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

Sayorn Chin*[a] and Ray Miller[b]

[a]Department of Economics, Lafayette College, Easton, PA
[b]Department of Economics, Colorado State University, Fort Collins, CO

May 7, 2023

Abstract

We estimate racial and ethnic 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 use 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 (1992-2016) for our analysis. Our measure broadly indicates that racial and ethnic 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 and ethnic 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 and ethnic 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 and ethnic 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, welfare, health, wealth, aging, consumption, mortality

---

*Corresponding author: chins@lafayette.edu. We wish to acknowledge helpful feedback from Sammy Zahran, David Mushinski, and Jude Bayham. All errors are our own.

1 Introduction

Racial and ethnic 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 and ethnic 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 and ethnic inequities. For instance, studies have found that income inequality across racial and ethnic groups is usually lower than wealth inequality, implying some underestimation of the broader racial and ethnic 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; Miller and Bairoliya, 2022; Miller et al., 2022). 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 and ethnic lines. While each such variable individually contributes to the gaps in racial and ethnic well-being, it remains unclear if the adoption of such narrowly defined metrics can adequately capture the true welfare inequality between racial and ethnic 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 and ethnic 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 and ethnic disparity measures based on cross-sectional income or consumption might underestimate the aggregate racial and ethnic 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 (2022). 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 (1992-2016) for our analysis. We then use the estimated system to simulate potential outcome paths by race/ethnicity 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/ethnicities, 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 and ethnic 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 and ethnically 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 and ethnic differences in dynamic processes after age sixty. This suggests that policies aimed at closing racial and ethnic 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 and ethnic 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 and ethnic 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 and ethnic 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 and ethnic 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 and ethnic welfare gaps at older ages.

The rest of this paper is divided into four sections. First, we will examine the data and statistical methods used in our study in Section 2. Second, we will discuss the construction of our welfare measure in Section 3. Third, in Section 4, we will present the results of our welfare analysis. Finally, we will conclude and discuss policy implications while also addressing limitations in Section 5.

2 Data and Methods

2.1 Data

We utilized data from the Health and Retirement Study (HRS), which is a national biennial longitudinal survey tracking individuals aged 50 and above in the United States across multiple cohorts. The HRS data includes seven birth cohorts, namely 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), the Children of Depression (CODA) cohort (born between 1924 and 1930), the War Baby (WB) cohort (born between 1942 and 1947), and the Early, Mid, and Late Baby Boomer cohorts (born after 1947). Our main data source was the publicly available 2016 RAND HRS Longitudinal File which includes data from 1992 to 2016. The file provided us with cleaned data on various individual characteristics such as race/ethnicity, health, mortality, economic outcomes, age, education, gender, birth cohort, region, and occupation. In the following section, we provide more detailed information on the variables employed in our analysis.

2.1.1 Race/Ethnicity Variables

In the HRS survey, respondents were asked two questions about their race/ethnicity: "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?" For our analysis, we categorized race/ethnicity into three groups: White, non-Hispanic; Black, non-Hispanic; and Hispanic, based on their answers. We excluded American Indian or Alaskan Native, Asian or Pacific Islander, and Unknown categories from the analysis, as they are not representative.

2.1.2 Health Outcomes

Importantly for older populations, our model incorporates data on comorbidities. Specifically, we include binary indicators for doctor's diagnosis of eight specific health problems as well as an indicator for ever reported difficulties with activities of daily living (ADLs). ADLs include activities such as bathing, getting dressed, walking across the room, and toileting. The health problems included 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. These health metrics are arguably more objective measures of health. However, self-rated health outcomes, where individuals rank their health on a five-point scale from poor (one) to excellent (five), have also been shown to be a good predictor of mortality even after controlling for other health conditions, health behavior, and socioeconomic characteristics (Idler and Benyamini, 1997). Therefore, we include self-rated health status in our model to test if people have significant private information about their health beyond 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 data provided by the Consumption Activities Mail Survey (CAMS), which was sent to a sub-sample of HRS respondents on off years of the core survey. The 2017 RAND CAMS data file provides a constructed estimate of total household consumption derived from household spending data on durables, non-durables, transportation, and housing collected from 2001-2015. We subtracted out-of-pocket health spending from total household consumption and then 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 (2022) 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 the Online Appendix for more 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).

We adapt the panel vector autoregressive (VAR) model of Miller and Bairoliya (2022) to estimate the joint evolution of consumption, leisure, health, mortality, and wealth (valued as bequests at death) across different racial/ethnic groups in late-life. Our proposed model enables us to: (1) accurately measure the racial/ethnic disparities in welfare within a given population; and (2) explore the extent to which these disparities could potentially be reduced through various counterfactual scenarios. The dynamics of the life-cycle are represented as a statistical process and estimated directly from the data. While explicitly modeling the maximization of lifetime utility would enable better policy analysis, it involves solving a complex intertemporal structural model that considers endogenous savings, labor supply, and multiple morbidity and health outcomes. Given that the primary goal of this paper is to develop a welfare measure that accurately reflects population well-being, we believe that a data driven statistical approach is more appropriate in this context.

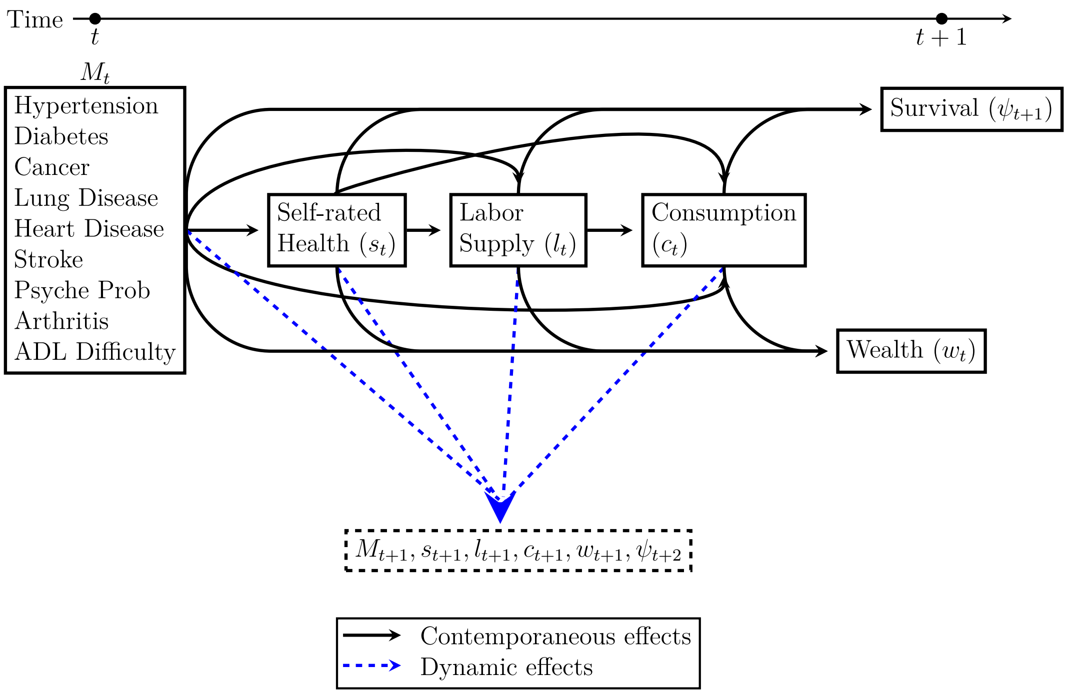


Figure 1. Simulation Model With One Period Lag

The core structure of the simulation 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, which affects their labor supply (i.e., their decision to retire) and, in turn, impacts consumption, wealth, and the likelihood of survival to the next time period. Note that the model allows both direct and indirect contemporaneous effects. For example, a stroke may influence retirement directly or through a change in self-rated health. Finally, general lagged effects are also included in the model (e.g., hypertension this period can impact the chance of heart disease next period). An important aspect of including lagged effects is that it allows for more recent diagnoses of a morbidity to have a different impact on health and economic changes than long-standing diagnoses.

2.1.4 Panel VAR Representation

While we allow for higher order lags in estimation, the following VAR(1) demonstrates the relevant structure of the model. In this model, Y_{it} represents a vector of outcomes for an individual i at time t. This vector includes log consumption c, retirement indicator r, self-rated health s, cube root of wealth w, and $n = 9$ morbidity states which are given by the $n \times 1$ vector M. We model each morbidity as an absorbing state to be consistent with the HRS data (e.g., ever diagnosed with hypertension). For simplicity, we also model retirement as an absorbing state (e.g., once retired always retired). We further include a $k \times 1$ vector of fixed individual characteristics X_{it} as exogenous predictors in our model.

Conditional on survival, the outcomes evolve according to the structural VAR(1) model:

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

where ϵ is a vector of independent and identically distributed (iid) shocks with zero mean, and the diagonal elements of matrix A are scaled to one. All parameters in the model are identical across individuals and time (e.g., $A_{it} = A$ for all i and t).

The model is estimated in five "blocks" of outcomes: morbidities, self-rated health, retirement, consumption, and wealth blocks. Setting aside the exogenous vector X_{it} for exposition, the VAR(1) model can be written in the following block matrix form:

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

where $n \times n$ matrix A_{11} has diagonal terms scaled to one. As illustrated in Figure 1, we assume the contemporaneous causal pathway runs from morbidities to self-rated health to retirement to consumption to wealth. This assumption is represented in the VAR(1) model by setting $A_{12} = A_{13} = A_{14} = A_{15} = 0$ 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. Note that health outcomes and retirement are allowed to affect all future outcomes through general lagged effects. We further allow lagged consumption to impact future wealth, but consumption and wealth are otherwise assumed not to have lagged effects.[1] By applying such block triangulation of the system, we eliminate simultaneity across blocks and allow for block-by-block estimation.

Exogenous characteristics X_{it} include a linear trend for calendar year and dummies for age, education, gender, census division, census occupation code, birth cohort and a post-2008 indicator to account for the great recession. We also include a time invariant individual fixed effect in the consumption equation (π^c) and in the wealth equation (π^w).

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

The unobserved fixed effect helps maintain the appropriate variance in consumption and wealth across time by acting as a person specific drift in the autoregressive process. The entry of exogenous characteristics in the VAR(1) can be explicitly written as:

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

Here we have excluded time invariant regressors from the consumption and wealth equations due to colinearity with the fixed effects. Time invariant socioeconomic characteristics are used instead of fixed effects in the health and retirement equations because absorbing states and ordinal models raise challenges in estimating dynamic panel models with fixed effects. Moreover, the model does well in replicating the dynamics of health and retirement even without unobserved fixed effects (see the Online Appendix for more details). Finally, note that we normalize c_{410} and c_{511} to one to allow identification of the unobserved fixed effects in the consumption and wealth blocks.

2.1.5 Morbidities

The system's block triangulation does not allow for the direct identification of the structural parameters in the morbidity block since there are nine separate outcomes. Therefore, the morbidity block is estimated as a reduced form VAR. To obtain the reduced form system, the structural system block is pre-multiplied by the inverse of matrix A_{11} as follows:

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

Denoting $-A_{11}^{-1} B_{1j} = \hat{B}_j$, $-A_{11}^{-1} [C_{11}, \ldots, C_{19}] = \hat{C}$ 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 have a direct contemporaneous effect on each other. However, there could be a potential 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. We assume that contemporaneous morbidity shocks follow a standard multivariate normal distribution with an $n \times n$ covariance matrix given by $\sum$.

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

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

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

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

It is important to note that the determination of each latent morbidity variable relies on lagged values of the other observed self-rated health and morbidity states. The morbidity block of equations takes the form of a multivariate probit model.

2.1.6 Self-Rated Health

Self-rated health is evaluated using a five-point scale. Therefore, similar to morbidity outcomes, predicting this measure is not a linear VAR process. We assume a continuous latent variable, denoted as s^*, underlies the observed self-rated health state. Accordingly, the relevant equation given in system (1) can be explicitly written as follows:

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

The observed health state is defined by the following equation:

$$s_{it} = \delta \quad if \; \kappa_{\delta-1} < s^*_{it} < \kappa_\delta \quad for \; \delta = 1, \ldots, 5.$$

Here, δ = 1 represents the poorest health state (poor), while $\delta = 5$ represents the best health state (excellent). To account for the persistence of general health shocks over the life-course, we assume that the latent self-rated health depends on the lagged value of the observed self-rated health category. We also assume that ϵ_2 is an iid shock with a standard normal distribution. Consequently, 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.

2.1.7 Retirement

We assume that retirement is a binary outcome, and that there is a continuous latent variable, denoted by r^*, which underlies the observed outcome. Specifically, we define r_{it} as follows:

$$\begin{aligned} r_{it} &= 0 \quad if \quad r^*_{it} \leq 0 \\ r_{it} &= 1 \quad if \quad r^*_{it} > 0. \end{aligned}$$

Assuming that the individual worked during the previous period (and setting $b_{33} = 0$), the retirement model, as defined in system (1), can be expressed as follows:

$$r^*_{it} \;=\; A_{31}M_{it} + a_{32}s_{it} + B_{31}M_{it-1} + b_{32}s_{it-1} + [c_{31}, \ldots, c_{39}]\, X_{it} + \epsilon_{3,it}. \tag{4}$$

Here, retirement is influenced by both current and lagged values of self-rated health and specific morbidities, as well as exogenous individual characteristics. We assume that ϵ_3 is an iid shock with a standard normal distribution, which implies that the retirement model has a standard probit structure.

2.1.8 Consumption and Wealth

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

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

Similarly, the equation for wealth can be given as:

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

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

2.1.9 Mortality

The last process to model is the survival from one life period to the next. Mortality probabilities are estimated separately from the VAR system mentioned earlier, as all other outcomes described are dependent on survival. Given that an individual is alive at time $t-1$, the survival to the next life period is modeled using the following equation:

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

Here, $\psi = 1$ indicates survival, X is the vector of previously defined observed individual characteristics, and u_{it} is an iid random shock with a standard normal distribution. The specification allows K lags of morbidity states, self-rated health, and retirement to affect the probability of survival.

2.1.10 Simulations

Our empirical analysis involves three steps, which utilize our forecasting model. Firstly, we estimate the parameters of the model using data from the HRS. The data includes all individuals aged fifty and older from all available waves of the HRS from 1992-2016, amounting to 40,708 unique individuals and 238,091 total individual-year observations. Additional details on the model estimation procedures and results can be found in the Online Appendix.

Secondly, we simulate remaining life-cycle paths for mortality, health, consumption, wealth, and leisure for a sub-sample of the HRS respondents using the estimated parameter values and age sixty data as initial conditions. The simulation sample consists of all individuals in the initial HRS cohort with age sixty data and the requisite lagged data for simulations. Further information on initial condition descriptives, sampling weights and representativeness, and simulation procedure is also provided in the Online Appendix.

Finally, we use our expected utility framework, detailed in the following section, to embed the simulated data and construct a measure of ex-ante welfare at age sixty for each individual in our simulation sample by racial/ethnic group.

3 Welfare Measure

We extend and modify the measure proposed by Miller and Bairoliya (2022) 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:

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

Here, c represents consumption (in thousands of dollars), l represents leisure, h represents health, b represents bequests, and Ψ is a survival indicator. We assume log utility over consumption and additive separability with leisure, allowing for a simple decomposition of results. We also report robustness checks where we relax these assumptions. The health measure h is a vector of indicators for each modelled morbidity and self-rated health. We assume that utility from consumption and leisure is scaled by the health function $\phi(h) \in [0,1]$. Note that $\phi(h) = 1$ represents the utility for a person in perfect health, and $\phi(h) = 0$ represents the utility for a person who is dead. By combining the survival indicator with the health function, we obtain 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 welfare gains from leaving bequests, as it could quantitatively contribute to driving inequalities across racial and ethnic groups, since bequests can be significant and are likely correlated with health and consumption.

We define welfare using a consumption-equivalent variation measure. In particular, we define welfare for an individual i at age j to satisfy the following condition:

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

Here, ψ_m, h_m, l_m, and b_m are reference levels of survival, health, leisure, and bequests chosen by the individual. The welfare λ_{ij} is 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 the individual. For instance, if $\lambda_{ij} = 20$, it means that the individual would be indifferent between receiving their own stochastic outcome profiles moving forward or receiving $\$20,000$ in annual consumption together with the reference profiles for health, leisure, bequests, and survival.

The welfare condition can be rearranged to yield an 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, Φ is defined as follows:

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

and $\tilde{\psi}$ is the reciprocal of the reference discounted quality-adjusted life expectancy, and E_{ψ} denotes expected values conditional on survival.

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

3.1 Calibration

To calibrate the preference parameters, we assume that the health utility is directly proportional to the health state vector, represented as $\phi(h_t) = \gamma h_t$. To determine the utility weights vector γ, we follow the methodology of Miller and Bairoliya (2022) and utilize the Health Utilities Index Mark 3 (HUI3) instrument. Further details regarding the calibration process can be found in the Online Appendix. The HUI3 has been extensively employed in the health utility literature (Furlong et al., 1998; Feeny et al., 2002; Horsman et al., 2003), and data from the year 2000 collection of a subset of Health and Retirement Study (HRS) respondents were used.

For retired individuals, we normalize leisure time to one, while for workers, we set leisure time to 0.66, assuming an endowment of $5,840$ hours per year (16 hours a day $\times$ 365 days), where workers supply $2,000$ hours of labor. We define preferences over leisure time using the function $v(l) = -\frac{\theta\epsilon}{1+\epsilon}(1-l)^{\frac{1+\epsilon}{\epsilon}}$, where ϵ represents the constant Frisch elasticity of labor supply. In line with Jones and Klenow (2016), we set $\epsilon = 1$ and derive a benchmark disutility weight of $\theta = 7.82$, such that the marginal cost of leisure is equated to the marginal benefit for the median individual in our sample.

Furthermore, we choose a discount factor of $\beta = 0.98$, which corresponds to an annual discount rate of one percent, in line with previous studies (De Nardi, 2004). We define preferences for bequests using the function $\zeta(b) = \Phi_1 \left(1 + \frac{b}{\Phi_2}\right)^{1-\sigma}$, where Φ_1 reflects the strength of the bequest motive and Φ_2 measures the extent to which bequests are a luxury good. Consistent with De Nardi (2004), we set $\Phi_1 = -9.5$, $\Phi_2 = 11.6$, and $\sigma = 1.5$ for our benchmark calibration.

With the preferences defined above, a retired individual will prefer life to death as long as the flow intercept $\bar{u}$ plus log consumption is positive. We set $\bar{u} = -log(2)$, which implies that $2,000 of consumption is needed for a retiree to maintain positive flow utility. This is approximately 10% of the 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 median value of remaining life for sixty-year-olds of about $60,000 per QALY in our simulation sample, which falls within the range of typical values reported in the literature (Ryen and Svensson, 2015; Kaplan and Bush, 1982). For more details, see the Online Appendix.

3.2 Reference Outcomes

To calculate welfare, we need to define reference profiles that will be used for all individuals. For leisure, we choose retirement by age sixty as our reference, meaning full leisure from age sixty onward. For health-adjusted welfare equivalents, the standard approach is to use a notion of "normal" or "good" health as the reference. This allows us to compare individuals based on their consumption differences. We follow the approach of Miller and Bairoliya (2022) and use a constant reference health level of $\phi(h_{ma}) = 0.8$

and a reference sixty-year-old life expectancy of 24 years. To ensure the robustness of our analysis, we also conduct a sensitivity analysis using a longer reference life expectancy. Finally, we choose a reference bequest level of $\$500,000$. In summary, we assume that we can compare the welfare of age 60 retirees who expect to live to age 84 in "good" health and leave a bequest of $\$500,000$ solely based on expected consumption profiles.

4 Welfare Results

Our presentation of welfare analyses across racial groups includes (1) age sixty descriptive statistics, (2) model estimates, (3) mean outcomes and welfare measures, (4) decomposition exercise, (5) selected morbidity counterfactuals, 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

	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.

Table 1. Simulation Sample Age Sixty Descriptive Statistics by Race/Ethnicity

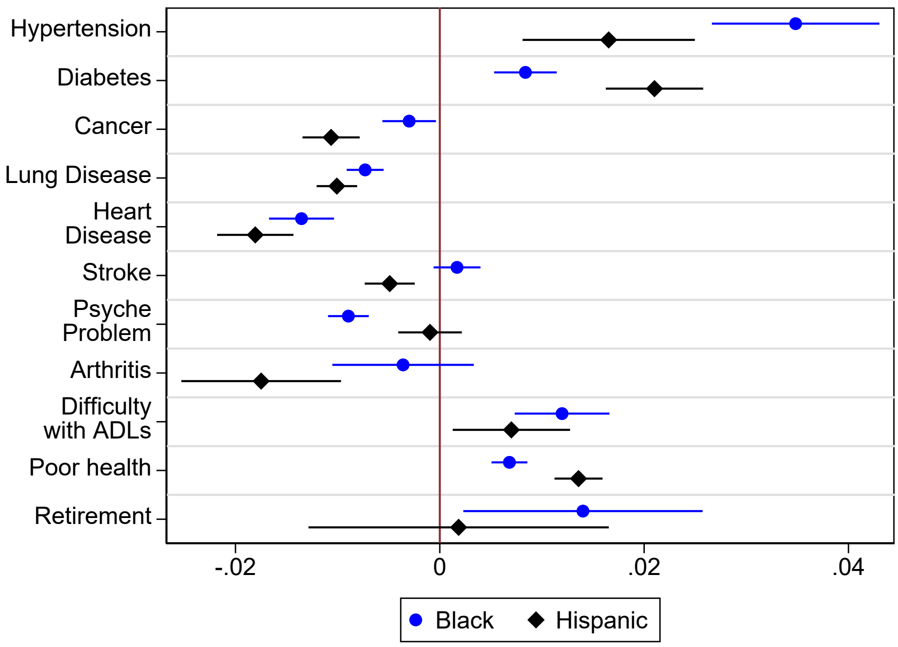
Table 1 summarizes the initial conditions at age sixty in the simulation sample, grouped by race/ethnicity. The prevalence of hypertension, diabetes, stroke, arthritis, and difficulty with activities of daily living (ADLs) was higher among Black and Hispanic respondents than White respondents. However, stroke and arthritis were exceptions for Hispanic respondents. For instance, the reported incidence of diabetes was 23% for Black respondents and 19% for Hispanic respondents, compared to only 10% for White respondents. This represents a 2.3-fold and 1.9-fold difference, respectively. In terms of self-reported health, 14% of Black respondents and 19% of Hispanic respondents reported poor health status, whereas only 6% of White respondents did so. Black and Hispanic respondents, on average, retired earlier than White respondents. Specifically, 56% of Black respondents and 60% of Hispanic respondents retired by age sixty, compared to only 50% of White respondents. Additionally, cross-sectional consumption at age sixty averaged $18,350 for Black respondents and $14,770 for Hispanic respondents, as opposed to $30,190 for White respondents. These differences correspond to a 1.6-fold and 2-fold difference, respectively. Finally, at age sixty, 48% of Black and 71% of Hispanic respondents reported less than a high school education, while only 24% of White respondents did so.

4.2 Model Estimates

This section presents some selected results from our simulation model aimed at gaining a better understanding of the correlation between race and other outcomes in the data. In particular, we use Figure 2 to illustrate the estimated average marginal effects of race on various health and retirement indicators. Our findings reveal that, in comparison to White respondents, Black and Hispanic respondents have a higher likelihood of experiencing health problems such as hypertension, diabetes, stroke, difficulty with ADLs, self-rated poor health, and early retirement (although stroke is an exception for Black respondents). For instance, compared to White respondents, Black and Hispanic respondents have a marginal increase in the probability of hypertension by about 2.8 and 1.8 percentage points (pp), respectively. We also discovered that race, in relation to morbidities, is associated with self-rated health. For example, the average marginal increase in the probability of reporting poor health is approximately 0.8 pp for Black respondents and 1.7 pp for Hispanic respondents.

4.3 Welfare Gaps in EHRS Cohort

This analysis examines the mean outcomes and distribution of a welfare measure across racial and ethnic groups of sixty-year-olds from the EHRS cohort, as presented in Table 2. Panel A displays the mean consumption, retirement, life expectancy, QALE, and expected bequests at age sixty, while Panel B shows the cumulative contribution of each factor to the welfare measure. Additionally, the mean Black-White and Hispanic-White outcome and welfare ratios are presented.



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

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

In Panel A, it is observed that the annual consumption for Black respondents at age sixty is approximately 61% of that for White respondents (Black-White ratio of 0.61). The corresponding estimate for Hispanic respondents is around 49% (Hispanic-White ratio of 0.49). Black and Hispanic respondents retire earlier than White respondents overall, with a Black-White and Hispanic-White ratio of about 1.11 (or 110%) and 1.20 (or 120%), respectively. Moreover, White respondents have higher life expectancy, QALE, and financial bequests. For instance, the life expectancy of Black respondents averages around 18.6 years 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 more significant in QALE, with about 14 years for Hispanic respondents compared to 17 years for White respondents. Additionally, the expected financial bequests for Black and Hispanic respondents are approximately 26% and 24% of that for White respondents, respectively.

Even if differences in expected leisure, life expectancy, health, and financial bequests are ignored, Panel B shows a substantial overall welfare gap between races/ethnicities. The "consumption" Black-White welfare ratio is approximately 0.62 (or 62%), while the Hispanic-White welfare ratio is 0.51 (or 51%). Furthermore, the 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.

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.

Table 2. Outcomes and Welfare by Race/Ethnicity

Slight adjustments for lost leisure due to working past age sixty slightly decrease the welfare gap, increasing 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 are expected to retire earlier than White respondents overall, but these differences have only small effects on the fully-adjusted measure of racial and ethnic 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 an average of $900 and $694 in expected annual consumption to retire at age sixty, respectively.

Health and life expectancy are essential 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 displays adjustments for leaving financial bequests, yielding the fully-adjusted welfare measure. Smaller financial bequests are almost as important to estimated welfare gaps as health and longevity. Adjusting for bequests lowers the Black-White welfare ratio by an additional 10 pp and the Hispanic-White ratio by an additional 8 pp.

Figure 3 presents the average expected life-cycle profiles, which can help us understand the differences between racial and ethnic groups. The mean gaps in consumption, retirement, and life expectancy are most significant at age sixty and decrease gradually as

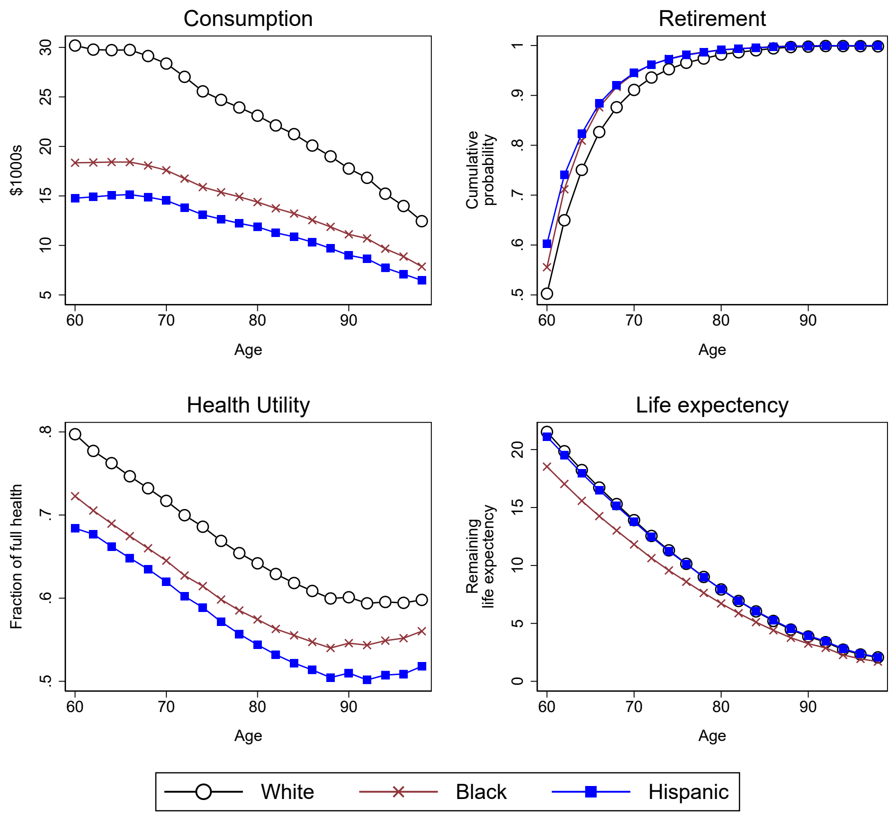


Figure 3. Average Life Cycle Profiles by Race/Ethnicity

individuals age. However, consumption gaps still exist even into the nineties. On the other hand, health gaps are significant at age sixty and persist throughout the remaining life. Our welfare results indicate that consumption, health, life expectancy, and bequests are crucial factors that contribute to racial and ethnic welfare inequality, while earlier retirement and leisure have a relatively minor impact.

Figure 4 illustrates the cumulative change in the distribution of log welfare at age sixty across racial and ethnic groups in greater detail. Adjusting for leisure, life expectancy, health, and bequests has a greater negative impact on the welfare distribution of Black and Hispanic respondents than White respondents. It is worth noting that the adjustments cause inequality *within* the Black and Hispanic respondent populations to increase more than the White population (i.e., the left tail of the welfare distribution becomes fatter). This is consistent with existing evidence on inequality, which shows that relative income disparity between the top and bottom 10 percent is particularly acute for Black Americans. Pew Research Center reported in 2016 that the 90th percentile of Black households earned nearly ten times as much as the 10th percentile (Pew Research Center, 2018).

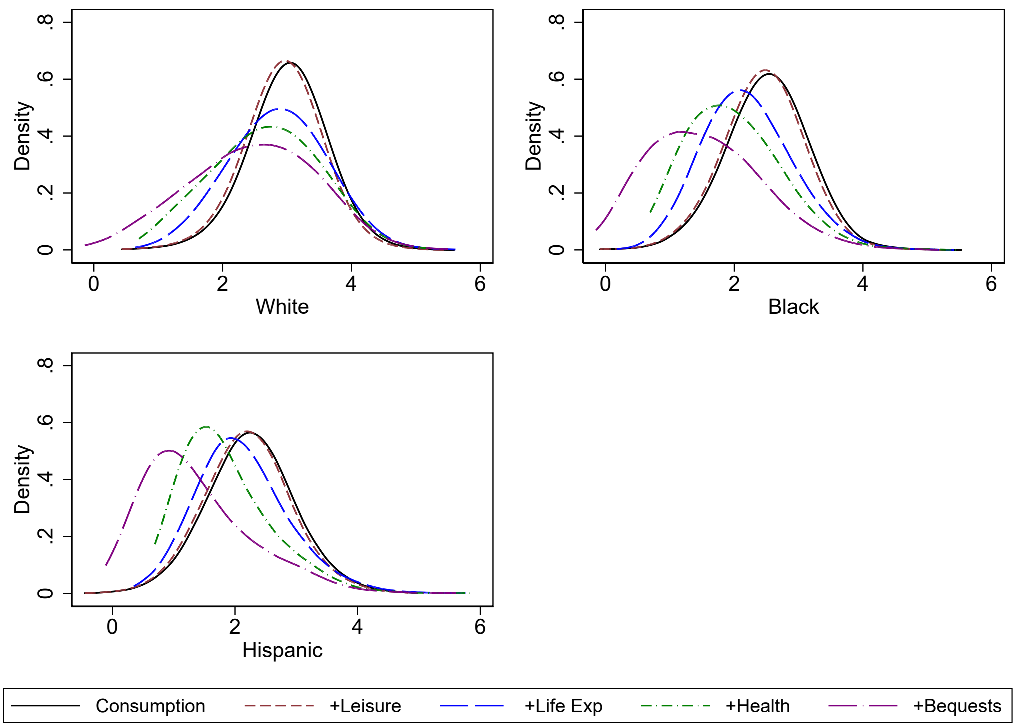


Figure 4. Cumulative Change in Distribution of Log Welfare by Race/Ethnicity

4.4 Decomposition

In our estimates, the increase in welfare inequality across racial and ethnic groups can be attributed to two potential factors: (1) differences in the distribution of initial conditions at age sixty across races and ethnicities and/or (2) differences in the stochastic processes experienced by each racial and ethnic group after age sixty. The main question we aim to answer in this analysis is: to what extent do initial conditions at age sixty versus differences in outcome dynamics after age sixty explain the racial and ethnic welfare gaps? To address this question, we conduct several experiments to estimate the impact of initial differences at age sixty as well as differential outcome evolutions across racial and ethnic 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 results are presented in Table 3, where 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.

In our first round of experiments, we assigned transition probabilities of White participants after age sixty to Black and Hispanic groups to investigate how the evolution of outcomes after sixty affects gaps in QALE, ELC, and welfare. However, as displayed in the second row of Table 3, the differences in the evolution of outcomes can only account

Experiment	QALE ratio		ELC ratio		Welfare ratio	
	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.

Table 3. Decomposition

for a small portion of the racial and ethnic welfare gaps. For instance, 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. Surprisingly, outcomes for Hispanic respondents become slightly worse when given White transition probabilities, with the QALE ratio decreasing by 4 pp, ELC ratio by 3.2 pp, and the fully-adjusted welfare ratio by 1.8 pp.

We then shifted our focus to the role of age sixty differences in explaining the estimated racial and ethnic welfare gaps. Our previous experiment only changed the evolution of outcomes after age sixty, while keeping the initial distribution of outcomes the same for each racial and ethnic group. As indicated in the last row of Table 3, when we instead assign the initial conditions of White respondents to Black and Hispanic groups, the estimated Black-White and Hispanic-White ratios in QALE, ELC, and fully-adjusted welfare measures increase significantly. Equating initial conditions enhances the Black-White welfare ratio by 50 pp and the Hispanic-White ratio by 70 pp. Our decomposition exercises indicate that the majority of the estimated welfare gaps are determined by age sixty initial conditions rather than racial and ethnic differences in dynamic processes after age sixty.

4.5 Morbidity Counterfactuals

This section aims to investigate how morbidities affect outcomes and welfare across different racial and ethnic groups. Table 4 displays the increase in quality-adjusted life expectancy (QALE) and expected lifetime consumption (ELC), the loss in bequests, and the Black-White/Hispanic-White welfare ratios when all hypertension or diabetes cases are eliminated after age sixty. We selected hypertension and diabetes since our model's estimates (refer to Figure 2) indicate that, among other health measures and exogenous characteristics, these two conditions have the most significant racial and ethnic disparities. Additionally, hypertension and diabetes are established risk factors for various downstream health issues, including 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).

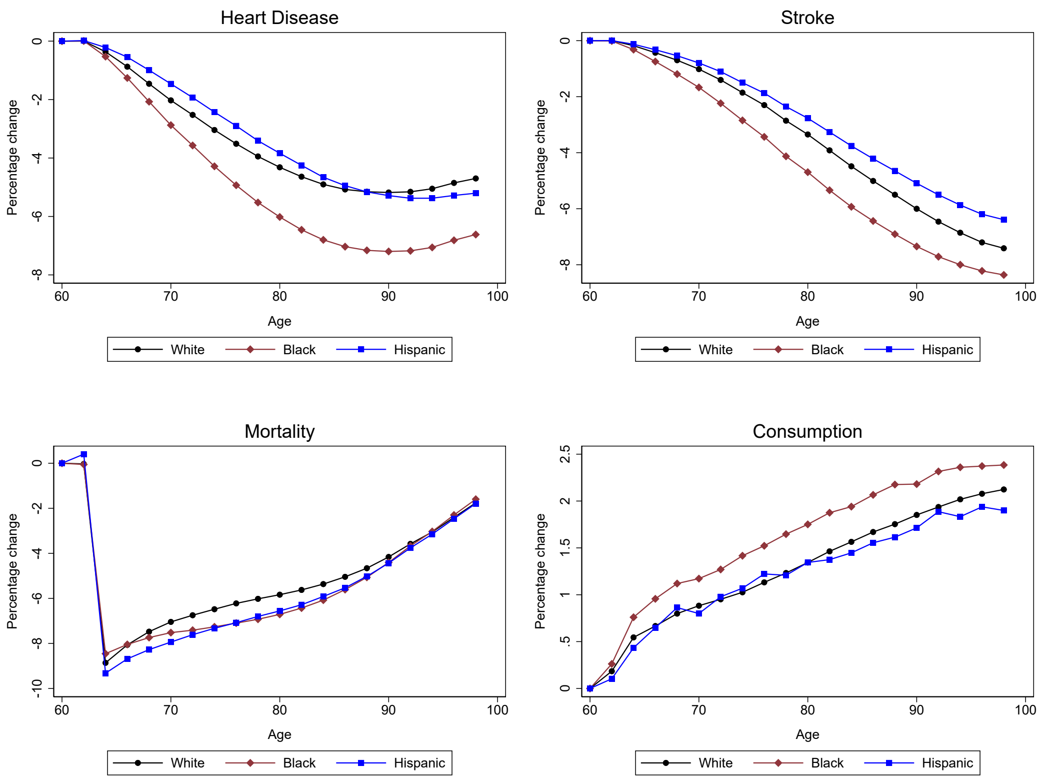
	Hypertension			Diabetes		
Outcomes and Welfare	White	Black	Hispanic	White	Black	Hispanic
QALE gain	1.241	1.462	1.409	0.699	1.040	1.251
ELC gain	32.152	26.594	19.484	17.926	19.788	17.016
Bequest loss	13.947	5.390	5.116	5.499	2.385	3.460
Welfare ratio	–	0.384	0.341	–	0.391	0.356
Baseline ratio	–	0.379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

Table 4. Eliminating Late-life Hypertension and Diabetes by Race/Ethnicity

Table 4 shows that eliminating hypertension resulted in Black and Hispanic respondents gaining slightly more QALE than White respondents at age sixty. Specifically, Black and Hispanic respondents gained about 1.5 and 1.4 years, respectively, compared to 1.2 years for White respondents. However, White respondents had a higher gain in ELC of \$32,152 compared to \$26,594 for Black and \$19,484 for Hispanic respondents due to their larger annual consumption. But this gain in lifetime consumption was partially offset by a larger decline in bequests for White respondents (\$13,947) compared to Black (\$5,390) and Hispanic (\$5,116) respondents. Eliminating late-life diabetes had similar patterns but with smaller effects. Black respondents gained more lifetime consumption than White respondents, and Hispanic respondents had higher QALE gains (and bequest losses) than Black respondents. However, overall, the counterfactual welfare ratios suggest that eliminating these diseases only marginally closes overall welfare gaps. Eliminating hypertension after age sixty increases the Black-White welfare ratio by 0.005 pp but 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.

To better understand how morbidities influence the dynamics of other outcomes in the system across racial and ethnic groups, Figures 5 and 6 illustrate the average percentage change in several expected outcomes with the exogenous elimination of hypertension and diabetes after age sixty. Eliminating hypertension after age sixty reduces the average probability of developing heart disease and stroke for all races and ethnicities, with the strongest changes for Black respondents. For example, Black respondents experienced a decreased probability of heart disease of about 6% by age eighty compared to approximately 4% for White and Hispanic respondents. Similarly, the probability of stroke by

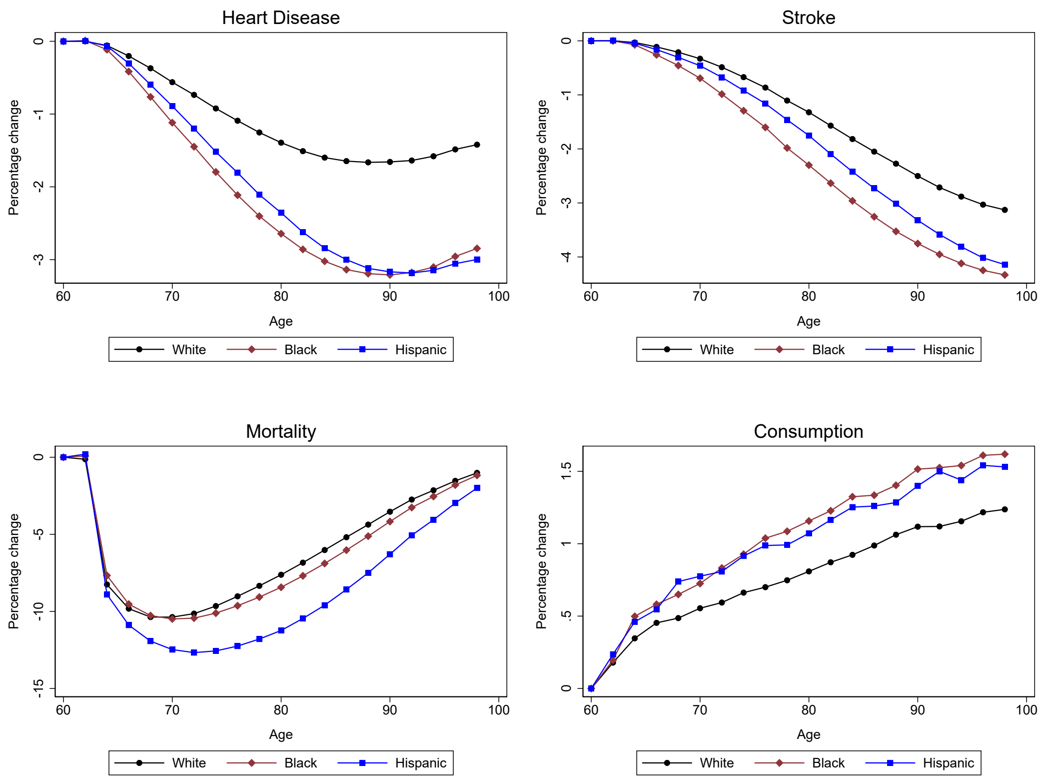


Notes: Results plot percentage difference in expected outcomes with the exogenous elimination of hypertension after age sixty relative to baseline. Sample includes all individuals in the simulation sample from the EHRS cohort. Expected outcomes are conditional on survival.

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

age eighty decreased by about 5% for Black respondents compared to 3% for White and Hispanic respondents. Interestingly, although Black respondents saw the largest gains in consumption, we see similar mortality gains for Hispanic respondents.

Compared to eliminating hypertension, general patterns are similar in the diabetes experiment, with the main difference being 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.



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.

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

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 and ethnicity specific forecasting model, a higher reference life expectancy and reference bequests, and alternate preference parameter values. 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.

4.6.1 Race/Ethnicity Specific Simulation Model

In our benchmark simulation model, we allowed dynamics to vary across race/ethnicity through a race/ethnicity intercept (or individual fixed effect for consumption and wealth). However, we assumed that other model parameters were the same for all racial and ethnic groups. For instance, we assumed that the direct effect of diabetes on self-rated health was identical for White, Black, and Hispanic respondents. In contrast, the "race-specific forecast" results in Table 5 were obtained by separately estimating a fore-

	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
Reference bequests	17.644	6.690	6.106	0.379	0.346
$\bar{u}$ = -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
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.

Table 5. Sensitivity of Mean Welfare by Race/Ethnicity

casting model for each of the three groups. This approach has the disadvantage of a loss in precision and fewer observations, especially for the Hispanic sample. However, we found that mean welfare only slightly increased for White and Hispanic respondents and slightly decreased for Black respondents when using this approach. As a result, the Black-White welfare ratio decreased by only 1 pp, and the Hispanic-White ratio increased by 2 pp compared to our benchmark results.

4.6.2 Reference Life Expectancy and Bequests

The third row of Table 5 shows sensitivity of results when we increase the reference age sixty life expectancy from 24 to 30 years. As is clear from equation (10), increasing reference life expectancy is more costly to log welfare for those with higher flow utility. Thus we see larger mean declines in welfare for White respondents, with a corresponding increase in the Black-White welfare ratio of 8 pp and in the Hispanic-White ratio of 4 pp. The next row in Table 5 provides results when the reference bequest level is increased from $\$500,000$ to one million dollars. Quantitatively, this has a much smaller effect on mean welfare than reference life expectancy, and welfare ratios are unchanged compared to the benchmark.

4.6.3 Preference Parameters

The remainder of Table 5 presents sensitivity results for our choice of calibrated preference parameter values. We first check the sensitivity of results to our choice of flow intercept $\bar{u}$. Specifically, we set $\bar{u} = -log(1.5)$, implying that $\$1,500$ of consumption is needed for a retiree to maintain positive flow utility compared to our benchmark value of

$\$2,000$. The change has only a small impact on estimated welfare inequality, decreasing both reported ratios by about 1 pp. With a lower time discount rate $\beta = 0.9$, anticipated gaps in future consumption and health are less important for welfare. As such, the Black-White welfare ratio increases about 6 pp and the Hispanic-White ratio 4 pp. The welfare ratios increase by a similar magnitude when we decrease the strength of the bequest motive Φ_1 by roughly half compared to the benchmark. Changes in Frisch elasticity of labor supply ϵ, disutility weight on labor supply θ, and the other bequest parameters Φ_2 and σ, each have very small impact on inequality results. Lastly, in our benchmark estimates we calibrated health utility weights by assuming that consumption and leisure were conceptualized as fixed across health states by HRS respondents that completed the HUI3 (see the Online Appendix for full discussion on this assumption and how it can be relaxed). The last row of Table 5 shows that results are largely insensitive to relaxing this assumption.

4.6.4 Consumption and Leisure Utility

In our study, we also investigate the reliability of our findings using a more general form of flow utility for consumption and leisure, represented by the following equation:

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

When $\gamma = 1$, this formula is equivalent to our benchmark case. However, when $\gamma > 1$, the curvature over consumption increases. This creates several challenges. Firstly, it becomes impossible to determine welfare for individuals at the very top of the health distribution, as increasing their consumption would never provide the same expected life-time utility as the reference life expectancy. Therefore, we report the median welfare instead of the mean welfare in Table 6. Secondly, higher curvature creates another issue: as γ increases, the implied value of life rises steeply. As shown in the first column of Table 6, the median value of life per QALY is $\$178,000$ per QALY when $\gamma = 2$, which is high but still reasonable. The estimated median Black-White and Hispanic-White welfare ratios also increase modestly to 0.39 and 0.36, respectively. When $\gamma = 3$, the value of life reaches about $\$557,000$ per QALY, and the welfare ratios increase more substantially to 0.58 and 0.54. However, only three out of 23 value of life studies surveyed by Ryen and Svensson (2015) estimated a mean value of life over $\$150,000$. Therefore, caution should be exercised when interpreting robustness results with high curvature values, as the value of life may be overstated. Nonetheless, higher curvature values provide an understanding of the robustness of key results.

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

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

Table 6. Sensitivity for Higher Curvature–Median Welfare by Race/Ethnicity

5 Conclusion

We propose and estimate an individual measure of welfare incorporating heterogeneity and uncertainty in future consumption, leisure, health, wealth and mortality at age sixty. Our measure broadly indicates that racial and ethnic 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/ethnic 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/ethnic 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/ethnic 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.

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 (2022). Health, Longevity, and Welfare Inequality of Older Americans. *(Forthcoming) The Review of Economics and Statistics*.

Miller, R., S. Chin, and A. K. Sedai (2022). The welfare cost of late-life depression. *Journal of Economic Behavior & Organization 204*, 15–36.

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.

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.