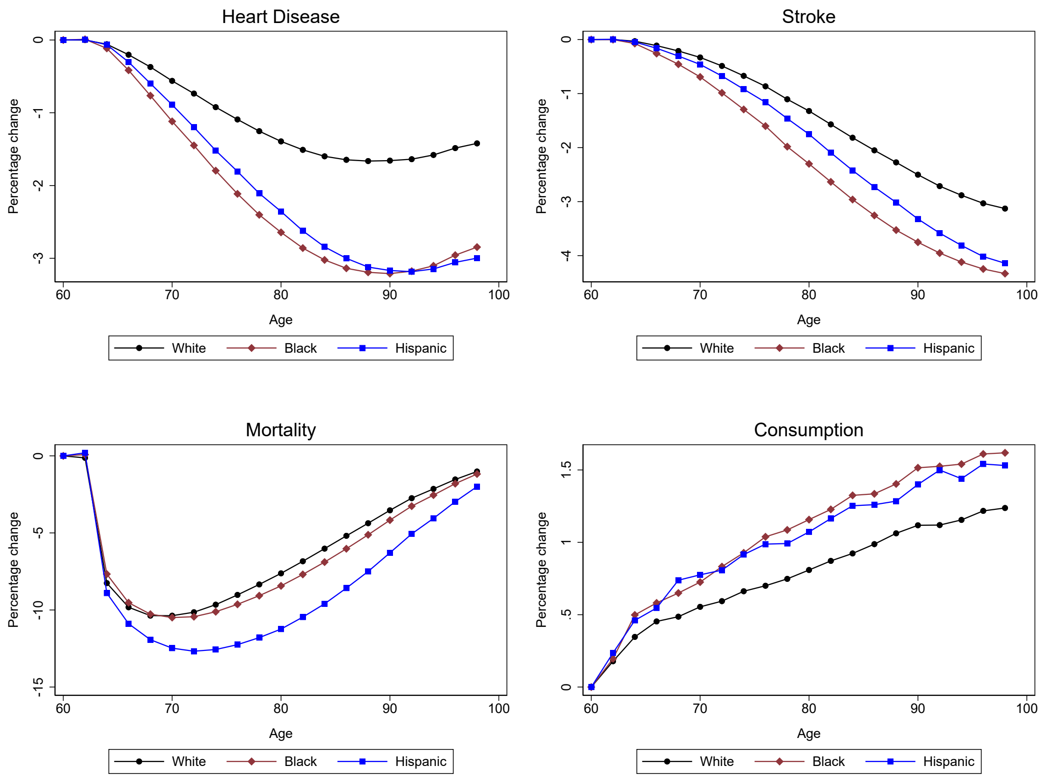
Figure 5: Impulse Response to Elimination of Hypertension after Age 60



Notes: Results plot percentage difference in expected outcomes with the exogenous elimination of hypertension after age sixty relative to baseline. Sample includes all individuals in the simulation sample from the EHRS cohort. Expected outcomes are conditional on survival.

Compared to eliminating hypertension, general patterns are similar in the diabetes experiment. The main difference is that effects are relatively stronger for Hispanic respondents. For example, Hispanic and Black respondents see similar improvements in heart disease incidence and consumption when eliminating late-life diabetes. However, Hispanic respondents clearly have the largest mortality gains. These patterns are consistent with the very strong association with increased diabetes risk among Hispanic respondents from our model estimates shown in Figure 2.

Figure 6: Impulse Response to Elimination of Diabetes after Age 60



Notes: Results plot percentage difference in expected outcomes with the exogenous elimination of diabetes after age sixty relative to baseline. Sample includes all individuals in the simulation sample from the EHRS cohort. Expected outcomes are conditional on survival.

4.6 Robustness

We estimated our main results under a variety of alternate modeling assumptions from our benchmark to gauge the sensitivity of our findings. These included using a race specific forecasting model, a higher reference life expectancy and reference bequests, alternate preference parameter values, and alternate health utility weights. Summary results are presented in Table 5.

While welfare levels are somewhat sensitive to robustness specifications, the Black-White ratio remains in the range of 0.36-0.45 and the Hispanic-White ratio in the range of 0.33-0.41.

Table 5: Robustness Results: Mean Welfare by Race

Measure	**Mean**				
	White	Black	Hispanic	Black-White-Ratio	Hispanic-White-Ratio
Benchmark	18.322	6.946	6.340	0.379	0.346
Race specific forecast	18.392	6.741	6.691	0.367	0.364
Reference life expectancy	11.655	5.221	4.744	0.448	0.407
$\bar{u}$ = -log(1.5)	18.448	6.714	6.211	0.364	0.337
β = 0.90	17.675	7.713	6.756	0.436	0.382
ϵ = 0.5	19.622	7.364	6.728	0.375	0.343
ϵ = 2	16.510	6.357	5.800	0.385	0.351
θ = 16	16.999	6.517	5.944	0.383	0.350
Reference bequests	17.644	6.690	6.106	0.379	0.346
Φ_1 = -5	19.111	8.036	7.284	0.421	0.381
Φ_2 = 6	18.473	7.095	6.444	0.384	0.349
σ = 2	18.448	7.100	6.415	0.385	0.348
Health utility weights	18.602	7.148	6.589	0.384	0.354

Notes: Estimates use base year respondent analysis weights. Welfare reported in \$1000s.

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

$$\phi\left(h\right)\left[\frac{c^{1-\gamma}}{1-\gamma}\left(1-\left(1-\gamma\right)\frac{\theta\epsilon}{1+\epsilon}\left(1-l\right)^{\frac{1+\epsilon}{\epsilon}}\right)^{\gamma}-\frac{\bar{u}^{1-\gamma}}{1-\gamma}\right] \tag{13}$$

which reduces to our benchmark case with $\gamma=1$. With $\gamma>1$ there is more curvature over consumption. Several problems arise with higher curvature over consumption. First, it is no longer possible to calculate welfare for those at the very top of the health distribution as no amount of consumption increase would provide the same expected life-time utility as the reference life expectancy. Thus we report median instead of mean welfare in Table 6. Second, as discussed by Murphy and Topel (2006), another problem that arises with higher curvature in this framework is that as γ rises, the implied value of life grows rapidly. In order to gain a sense of this issue, the first column in Table 6 shows the median value of life per QALY with higher curvatures. With $\gamma = 2$, the median value of life is high but not completely implausible at $\$178,000$ per QALY. The estimated median Black-White and Hispanic-White welfare ratios climb to 0.39 and 0.36 with the higher curvature. When $\gamma = 3$, the value of life reaches about $\$557,000$ per QALY and the welfare ratios reach 0.58 and 0.54. Only three out 23 value of life studies surveyed by Ryen and Svensson (2015) estimated a mean value of life over $\$150,000$. 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.

Table 6: Robustness Results for Higher Curvature: Median Welfare by Race

Measure			**Mean**			
	VOL	White	Black	Hispanic	Black-White-Ratio	Hispanic-White-Ratio
γ = 1.0	59.805	12.231	4.277	3.223	0.350	0.264
γ = 1.5	104.601	7.416	2.421	1.969	0.326	0.265
γ = 2.0	178.162	3.729	1.472	1.336	0.395	0.358
γ = 3.0	557.168	1.756	1.012	0.951	0.576	0.541

Notes: Estimates use base year respondent analysis weights. Welfare reported in $1000s.

5. Conclusion

We propose and estimate an individual measure of racial welfare incorporating heterogeneity and uncertainty in future consumption, leisure, health, wealth and mortality at age sixty. Our measure broadly indicates that racial inequality is larger than suggested by other welfare metrics such as income or consumption. We also find health, mortality, and wealth gaps are important in explaining the level of racial welfare inequality among the older Americans in our sample, with leisure playing a comparatively minor role.

Our decomposition exercises show that a majority of the estimated welfare gaps are determined by age sixty initial conditions as opposed to racial differences in dynamic processes after age sixty. Our morbidity counterfactuals further suggest that eliminating common heath risk factors such as hypertension or diabetes in late-life only marginally closes overall welfare gaps. These simulations suggest that policies aimed at closing racial gaps in late-life may be more successful and efficient if targeted earlier in the life-cycle. In other words, outside of direct wealth transfers, it may largely be too late to target such interventions directly at older populations.

Our approach is not without limitations. We do not explicitly account for morbidity spillover effects such as the cost of caregiver time and the numerous costs associated with the loss of a spouse. Likewise, we abstract from other potentially important inputs into late-life welfare such as social interactions and end-of-life care. We assume institutions and relevant policies remain fixed moving forward and past trends in late-life health, retirement, and consumption continue into the future. For example, significant anticipated changes to Social Security or Medicare programs or exponential advances in medicine could alter the distribution of our welfare measure. Nonetheless, our framework provides important insights into the sources and scope of racial welfare gaps.

References

Adams, K., S. Leibbrandt, and H. Moon (2011). A Critical Review of the Literature on Social and Leisure Activity and Wellbeing in Later Life. *Ageing & Society 31*(4), 683–712.

Aliprantis, D., D. Carroll, and E. Young (2019). The Dynamics of the Racial Wealth Gap. *Federal Reserve Bank of Cleveland Working Paper No. 19-18*.

Ashman, H. and S. Neumuller (2020). Can Income Differences Explain the Racial Wealth Gap: A Quantitative Analysis. *Review of Economic Dynamics 35*, 220–239.

Benhabib, J., A. Bisin, and M. Luo (2017, May). Earnings Inequality and Other Determinants of Wealth Inequality. *American Economic Review 107*(5), 593–97.

Bhutta, N., A. C. Chang, L. Dettling, and J. Hsu (2020). Disparities in Wealth by Race and Ethnicity in the 2019 Survey of Consumer Finances. Technical report, FEDS Notes. Washington: Board of Governors of the Federal Reserve System.

Conley, D. (2000). The Racial Wealth Gap: Origins and Implications for Philanthropy in the African American Community. *Nonprofit and Voluntary Sector Quarterly 29*(4), 530–540.

Darity Jr, W. A. and S. L. Myers Jr (1998). *Persistent Disparity: Race and Economic Inequality in the United States since 1945*. Edward Elgar Publishing.

De Nardi, M. (2004). Wealth Inequality and Intergenerational Links. *The Review of Economic Studies 71*(3), 743–768.

Deaton, A. S. and C. H. Paxson (1998). Aging and Inequality in Income and Health. *American Economic Review 88*(2), 248–253.

Feeny, D., W. Furlong, G. W. Torrance, C. H. Goldsmith, Z. Zhu, S. DePauw, M. Denton, and M. Boyle (2002). Multiattribute and Single-Attribute Utility

Functions for the Health Utilities Index Mark 3 System. *Medical Care 40*(2), 113–128.

Fleurbaey, M. (2005). Health, Wealth, and Fairness. *Journal of Public Economic Theory 7*(2), 253–284.

Fleurbaey, M. (2009). Beyond GDP: The Quest for a Measure of Social Welfare. *Journal of Economic Literature 47*(4), 1029–75.

Fleurbaey, M. and G. Gaulier (2009). International Comparisons of Living Standards by Equivalent Incomes. *Scandinavian Journal of Economics 111*(3), 597–624.

Fleurbaey, M., S. Luchini, C. Muller, and E. Schokkaert (2013). Equivalent Income and Fair Evaluation of Health Care. *Health Economics 22*(6), 711–729.

Furlong, W., D. Feeny, G. Torrance, C. Goldsmith, S. DePauw, Z. Zhu, M. Denton, M. Boyle, 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.

Goetz, S. J. (1991). Single-and Multi-Dimensional Theil Measures of Inequality: An Application to Farm Households. Technical report.

Han, J. and I. Patterson (2007). An Analysis of the Influence that Leisure Experiences Have on a Person's Mood State, Health and Wellbeing. *Annals of Leisure Research 10*(3-4), 328–351.

Honaker, J. and G. King (2010). What to Do about Missing Values in Time-Series Cross-Section Data. *American Journal of Political Science 54*(2), 561–581.

Horsman, J., W. Furlong, D. Feeny, and G. Torrance (2003). The Health Utilities Index (HUI®): Concepts, Measurement Properties and Applications. *Health and Quality of Life Outcomes 1*(1), 1–13.

Hribernik, J. and A. J. Mussap (2010). Research Note: Leisure Satisfaction and Subjective Wellbeing. *Annals of Leisure Research 13*(4), 701–708.

Huang, Y., X. Cai, J. Zhang, W. Mai, S. Wang, Y. Hu, H. Ren, and D. Xu (2014). Prehypertension and Incidence of ESRD: A Systematic Review and Meta-Analysis. *American Journal of Kidney Diseases 63*(1), 76–83.

Idler, E. L. and Y. Benyamini (1997). Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. *Journal of Health and Social Behavior*, 21–37.

Jones, C. I. and P. J. Klenow (2016). Beyond GDP? Welfare Across Countries and Time. *American Economic Review 106*(9), 2426–57.

Kaplan, R. M. and J. W. Bush (1982). Health-Related Quality of Life Measurement for Evaluation Research and Policy Analysis. *Health Psychology 1*(1), 61.

Kokubo, Y. and Y. Iwashima (2015). Higher Blood Pressure as a Risk Factor for Diseases other than Stroke and Ischemic Heart Disease. *Hypertension 66*(2), 254–259.

Lepinteur, A. (2019). The Shorter Workweek and Worker Wellbeing: Evidence from Portugal and France. *Labour Economics 58*, 204–220.

Lewington, S. (2002). Prospective Studies Collaboration. Age-Specific Relevance of Usual Blood Pressure to Vascular Mortality: A Meta-Analysis of Individual Data for One Million Adults in 61 Prospective Studies. *Lancet 360*, 1903–1913.

Lynch, S. M. (2008). Race, Socioeconomic Status, and Health in Life-Course Perspective: Introduction to the Special Issue. *Research on Aging 30*(2), 127–136.

Maasoumi, E. (1986). The Measurement and Decomposition of Multidimensional Inequality. *Econometrica*, 991–997.

Maasoumi, E. and G. Nickelsburg (1988). Multivariate Measures of Well-Being and an Analysis of Inequality in the Michigan Data. *Journal of Business & Economic Statistics 6*(3), 327–334.

Maasoumi, E. and G. Nickesburg (1983). Measuring Joint Inequality in Incomes, Wealth and Schooling and a Decomposition. *Department of Economics Working Paper, 83-4, Indiana University, Bloomington*.

Manduca, R. (2018). Income Inequality and the Persistence of Racial Economic Disparities. *Sociological Science 5*, 182–205.

Manton, K. G., P. C. H. . J. K. W. (1987). Health Differentials between Blacks and Whites: Recent Trends in Mortality and Morbidity. *The Milbank Quarterly 65*, 129–199.

Margo, R. A. (2016). Obama, Katrina, and the Persistence of Racial Inequality. *The Journal of Economic History 76*(2), 301–341.

Martin, L. G. and B. J. Soldo (1997). *Racial and Ethnic Differences in the Health of Older Americans*. National Academies Press.

Miller, R. and N. Bairoliya (2021). Health, Longevity, and Welfare Inequality of Older Americans. *Forthcoming in The Review of Economics and Statistics*.

Minkler, M., E. Fuller-Thomson, and J. M. Guralnik (2006). Gradient of Disability across the Socioeconomic Spectrum in the United States. *New England Journal of Medicine 355*(7), 695–703.

Murphy, K. M. and R. H. Topel (2006). The Value of Health and Longevity. *Journal of Political Economy 114*(5), 871–904.

Ozawa, M. N. and H.-y. Tseng (2000). Differences in Net Worth between Elderly Black People and Elderly White People. *Social Work Research 24*(2), 96–108.

Pager, D. and H. Shepherd (2008). The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets. *Annual Review of Sociology 34*, 181–209.

Patton, G. C., S. M. Sawyer, J. S. Santelli, D. A. Ross, R. Afifi, N. B. Allen, M. Arora, P. Azzopardi, W. Baldwin, C. Bonell, et al. (2016). Our Future: A Lancet Commission on Adolescent Health and Wellbeing. *The Lancet 387*(10036), 2423–2478.

Pew Research Center (2018). Income Inequality in the U.S. is Rising Most Rapidly Among Asians. https://www.pewresearch.org/social-trends/wp-content/uploads/sites/3/2018/07/Pew_Research_Center_Inequality-Report_FINAL.pdf.

Pollack, C. E., S. Chideya, C. Cubbin, B. Williams, M. Dekker, and P. Braveman (2007). Should Health Studies Measure Wealth?: A Systematic Review. *American Journal of Preventive Medicine 33*(3), 250–264.

Raghavan, S., J. L. Vassy, Y.-L. Ho, R. J. Song, D. R. Gagnon, K. Cho, P. W. Wilson, and L. S. Phillips (2019). Diabetes Mellitus–Related All-Cause and Cardiovascular Mortality in a National Cohort of Adults. *Journal of the American Heart Association 8*(4), e011295.

Rapsomaniki, E., A. Timmis, J. George, M. Pujades-Rodriguez, A. D. Shah, S. Denaxas, I. R. White, M. J. Caulfield, J. E. Deanfield, L. Smeeth, et al. (2014). Blood Pressure and Incidence of Twelve Cardiovascular Diseases: Lifetime Risks, Healthy Life-Years Lost, and Age-Specific Associations in 1· 25 Million People. *The Lancet 383*(9932), 1899–1911.

Rohde, N. and R. Guest (2013). Multidimensional Racial Inequality in the United States. *Social Indicators Research 114*(2), 591–605.

Ryen, L. and M. Svensson (2015). The Willingness to Pay for a Quality Adjusted Life Year: A Review of the Empirical Literature. *Health Economics 24*(10), 1289–1301.

Samson, A.-L., E. Schokkaert, C. Thébaut, B. Dormont, M. Fleurbaey, S. Luchini, and C. Van de Voorde (2018). Fairness in Cost-Benefit Analysis: A Methodology for Health Technology Assessment. *Health Economics 27*(1), 102–114.

Schokkaert, E., C. Van de Voorde, B. Dormont, M. Fleurbaey, S. Luchini, A.-L. Samson, and C. Thébaut (2013). Equity in Health and Equivalent Incomes. In *Health and Inequality*, Volume 21, pp. 131–156. Emerald Group Publishing Limited.

Shapiro, T. and J. Kenty-Drane (2005). The Racial Wealth Gap. In *African Americans in the US Economy*, Volume 175, pp. 177. Rowman & Littlefield Lanham.

Shea, D. G., T. Miles, and M. Hayward (1996). The Health-Wealth Connection: Racial Differences. *The Gerontologist 36*(3), 342–349.

Smith, G. D. and M. Egger (1993). Socioeconomic Differentials in Wealth and Health. *British Medical Journal 307*(6912), 1085.

Smith, J. P. et al. (1997). Wealth Inequality Among Older Americans. *Journals of Gerontology Series B 52*, 74–81.

Steptoe, A., A. Deaton, and A. A. Stone (2015). Subjective Wellbeing, Health, and Ageing. *The Lancet 385*(9968), 640–648.

Strife, S. and L. Downey (2009). Childhood Development and Access to Nature: A New Direction for Environmental Inequality Research. *Organization & Environment 22*(1), 99–122.

Vespa, J., D. M. Armstrong, and L. Medina (2018). *Demographic Turning Points for the United States: Population Projections for 2020 to 2060*. US Department of Commerce, Economics and Statistics Administration, US Census Bureau.

Williams, D., C. Wilson, R. Binstock, and L. George (2001). Race, Ethnicity, and Aging. In *Handbook of Aging and Social Sciences, 5th Edition*, pp. 160–178. San Diego, CA: Academic Press.