Mortality Risk, Insurance, and the Value of Life*

Daniel Bauer
University of Alabama

Darius Lakdawalla
University of Southern California and NBER

Julian Reif
University of Illinois

Abstract. We develop a novel economic framework for valuing improvements in health and apply it to data in order to explore the relationships between annuity programs, the value of life-extension, and the value of preventing illness. Incorporating incomplete annuitization and stochastic mortality into the conventional economic theory of life-extension generates several new findings. First, public annuity programs boost the demand for life-extension. For instance, US Social Security adds $11.5 trillion (10.6 percent) to the current value of post-1940 longevity gains. Second, in contrast to the conventional theory, a given mortality improvement may be worth more, not less, to patients facing shorter lives. Holding income and wealth constant, the value of statistical life (VSL) can soar by over $1 million following a health shock that lowers life expectancy. Thus, existing economic analysis may be undervaluing treatment of severe illnesses relative to mild ones. This result also reconciles an empirical puzzle with the economic approach to valuing life, because consumers often report a preference for extending life among people with the bleakest survival prospects. Finally, we introduce a new concept, the value of statistical illness (VSI), which quantifies the willingness-to-pay to avoid falling ill and includes VSL as a special case. Our framework implies that treatments are worth more than prevention, all else equal. Using real-world data, we calculate that treating illnesses such as cancer and heart disease is worth two to five times more than saving an equivalent number of life-years by preventing these conditions.

* We are grateful to Dan Bernhardt, Tatyana Deryugina, Sonia Jaffe, Nolan Miller, Alex Muermann, George Pennacchi, Mark Shepard, Dan Silverman, George Zanjani, and participants at the AEA/ARIA meeting, the NBER Insurance Program Meeting, Temple University, the University of Chicago Applications Workshop, and the University of Wisconsin-Madison for helpful comments. Bauer acknowledges financial support from the Society of Actuaries. Lakdawalla discloses that he is the Chief Scientific Officer of Precision Health Economics (PHE) and an investor in its parent company, Precision Medicine Group. PHE provides research and consulting services to firms in the life sciences and health insurance industries.

I. INTRODUCTION

The economic analysis of risks to life and health has made enormous contributions to both academic discussions and public policy. Economists have used the standard tools of life-cycle consumption theory to propose a transparent framework that measures the value of improvements to both health and longevity (Arthur 1981; Rosen 1988; Murphy and Topel 2006). Economic concepts such as the value of statistical life now play central roles in public policy discussions surrounding investments in medical care, public safety, workplace safety, environmental hazards, and countless other arenas.

The standard framework has typically assumed complete annuitization and deterministic mortality risk. While analytically convenient and useful for illustrating some of the underlying economics, these assumptions are not realistic: it is well known that most people are incompletely annuitized (Brown et al. 2008), and that mortality risk changes over time according to one's health state. Moreover, these assumptions hamper the standard model's predictive power in several ways: they gloss over policy-relevant relationships between the value of life and the structure of the annuity market, cannot investigate what happens to the value of life upon falling ill, and cannot meaningfully distinguish between preventive care and medical treatment.

This paper develops a general economic framework for valuing health improvements and applies it to data. We establish three main results. First, we calculate that the US Social Security program added \$11.5 trillion (10.6 percent) to the value of post-1940 longevity gains. Second, we derive conditions under which the value of life can *rise* following a negative shock to life expectancy, and demonstrate that this effect is economically significant: for example, the value of statistical life (VSL) for a healthy 50-year-old soars by over \$1 million following a health shock that lowers life expectancy by several years. Third, we introduce the value of statistical illness (VSI), which quantifies the willingness-to-pay to avoid falling ill and includes VSL as a special case. Using this new concept, we calculate that—holding wealth constant—a sick individual's willingness-to-pay for medical treatments is two to five times greater than a healthy individual's willingness-to-pay for equally effective preventive treatments.

Incomplete annuity markets drive all three of these results. While complete annuity markets shield an individual's consumption against longevity risk, an incompletely annuitized consumer will have a different consumption profile, which in turn affects her value of life. A very simple example illustrates the intuition. Imagine a 60-year-old retiree with no bequest motive and a flat optimal consumption profile. If she fully annuitizes her savings, her consumption remains flat at, say, \$30,000 annually. Now suppose she cannot annuitize any of her wealth. It is well known that in this case it is optimal to shift consumption forward (Yaari 1965), because consumption allocated to later time periods will not be enjoyed in the event of an early death (see Figure 1). Because VSL depends greatly on consumption, it too will shift forward. Thus, an increase in the annuitization rate will raise VSL at older ages. If this in turn increases spending on elderly healthcare, it will generate a positive relationship between public spending on annuity and elderly healthcare programs.

Our other results follow from the simple observation that it is optimal for a non-annuitized individual to shift her consumption forward, i.e., to spend down her wealth, following an adverse stochastic shock to mortality. At least for some initial period of time, a rise in mortality risk increases consumption, and thus reduces the marginal utility of consumption. An important insight of our paper is that although this rise in mortality risk always reduces lifetime utility, the accompanying reduction in the contemporaneous marginal utility of consumption can be large enough to cause the value of statistical life to *increase* even when life expectancy falls. Indeed, we show that the value of statistical life is frequently higher for an individual diagnosed with a more fatal illness, and vice-versa. This is in stark contrast to the conventional model with full annuitization, where a reduction in survival always reduces the value of statistical life.

The first half of this paper provides a formal framework that confirms these insights. We first review the conventional model employed by prior studies and show that relaxing its assumption of full annuitization causes both consumption and the value of life to shift forward. We then further generalize this framework by taking the more realistic perspective that an individual faces uncertainty over her future mortality risk. Allowing mortality to be stochastic produces additional insights. First, we demonstrate that consumption increases following an adverse shock to mortality, and provide a set of sufficient conditions under which the shock also generates an accompanying increase in the value of statistical life.[1] Second, although the conventional model employed in prior studies quantifies the value of statistical life, it has little to say about the continuum of health events that precede death. Our framework by contrast lends itself naturally to a more general concept, the value of statistical illness (VSI), which quantifies an individual's willingness to pay to avoid an increase in the risk of acquiring an illness. This allows for the first time an economic comparison of the value of prevention to the value of treatment. In contrast to the convention among health practitioners, we show how the value of treatment technologies, which are consumed after an illness occurs, is larger than the value of preventive technologies, even when both increase life expectancy by identical amounts. This is because treatments are administered in bleaker health states, where the marginal utility of consumption is low.

The second half of the paper applies our model to real-world data. Our first empirical exercise illustrates the connections between public annuity programs and the societal value of mortality reductions. Combining datasets on life-cycle earnings and historical mortality rates, we calculate that the US Social Security program added \$11.5 trillion (10.6 percent) to the value of post-1940 longevity gains. To put this in perspective, this gain is worth over \$35,000 per person to the current population, or about half as much as the longevity insurance value of Social Security. We also calculate that Social Security has increased the aggregate value of reducing future mortality risks by over 10 percent, so that a 1 percent reduction in population-wide mortality is \$143 billion more valuable than it would have been without the program. Moreover, increasing the size of Social Security pensions by 50 percent would add a further \$49 billion of value to this mortality decline. Intuitively, annuitization tends to raise the value of life-extension at older ages where people might otherwise have outlived their wealth. Since the absolute number of deaths is quite high at those ages, this also tends to boost the value of proportional reductions in mortality.

Our second set of empirical exercises employs data from the Future Elderly Model (FEM), a well-published microsimulation model of health, mortality, and aging. The FEM provides age-specific data on mortality, quality of life, and medical spending for twenty different health states. We incorporate these detailed data into a stochastic life-cycle model that allows for shocks to mortality, quality of life, and wealth. The model illustrates that VSL depends on an individual's health history. For instance, we find that VSL soars from \$1.5 million to \$2.9 million for an 80-year-old who suffers a debilitating health shock that reduces her life expectancy by 7 years. In many empirically realistic cases, we demonstrate the surprising theoretical result that the value of a statistical life can go up when life expectancy falls. This prediction is impossible under complete annuitization, but theoretically possible without it. This relationship between mortality shocks and VSL generates substantial variability in VSL in the aggregate:

[1] Intuitively, the sign depends on whether the loss in lifetime utility is offset by a corresponding decrease in marginal utility. Specifically, it depends on a trade-off between the elasticity of intertemporal substitution, which measures the curvature of the utility function, and prudence, which measures the curvature of the *marginal* utility function. An adverse mortality shock increases VSL when demand for current consumption is sufficiently inelastic, or when the marginal utility of demand is sufficiently linear. See **Proposition 6** for a formal proof.

a Monte Carlo simulation of a set of initially healthy, identical 50-year-olds estimates that stochastic health shocks generate an inter-vigintile (middle 90 percent) VSL range of $4-5 million by age 60.[2]

We also use the FEM data to examine how VSL varies across different health states, and find that the value of life-years gained through medical treatment is higher in states with lower remaining life expectancy. Finally, we calculate that the value of treating life-threatening conditions like cancer is worth 2 to 5 times more than equivalent preventive treatments that add the same number of years to an individual's life expectancy.

Our primary contribution is the development and application of a novel model of the value of health improvements. Jettisoning the unrealistic assumptions of full annuitization and deterministic mortality commonly maintained in the prior literature permits us to address three new research questions: How does the value of life change following a health diagnosis? How does the value of treatment compare to the value of prevention? What is the relationship between annuitization and the value of life? In doing so, our study connects the large literature on the value of life (Arthur 1981; Shepard and Zeckhauser 1984; Murphy and Topel 2006; Hall and Jones 2007) with the vast literature on annuities and life-cycle consumption models that goes back to Yaari (1965). It is well known that annuitization provides substantial value by insuring individuals against longevity risk. We show that it also increases the value of statistical life at older ages, and the value of mortality reductions in the aggregate. Our results suggest that researchers and policymakers should pay more attention to the public finance interactions between pension and healthcare systems.

Our findings have two significant implications for cost-effectiveness analysis, which governs the allocation of healthcare resources in many "single-payer" countries, including the United Kingdom, Canada, and Australia (Dranitsaris and Papadopoulos 2015), and which continues to grow in importance in the multi-payer US healthcare marketplace (Goldman, Nussbaum, and Linthicum 2016). First, standard cost-effectiveness analysis asserts that the value of extending life is insensitive to the severity of illness. For instance, it implies equivalence between providing X aggregate life-years to a very large population of hypertension patients and providing X aggregate life-years by extending life substantially for a proportionally smaller population of cancer patients. Our model finds that this equivalence is incorrect when individuals are not fully annuitized. In particular, it suggests that the cost-effectiveness approach to healthcare resource allocation underinvests in the treatment of the most life-threatening illnesses relative to less severe conditions. This insight is consistent with data on how consumers view the value of life-extension (Nord et al. 1995; Green and Gerard 2009; Linley and Hughes 2013), and can better inform the way health economists and healthcare payers assess the value of medical technologies.

Second, cost-effectiveness analysis traditionally values life-years gained by prevention and treatment equally (Drummond et al. 2005a). However, we demonstrate theoretically that baseline health status affects the value of life-years gained, and this creates a wedge between prevention and treatment. In contrast to the old adage, we find that treatment is often significantly more valuable than prevention, even when they produce the same longevity gain.

Finally, extending the value of life analysis to a stochastic mortality setting requires us to rely on tools from continuous-time stochastic optimal control. To derive the expressions for VSL and VSI, we rely on a "stochastic" version of the Pontryagin maximum principle following recent developments in the

[2] The Monte Carlo simulation is repeated 10,000 times. We solve this large number of stochastic life-cycle models using a recursive analytical formula, which allows for quick and exact calculations. A complete derivation is available in Appendix C2.

systems and control literature (Parpas and Webster 2013). The resulting expression for VSL generalizes the deterministic versions in the earlier literature (Rosen 1988; Murphy and Topel 2006), and VSI can in turn be interpreted as a generalization of the concept of VSL. We view our application of these tools as a useful demonstration for other researchers working in stochastic settings.

Section II reviews the predictions of the conventional model for the returns to life-extension and demonstrates how relaxing the perfect annuity assumption alters these predictions. Section III then generalizes the framework further by incorporating the more realistic assumption of stochastic mortality. Section IV presents empirical analysis that: (1) estimates the effect of public annuity programs on the value of statistical life; (2) quantifies how health shocks change the value of statistical life when annuity markets are incomplete; (3) illustrates how more severe health shocks cause consumers to place higher value on a given mortality reduction; and (4) calculates the value of preventing different kinds of illness. Section V concludes.

II. THE VALUE OF LIFE WHEN MORTALITY IS DETERMINISTIC

Consider an individual who faces a mortality risk. We are interested in analyzing the value of a marginal reduction in this risk. We first quantify this value in the conventional setting where markets are complete and the consumer has access to actuarially fair annuities (Rosen 1988; Murphy and Topel 2006). We then repeat this exercise in a "Robinson Crusoe" economy where the consumer cannot purchase annuities to insure against her uncertain lifetime (Shepard and Zeckhauser 1984; Ehrlich 2000; Johansson 2002). We compare our findings for these two polar cases to illustrate the basic insights of the paper. We focus on improvements in longevity and their relationship to annuity insurance markets, but allow for improvements in quality of life as well. Section III then extends the model to accommodate stochastic mortality and introduces the value of statistical illness.

Although it is optimal for a consumer to fully annuitize, real-world annuitization rates are quite low. This "annuity puzzle" is the subject of numerous papers. Many explanations have been suggested, but there is no consensus on what drives incomplete annuitization (Brown et al. 2008). Our model takes the low rate of annuitization as a given empirical fact and illustrates its significance for the value of life. Section IV uses a numerical model to probe the sensitivity of our results to different assumptions about consumer preferences, such as the presence of a bequest motive, which prior studies have argued might rationalize low observed rates of annuitization. There continues to be debate over why real-world consumption profiles and annuity purchase decisions look the way they do. However, as we show, the implications for life-extension depend primarily on the real-world consumption profiles themselves, not the reasons that lie beneath.

We focus throughout this paper on the willingness-to-pay for a marginal reduction in mortality risk. The extent to which this translates into an increase in health spending depends on the health production function. See Hall and Jones (2007) for additional discussion.

II.A. The fully annuitized value of life

Let $c(t)$ be consumption at time t, W_0 be baseline wealth, $m(t)$ be exogenously determined income, ρ be the rate of time preference, and r be the rate of interest.[3] Finally, define $q(t)$ as health-related quality of

[3] It is straightforward to incorporate endogenous labor supply (Murphy and Topel 2006). One could also allow income to depend on quality of life or mortality risk. This would have no qualitative effect on the results we present for the fully annuitized model. In the uninsured model presented later in the paper, it would reduce the value of life for working-age individuals who fall ill.

life at time t. Since it sacrifices little generality in our application, we take the life-cycle quality of life profile $q(t)$ as exogenous. As needed, one can consider any relevant quality of life profile in concert with a given profile of mortality. The maximum lifespan of a consumer is T, and her mortality (hazard) rate at any point in time is given by $\mu(t)$, where $0 \leq t \leq T$. The probability that a consumer will be alive at time t is:

$$S(t) = \exp\left[-\int_0^t \mu(s)ds\right]$$

At time $t = 0$, the consumer fully annuitizes. We assume that annuitization is actuarially fair. The consumer's maximization problem is:

$$V(0) = \max_{c(t)} \int_0^T e^{-\rho t} S(t) u(c(t), q(t)) dt$$
$$s.t. \int_0^T e^{-rt} S(t) c(t) dt = W_0 + \int_0^T e^{-rt} S(t) m(t) dt$$

The consumer's utility function, $u(c(t), q(t))$, depends on both consumption and health-related quality of life. We assume $u(\cdot)$ is strictly increasing and concave in its first argument, and twice continuously differentiable. Let $u_c(\cdot)$ denote the marginal utility of consumption. Associating the multiplier θ with the wealth constraint, optimal consumption is characterized by the first-order condition:

$$\frac{\partial V(0)}{\partial W} = \theta = e^{(r-\rho)t} u_c(c(t), q(t))$$

To analyze the value of life, let $\delta(t)$ be a perturbation on the mortality rate with $\int_0^T \delta(t)dt = 1$, and consider

$$S^{\varepsilon}(t) = \exp\left[-\int_0^t (\mu(s) - \varepsilon\delta(s))ds\right], \varepsilon > 0$$

Let $c^{\varepsilon}(t)$ represent the equilibrium variation in $c(t)$ caused by this perturbation. As shown in Rosen (1988), the marginal utility of this life-extension is given by

$$\left.\frac{\partial V}{\partial \varepsilon}\right|_{\varepsilon=0} = \left.\frac{\partial}{\partial \varepsilon} \int_0^T e^{-\rho t} S^{\varepsilon}(t) u\big(c^{\varepsilon}(t), q(t)\big) dt\right|_{\varepsilon=0}$$
$$= \int_0^T \left[e^{-\rho t} u(c(t), q(t)) + e^{-rt}\theta\big(m(t) - c(t)\big)\right] \left[\int_0^t \delta(s)ds\right] S(t)dt$$

The marginal value of life-extension is equal to the marginal rate of substitution between longer life and wealth:

$$\frac{\partial V/\partial \varepsilon}{\partial V/\partial W} = \int_0^T e^{-rt} S(t) \left(\frac{u(c(t), q(t))}{u_c(c(t), q(t))} + m(t) - c(t)\right) \left[\int_0^t \delta(s)ds\right] dt \qquad (1)$$

The value of a life-year is the value of a one-period change in survival from the perspective of current time:

$$v(t) = \frac{u\big(c(t), q(t)\big)}{u_c\big(c(t), q(t)\big)} + m(t) - c(t) \qquad (2)$$

The value of a life-year, $v(t)$, is equal to the value of consumption in that year plus net savings, $m(t) - c(t)$. The net savings term is a consequence of the requirement that annuities be actuarially fair. The value of a life-year can be rewritten as:

$$v(t) = m(t) + c(t)\left(\frac{u\big(c(t), q(t)\big)}{c(t)u_c\big(c(t), q(t)\big)} - 1\right) = m(t) + c(t)\phi(c, q)$$

where $\phi(c, q)$ represents the consumer surplus value per unit of consumption. It is positive if average utility exceeds marginal utility. A life-year adds value through two different channels: an increase in earnings, which can finance additional consumption, and an increase in consumer surplus.[4]

A canonical choice for $\delta(\cdot)$ in equation (1) is the Dirac delta function, so that the mortality rate is perturbed at $t = 0$ and remains unaffected otherwise. This then yields an expression that is commonly called the value of statistical life (VSL):

$$VSL \equiv \int_0^T e^{-rt} S(t) v(t) dt \tag{3}$$

VSL corresponds to the value that the individual places on a marginal reduction in risk of death in the current period. For example, it is the amount that 1,000 people would be collectively willing to pay to eliminate a current risk that is expected to kill one of them. It is equal to the present discounted value of lifetime consumption, plus the change in net savings. Holding wealth constant, VSL increases with survival, which implies increasing returns in health improvements (Murphy and Topel 2006). Conversely, this leads to the conventional wisdom that VSL falls when mortality rises.

The value of statistical life depends on how substitutable consumption is at different ages, i.e., on how easily an individual can reallocate consumption over time. Intuitively, if present consumption is a good substitute for future consumption, then living longer is less valuable. Define the elasticity of intertemporal substitution, σ, as:

$$\frac{1}{\sigma} \equiv -\frac{u_{cc}c}{u_c}$$

In addition, define the elasticity of quality of life with respect to the marginal utility of consumption as:

$$\eta \equiv \frac{u_{cq}q}{u_c}$$

When this term is positive, the marginal utility of consumption is higher in healthier states, and vice-versa. Taking logarithms of the first-order condition for consumption and differentiating with respect to time yields the rate of change for consumption over the life cycle:

$$\frac{\dot{c}}{c} = \sigma(r - \rho) + \sigma\eta\frac{\dot{q}}{q} \tag{4}$$

[4] Positive consumer surplus may require that consumption remain above a "subsistence" level, $\underline{c} > 0$.

If one assumes that $r > \rho$, and that the marginal utility of consumption is higher when health status is better, then life-cycle consumption will have the inverted U-shape observed in real-world data.[5]

Note the crucial feature of the conventional model that consumption growth over the life-cycle is independent of mortality risk, because the individual is fully insured against that risk. This feature in turn implies that the rate of change in the value of a life-year is also not a function of mortality risk:

$$\frac{\dot{v}}{v} = \left(\frac{1}{\sigma v}\frac{u}{u_c}\right)\frac{\dot{c}}{c} + \left(\frac{-\eta}{v}\frac{u}{u_c} + \frac{q}{v}\frac{u_q}{u_c}\right)\frac{\dot{q}}{q} + \frac{\dot{m}}{v}$$

In sum, we have identified two major features of the conventional, fully annuitized and deterministic model of mortality:

- The relative value of a life-year within a lifetime is independent of mortality risk.
- The value of statistical life falls when mortality rises.

II.B. The uninsured value of life

To illustrate the effects of annuitization, we consider a model without any annuitization possibilities. In our numerical exercises later, we will consider various partial annuitization schemes. To characterize the model without annuitization, we employ the Yaari (1965) model of consumption behavior under mortality risk. The consumer's maximization problem is:

$$\begin{aligned} V\big(0, W(0)\big) &= \max_{c(t)} \int_0^T e^{-\rho t} S(t) u(c(t), q(t)) dt \\ s.t.\, W(0) &= W_0, \\ W(t) &\geq 0, W(T) = 0, \\ \frac{\partial W(t)}{\partial t} &= rW(t) + m(t) - c(t) \end{aligned}$$

If the non-negative wealth constraint binds, then the solution to the consumer's problem is to set $c(t) = m(t)$. Otherwise, the solution is to maximize subject to the constraint on the law of motion for wealth. We focus here on the latter, nontrivial case.

Optimal consumption is again characterized by the first-order condition:

$$\frac{\partial V\big(0, W(0)\big)}{\partial W(0)} = \theta = e^{(r-\rho)t} S(t) u_c(c(t), q(t))$$

Unlike in the case of perfect markets, the survival function enters the consumer's first-order condition for optimal consumption. Instead of setting the discounted marginal utility of consumption equal to the marginal utility of wealth, the consumer sets the *expected* discounted marginal utility of consumption at time t equal to the marginal utility of wealth. This effectively shifts consumption to earlier ages in the life-cycle. This is rational because consumption allocated to later time periods will not be enjoyed in the event of an early death.

The expression for the marginal utility of life-extension is:

[5] Consumption climbs early in life as the benefits to savings diminish. It declines later in life when quality of life deteriorates. This inverted U-shape for the age profile of consumption has been widely documented across different countries and goods (Carroll and Summers 1991; Banks et al. 1998; Fernandez-Villaverde and Krueger 2007).

$$
\begin{aligned}
\left.\frac{\partial V}{\partial \varepsilon}\right|_{\varepsilon=0} &= \left.\frac{\partial}{\partial \varepsilon}\int_0^T e^{-\rho t} S^{\varepsilon}(t) u\big(c^{\varepsilon}(t), q(t)\big) dt\right|_{\varepsilon=0} \\
&= \int_0^T e^{-\rho t}\left[\int_0^t \delta(s) ds\right] S(t) u(c(t), q(t)) dt + \int_0^T e^{-\rho t} S(t) u_c(c(t), q(t)) \left.\frac{\partial c^{\varepsilon}(t)}{\partial \varepsilon}\right|_{\varepsilon=0} dt \\
&= \int_0^T e^{-\rho t}\left[\int_0^t \delta(s) ds\right] S(t) u(c(t), q(t)) dt + \theta \frac{\partial}{\partial \varepsilon}\int_0^T e^{-rt} c^{\varepsilon}(t) dt \\
&= \int_0^T e^{-\rho t}\left[\int_0^t \delta(s) ds\right] S(t) u(c(t), q(t)) dt,
\end{aligned}
$$

where the last equality follows from application of the budget constraint.[6]

Dividing this result by the marginal utility of wealth, θ, then yields the marginal value of life-extension:

$$
\begin{aligned}
\frac{\partial V/\partial \varepsilon}{\partial V/\partial W} &= \int_0^T e^{-\rho t}\left[\int_0^t \delta(s) ds\right] S(t) \frac{u\big(c(t), q(t)\big)}{u_c\big(c(0), q(0)\big)} dt \\
&= \int_0^T e^{-rt}\left[\int_0^t \delta(s) ds\right] \frac{u\big(c(t), q(t)\big)}{u_c\big(c(t), q(t)\big)} dt
\end{aligned} \tag{5}
$$

In this setting, the value of a life-year from the perspective of current time is:

$$
v(t) = \frac{u\big(c(t), q(t)\big)}{u_c\big(c(t), q(t)\big)} \tag{6}
$$

When the consumer is uninsured, the value of a life-year depends only on the value of consumption. The net savings term is absent in equation (6) because life-extension has no effect on the consumer's budget constraint.[7]

Choosing again the Dirac delta function for $\delta(\cdot)$ yields an expression for VSL that differs from the perfect markets case:

$$
VSL = \int_0^T e^{-rt} v(t) dt \tag{7}
$$

The value of statistical life is proportional to (expected) lifetime utility, and inversely proportional to the marginal utility of consumption. It is well known that removing annuity markets lowers lifetime utility (Yaari 1965). As we show more formally below, removing these markets also shifts consumption to earlier ages, thereby lowering the marginal utility of consumption, at least at those ages. When consumers shift consumption forward, the near-term life-years rise in value but distant life-years fall in value. Thus, the net effect of annuity markets on VSL is in general ambiguous. Put differently, exposure to longevity risk does not necessarily lower VSL. In the next section, we will show that this basic insight extends to

---

[6] The budget constraint $W(T) = 0$ implies $\int_0^T e^{-rt} c^{\varepsilon}(t) dt = W_0 + \int_0^T e^{-rt} m(t) dt$, a value which does not depend on survival and thus is unaffected by life extension.

[7] Unless the consumer survives until period T, she will die with positive wealth. Although this remaining wealth has no value to an individual with no bequest motive, it may be of value to society. When calculating the *social* value of life-extension, we account for the effect of increased longevity on bequests by including a net savings term, defined to be the expected increase in future earnings net of consumption, as in equation (2). This term reflects the external effect on society's aggregate wealth due to increased longevity.

exposing a consumer to a mortality "shock." We emphasize that in both cases the result depends critically on whether consumers are fully annuitized.

Unlike the perfect markets case, the life-cycle consumption profile of the non-annuitized individual depends explicitly on mortality risk. Taking logarithms of the first-order condition for consumption and differentiating with respect to time yields:

$$\frac{\dot{c}}{c} = \sigma(r-\rho) + \sigma\eta\frac{\dot{q}}{q} - \sigma\mu(t) \tag{8}$$

Comparing this result to the standard case, given by equation (4), reveals both similarities and differences. As in the standard, fully annuitized model, the non-annuitized consumption profile described by equation (8) changes shape when the rate of time preference is above or below the rate of interest and when the quality of life changes. Unlike in the standard model, the consumption profile described by equation (8) depends explicitly on the mortality rate, $\mu(t)$. Higher rates of mortality depress the rate of consumption growth over the life-cycle. This rate of growth is always higher in the fully annuitized case, in which the last term drops out of the consumption growth equation (8). Put another way, removing the annuity market "pulls consumption earlier" in the life-cycle.

An appealing feature of the uninsured model is that it generates an inverted U-shape for the profile of consumption under quite natural assumptions. Low income early in life and high mortality risk later in life are sufficient conditions for the inverted U-shape consumption profile. One need not impose the ad hoc assumptions on the signs of $r-\rho$ or η that are necessary in the fully annuitized model (Murphy and Topel 2006).

The life-cycle profile of the value of a life-year is:

$$\frac{\dot{v}}{v} = \left(\frac{1}{\sigma} + \frac{c}{v}\right)\frac{\dot{c}}{c} + \left(\frac{qu_q}{u} - \eta\right)\frac{\dot{q}}{q} \tag{9}$$

An important implication of (9) is that willingness to pay for longevity depends on the life-cycle mortality profile because of its dependence on the rate of change in consumption. Holding quality of life constant, it is evident from equation (6) that increases in the mortality rate—which shift consumption forward—will raise v, the current value of a life-year. That is, mortality also shifts forward the value of life. All else equal, individuals who face high mortality risks will pay more for a marginal (near-term) life-year, but less for a distant life-year, than healthy peers who face low mortality risks. This differs from the implications of the conventional model, in which higher mortality reduces the values of life-years but has no impact on their relative values.

At the aggregate level, as societies become richer and live longer, the fraction of wealth spent on health will depend not just on the income elasticity of health, but also on the degree of survival uncertainty they face. Furthermore, our results imply that public programs that increase annuitization rates, such as Social Security, will affect society's willingness to pay for longevity, thereby creating a feedback loop that could dampen or increase program expenditures.[8] In our numerical exercises, we will quantify how the degree of annuitization influences the value of statistical life.

To summarize the findings for this uninsured model, we have identified the following two properties that contrast with those of the fully annuitized model:

[8] Philipson and Becker (1998) make the important, but distinct, point that the moral hazard effects of public annuity programs also increase an individual's willingness to pay for longevity gains.

- The values of near-term life-years rise, and distant life-years fall, when mortality rises.
- The value of statistical life may rise or fall when mortality rises.

In the next section, we allow mortality to be stochastic so that we can investigate formally the effect of disease and other health shocks on the value of life. Before turning to that analysis, we pause to note that suffering a health shock is similar to removing access to annuity markets, which exposes an individual to mortality risk. We have shown here that this shifts the value of life-years forward, with an ambiguous net effect on VSL. As we shall see, health shocks have a similar effect.

III. THE VALUE OF LIFE WHEN MORTALITY IS STOCHASTIC

The previous analysis demonstrates that mortality risk affects the value of life when annuity markets are incomplete. Prior studies have overlooked this relationship by assuming complete annuitization. However, the conventional framework is ill-equipped to study the influence of mortality risk for another reason as well. Prior analysis, just like our deterministic model above, treats the mortality rate as a nonrandom parameter (Murphy and Topel, 2006). Thus, shifts in mortality risk reflect preordained and anticipated changes in mortality. In the real world, however, neither the timing nor the size of shifts in mortality risk is known. As a related matter, the conventional framework does not allow for different health states. This omission precludes a meaningful analysis of the value of preventing health deterioration.

This section extends our analysis to allow for stochastic mortality. Specifically, we assume that the mortality rate now depends on the individual's health state. Let Y_t be a continuous-time Markov chain with finite state space $Y = \{1,2,\dots,n\}$. Denote the transition intensities by:

$$\lambda_{ij}(t) = \lim_{h\to 0}\frac{1}{h}\mathbb{P}[Y_{t+h} = j | Y_t = i], j \neq i,$$

$$\lambda_{ii}(t) = -\sum_{j\neq i}\lambda_{ij}(t)$$

The mortality rate at time t is defined as

$$\mu(t) = \sum_{j=1}^{n}\overline{\mu}_j(t)\mathbf{1}\{Y_t = j\}$$

where $\left\{\overline{\mu}_j(t)\right\}$ are exogenous and $\mathbf{1}\{Y_t = j\}$ is an indicator variable equal to 1 if the individual is in state j at time t and 0 otherwise. Without meaningful loss of generality, we assume that individuals can transition only to higher-numbered states, i.e., $\lambda_{ij}(t) = 0\ \forall j < i$, so that the probability that a consumer in state i at time 0 remains in state i at time t is equal to:[9]

$$\tilde{S}(i,t) = \exp\left[-\int_0^t\left(\overline{\mu}_i(s) + \sum_{j\neq i}\lambda_{ij}(s)\right)ds\right]$$

[9] That is, an individual can transition from state i to j, $i < j$, but not vice versa. This does not meaningfully limit the generality of our model, because one can always define a new state $k > j$ where $\overline{\mu}_k(t) = \overline{\mu}_i(t)\forall t$.

A complete annuities market allows the consumer to insure fully against mortality risk even when mortality is stochastic.[10] Appendix C provides a full derivation for a setting with complete markets and demonstrates that stochastic mortality, by itself, does not alter the basic insights regarding VSL offered by the prior literature as long as one maintains the assumption of full annuitization. Appendix C also derives expressions for the value of preventing illness when the consumer is fully annuitized. We defer discussion of those results until later in this section.

Here, we focus on the uninsured case. The consumer's maximization problem is:

$$V(0, W(0), Y_0) = \max_{c_{Y_t}(t)} \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) u\big(c_{Y_t}(t), q_{Y_t}(t)\big) dt \,\middle|\, Y_0\right] \tag{10}$$

$$\begin{aligned} s.t.\, & W(0) = W_0, \\ & W(t) \geq 0, W(T) = 0, \\ & \frac{\partial W(t)}{\partial t} = rW(t) + m_{Y_t}(t) - c_{Y_t}(t) \end{aligned}$$

As in the deterministic model presented in Section II.B, we focus on the non-trivial case where the non-negative wealth constraint does not bind. Define the consumer's objective function at time u as:

$$J(u, i) = \mathbb{E}\left[\int_0^{T-u} e^{-\rho t} \exp\left\{-\int_0^t \mu(u+s) ds\right\} u\big(c_{Y_{u+t}}(u+t), q_{Y_{u+t}}(u+t)\big) dt \,\middle|\, Y_u = i\right] \tag{11}$$

We can then write the objective function recursively as:

$$J(u, i) = \int_0^{T-u} e^{-\rho t} \exp\left\{-\int_0^t \left(\bar{\mu}_i(u+s) + \sum_{j \neq i} \lambda_{ij}(u+s)\right) ds\right\} \left(u(c_i(u+t), q_i(u+t)) + \sum_{j \neq i} \lambda_{ij}(u+t) J(u+t, j)\right) dt$$

Define the optimal value function as:

$$V(t, W(t), i) = \max_{c.(s), s \geq t} \{J(t, i)\}$$

Under conventional regularity conditions, we know that if V and its partial derivatives are continuous, then V satisfies the following Hamilton-Jacobi-Bellman (HJB) system of equations:

$$\begin{aligned} \big(\rho + \bar{\mu}_i(t)\big) V(t, W(t), i) \\ = \max_{c_i(t)} \Bigg\{ & u\big(c_i(t), q_i(t)\big) + \frac{\partial V(t, W(t), i)}{\partial W(t)} [rW(t) + m_i(t) - c_i(t)] \\ & + \frac{\partial V(t, W(t), i)}{\partial t} + \sum_{j \neq i} \lambda_{ij}(t) [V(t, W(t), j) - V(t, W(t), i)] \Bigg\}, i = 1, \dots, n \end{aligned} \tag{12}$$

We are interested in understanding how optimal consumption, and thus the value of life, changes over the life-cycle in this problem. We follow Parpas and Webster (2013), who demonstrate that it is possible to

[10] Reichling and Smetters (2015) show that when annuity markets are incomplete, stochastic mortality and correlated medical costs can explain the puzzling observation that many households do not fully annuitize their wealth. They take the positive correlation between health shocks and medical spending as a given. Our study sheds light on *why* these two phenomena are positively correlated.

reformulate a stochastic optimization problem as a deterministic problem that takes $V(t, W(t), j), j \neq i$, as exogenous. This then allows us to apply the Pontryagin maximum principle and derive analytic expressions.

Lemma 1:

The optimal value function for $Y_0 = i, V(0, W(0), i)$, for the following deterministic optimization problem also satisfies the HJB given by (12), for each $i \in \{1, \dots, n\}$:

$$V(0, W_0, i) = \max_{c_i(t)} \left[\int_0^T e^{-\rho t} \tilde{S}(i,t) \left(u(c_i(t), q_i(t)) + \sum_{j \neq i} \lambda_{ij}(t) V(t, W(t), j) \right) dt \right] \tag{13}$$

$$s.t. \frac{\partial W(t)}{\partial t} = rW(t) + m_i(t) - c_i(t)$$

where $V(t, W(t), j)$ are taken as exogenous.

Proof of Lemma 1: see Appendix A

Following Bertsekas (2005), the present value Hamiltonian corresponding to (13) is

$$H\big(W(t), c_i(t), p_t^{(i)}\big) = e^{-\rho t} \tilde{S}(i,t) \left(u\big(c_i(t), q_i(t)\big) + \sum_{j \neq i} \lambda_{ij}(t) V(t, W(t), j) \right) + p_t^{(i)} [rW(t) - c_i(t) + m_i(t)]$$

where $p_t^{(i)}$ is the costate variable for state i. The necessary costate equation is:

$$\dot{p}_t^{(i)} = -p_t^{(i)} r - e^{-\rho t} \tilde{S}(i,t) \sum_{j \neq i} \lambda_{ij}(t) \frac{\partial V(t, W(t), j)}{\partial W(t)}$$

The solution to the costate equation can be obtained using the variation of the constant method:

$$p_t^{(i)} = \left[\int_t^T e^{(r-\rho)s} \tilde{S}(i,s) \sum_{j \neq i} \lambda_{ij}(s) \frac{\partial V(s, W(s), j)}{\partial W(s)} ds \right] e^{-rt} + \theta^{(i)} e^{-rt}$$

where $\theta^{(i)}$ is a constant. The necessary first-order condition for consumption is:

$$p_t^{(i)} = e^{-\rho t} \tilde{S}(i,t) u_c\big(c_i(t), q_i(t)\big) \tag{14}$$

where the marginal utility of wealth at time $t = 0$ is $\frac{\partial V(0, W_0, i)}{\partial W_0} = p_0^{(i)} = u_c\big(c_i(0), q_i(0)\big)$. Since the Hamiltonian is concave in c and linear in W, the necessary conditions for optimality are also sufficient (Seierstad and Sydsaeter 1977).

To analyze the value of life, we let $\delta(t)$ be a perturbation on the mortality rate in state i with $\int_0^T \delta(t) dt = 1$ and consider

$$\tilde{S}^{\varepsilon}(i,t) = \exp\left[-\int_0^t \big(\overline{\mu}_i(s) - \varepsilon\delta(s)\big) + \sum_{j \neq i} \lambda_{ij}(s)\, ds \right], \text{where } \varepsilon > 0$$

We first derive an expression for the effect of this perturbation on expected lifetime utility.

Lemma 2:

The marginal utility of life extension in state i is equal to:

$$\left.\frac{\partial V}{\partial \varepsilon}\right|_{\varepsilon=0} = \int_0^T \left[e^{-\rho t} \left(\int_0^t \delta(s) ds \right) \tilde{S}(i,t) \left(u(c_i(t), q_i(t)) + \sum_{j \neq i} \lambda_{ij}(t)\, V(t, W(t), j) \right) \right] dt$$

Proof of Lemma 2:

From (13), the marginal utility of life-extension is

$$\begin{aligned}
\left.\frac{\partial V}{\partial \varepsilon}\right|_{\varepsilon=0} &= \frac{\partial}{\partial \varepsilon} \int_0^T e^{-\rho t} \exp\left\{ -\int_0^t \big(\mu(s) - \varepsilon\delta(s)\big) + \sum_{j \neq i} \lambda_{ij}(s)\, ds \right\} \Bigg(u\big(c_i^{\varepsilon}(t), q_i(t)\big) \\
&\qquad + \sum_{j \neq i} \lambda_{ij}(t) V\big(t, W^{\varepsilon}(t), j\big) \Bigg) dt \Bigg|_{\varepsilon=0} \\
&= \int_0^T e^{-\rho t} \left(\int_0^t \delta(s) ds \right) \tilde{S}(i,t) \left(u\big(c_i(t), q_i(t)\big) + \sum_{j \neq i} \lambda_{ij}(t) V(t, W(t), j) \right) dt \\
&\qquad + \int_0^T e^{-\rho t} \tilde{S}(i,t) \left(u_c\left(c_i(t), q_i(t)\right) \frac{\partial c_i^{\varepsilon}(t)}{\partial \varepsilon} + \sum_{j \neq i} \lambda_{ij}(t) \frac{\partial V(t, W(t), j)}{\partial W} \frac{\partial W^{\varepsilon}(t)}{\partial \varepsilon} \right) dt \Bigg|_{\varepsilon=0}
\end{aligned}$$

where $c_i^{\varepsilon}(t)$ and $W^{\varepsilon}(t)$ represent the equilibrium variations in $c_i(t)$ and $W(t)$ caused by this perturbation. We conclude the proof by showing that the second term in the last equality is equal to 0. Note that along this path, wealth at time t is equal to

$$W(t) = W_0 e^{rt} + \int_0^t e^{r(t-s)} m_i(s) ds - \int_0^t e^{r(t-s)} c_i(s) ds,$$

which implies $\frac{\partial W^{\varepsilon}(t)}{\partial \varepsilon} = -\int_0^t e^{r(t-s)} \frac{\partial c_i^{\varepsilon}(s)}{\partial \varepsilon} ds$. From the solution to the costate equation, we know that

$$e^{-\rho t} \tilde{S}(i,t) u_c\big(c_i(t), q_i(t)\big) = \left[\int_t^T e^{(r-\rho)s} \tilde{S}(i,s) \sum_{j \neq i} \lambda_{ij}(s) \frac{\partial V(s, W(s), j)}{\partial W(s)} ds \right] e^{-rt} + \theta^{(i)} e^{-rt}$$

Thus, we can rewrite the second term in the expression for $\left.\frac{\partial V}{\partial \varepsilon}\right|_{\varepsilon=0}$ above as

$$\begin{aligned}
&\int_0^T \left[\int_t^T e^{(r-\rho)s} \tilde{S}(i,s) \sum_{j \neq i} \lambda_{ij}(s) \frac{\partial V(s, W(s), j)}{\partial W(s)} ds + \theta^{(i)} \right] e^{-rt} \frac{\partial c_i^{\varepsilon}(t)}{\partial \varepsilon} dt \\
&\qquad - \int_0^T e^{-\rho t} \tilde{S}(i,t) \sum_{j \neq i} \lambda_{ij}(t) \frac{\partial V(t, W(t), j)}{\partial W} \int_0^t e^{r(t-s)} \frac{\partial c_i^{\varepsilon}(s)}{\partial \varepsilon} ds\, dt \Bigg|_{\varepsilon=0}
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^T \left[\int_t^T e^{(r-\rho)s} \tilde{S}(i,s) \sum_{j\neq i} \lambda_{ij}(s) \frac{\partial V(s,W(s),j)}{\partial W(s)} ds \right] e^{-rt} \frac{\partial c_i^\varepsilon(t)}{\partial \varepsilon} dt \\
&\qquad - \int_0^T \left[\int_t^T e^{(r-\rho)s} \tilde{S}(i,s) \sum_{j\neq i} \lambda_{ij}(s) \frac{\partial V(s,W(s),j)}{\partial W(s)} ds \right] e^{-rt} \frac{\partial c_i^\varepsilon(t)}{\partial \varepsilon} dt + \int_0^T \theta^{(i)} e^{-rt} \frac{\partial c_i^\varepsilon(t)}{\partial \varepsilon} dt \Bigg|_{\varepsilon=0} \\
&= \theta^{(i)} \frac{\partial}{\partial \varepsilon} \underbrace{\int_0^T e^{-rt} c_i^\varepsilon(t) dt}_{W_0 + \int_0^T e^{-rt} m_i(t) dt} \Bigg|_{\varepsilon=0} \\
&= 0
\end{aligned}
$$

QED

In order to facilitate comparison to the deterministic case, it is useful to derive an expression for the marginal utility of wealth at time t.

Lemma 3:

The expected marginal utility of wealth in state i at time t is equal to:

$$
\frac{\partial V(t,W(t),i)}{\partial W(t)} = u_c\big(c_i(t), q_i(t)\big) = \mathbb{E}\left[e^{(r-\rho)(\tau-t)} \exp\left\{ -\int_t^\tau \mu(s) ds \right\} u_c\big(c_{Y_\tau}(\tau), q_{Y_\tau}(\tau)\big) \middle| Y_t = i \right]
$$

Proof of Lemma 3: see Appendix A

Our next result demonstrates that the value of statistical life takes the same basic form as in the deterministic case.

Proposition 4:

Choosing once again the Dirac delta function for $\delta(\cdot)$ simplifies the expression for the marginal utility of life-extension:

$$
\begin{aligned}
\frac{\partial V}{\partial \varepsilon}\bigg|_{\varepsilon=0} &= \int_0^T \left[e^{-\rho t} \tilde{S}(i,t) \left(u\big(c_i(t), q_i(t)\big) + \sum_{j\neq i} \lambda_{ij}(t)\, V(t,W(t),j) \right) \right] dt \\
&= \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) u\big(c_{y_t}(t), q_{y_t}(t)\big) dt \middle| Y_0 = i \right]
\end{aligned}
$$

Dividing the result by the marginal utility of wealth at time $t = 0$ and then applying **Lemma 3** shows that the value of statistical life takes the same basic form as in the deterministic case:

$$
VSL(i) = \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) \frac{u\big(c_{y_t}(t), q_{y_t}(t)\big)}{u_c\big(c_{Y_0}(0), q_{Y_0}(0)\big)} dt \middle| Y_0 = i \right] = \int_0^T e^{-rt} v(i,t) dt \tag{15}
$$

where the value of a statistical life-year is equal to the expected utility of consumption normalized by the expected marginal utility of consumption:

$$
v(i,t) = \frac{\mathbb{E}\left[S(t) u\left(c_{y_t}(t), q_{y_t}(t)\right) \middle| Y_0 = i \right]}{\mathbb{E}\left[S(t) u_c\left(c_{y_t}(t), q_{y_t}(t)\right) \middle| Y_0 = i \right]}
$$

Proof of Proposition 4: see Appendix A

As before, the value of statistical life is proportional to the expected discounted (lifetime) utility of consumption, and inversely proportional to the marginal utility of consumption. As we shall show below, a negative health shock increases current consumption, causing the net effect on VSL to be ambiguous. This parallels the result we showed previously that removing access to annuitization, thereby exposing a consumer to mortality risk, has an ambiguous effect on VSL.

We can derive an expression for the life-cycle profile of consumption from (14), the first-order condition for p_t. Differentiating with respect to t, plugging in the result for the costate equation and its solution, and rearranging yields

$$\frac{\dot{c}_i}{c_i} = \sigma(r-\rho) + \sigma\eta\frac{\dot{q}}{q} - \sigma\overline{\mu}_i(t) - \sigma\sum_{j\neq i}\lambda_{ij}(t)\left[1 - \frac{u_c\left(c_j(t), q_j(t)\right)}{u_c\left(c_i(t), q_i(t)\right)}\right] \tag{16}$$

As in the deterministic case, the rate of change is a declining function of the individual's current mortality rate, $\overline{\mu}_i(t)$: removing the annuity market "pulls consumption earlier" in the life-cycle. Unlike in the deterministic case, there is now an additional source of mortality risk, captured by the fourth term in equation (16). This term represents the possibility that the consumer might transition to a different health state in the future, and shifts consumption further still if the consumer is likely to fall ill in the future.

We caution that equation (16) is specific to an individual's health state i, and cannot be easily aggregated across health states. That is, one cannot infer from equation (16) whether stochastic mortality *on average* causes consumption to shift forward relative to deterministic mortality. That said, one should expect stochastic mortality to shift consumption forward by less than in the deterministic case. Intuitively, this is because a stochastic environment allows an individual to react to unanticipated health shocks by adjusting her consumption. Put differently, a deterministic model is equivalent to a stochastic model where the consumer is forced to keep consumption constant across states. Consumers prefer the ability to adjust consumption, so that they can consume less in healthy states and more in sick states. We have confirmed this intuition in (unreported) empirical exercises that assume CRRA utility: on net, stochastic mortality causes consumers to shift consumption forward a bit less than deterministic mortality.

What happens when an individual transitions to a new health state? Because the consumer is not insured against mortality or quality of life risks, consumption will jump. The sign of the jump can be positive or negative, depending on the characteristics of the new health state relative to the old state. Because there is no consensus regarding the sign of health state dependence ($u_{cq}(\cdot)$), let alone the magnitude, we hold quality of life constant for the time being, and return to this issue in our empirical analysis.[11] Focusing on mortality, the model predicts that transitioning to a state where the current mortality and future expected mortality are high will shift consumption forward (see Figure 6), and vice versa. Our next result proves this formally for a two-state case.[12]

[11] Finkelstein et al. (2013), Sloan et al. (1998), and Viscusi and Evans (1990) find evidence of negative state dependence. Edwards (2008) and Lillard and Weiss (1988) find evidence of positive state dependence. Evans and Viscusi (1991) find no evidence of state dependence. Murphy and Topel (2006) assume negative state dependence when performing their calibration exercises, while Hall and Jones (2007) assume state independence.

[12] The proof can be extended to allow for a larger number of states, but the conditions required to sign the jump in consumption then become a complicated function of the matrix of transition probabilities and state-specific mortality rates. The two-state case conveys the basic result without a meaningful loss of generality.

Proposition 5:

Let there be $n = 2$ states with identical quality of life profiles, so that $q_1(s) = q_2(s)\ \forall s$. Assume that $\overline{\mu}_1(s) < \overline{\mu}_2(s) \forall s$, so that state 1 is "healthy" and state 2 is "sick." Suppose that the consumer transitions from state 1 to state 2 at time t, with no accompanying decrease in income (i.e., $m_1(t) \leq m_2(t)$). Then $c_1(t) < c_2(t)$.

Proof of Proposition 5: see Appendix A

It follows immediately from **Proposition 5** that the value of near-term life-years will increase, and the value of distant life-years will decrease, when transitioning from a healthy state with low mortality to a sick state with higher mortality. Whether VSL rises or falls is ambiguous, however. A rise in mortality risk lowers lifetime utility, which reduces VSL, but it also reduces the marginal utility of consumption, which increases VSL. Thus, the net effect depends on the curvature of the utility function relative to the curvature of the marginal utility function. The elasticity of intertemporal substitution, σ, is a common measure of the utility curvature. The analogous measure for the curvature of marginal utility is prudence (Kimball 1990). Define relative prudence as

$$\pi \equiv -\frac{c u_{ccc}(\cdot)}{u_{cc}(\cdot)}$$

Our next result provides a sufficient condition for VSL to rise following an adverse mortality shock.

Proposition 6:

Consider a two-state setting with assumptions set out in **Proposition 5**. Assume further that preferences satisfy the additional condition

$$\pi < \frac{2}{\sigma}$$

Suppose that the consumer transitions from state 1 to state 2 at time t, and that $\lambda_{12}(\tau) = 0 \forall \tau > t$. Then $VSL(1,t) < VSL(2,t)$.

Proof of Proposition 6: see Appendix A

The condition specified in **Proposition 6** is satisfied by many common preferences, such as CRRA with $\sigma < 1$ (which we employ in our numerical exercises) and quadratic preferences. Consumers with inelastic demand, i.e., preferences with a low value for σ, find it costly to reallocate consumption over time. They therefore have a high willingness-to-pay for life-extension and are more likely to exhibit a rise in VSL following an adverse mortality shock. Likewise, consumers with low levels of prudence have nearly-linear marginal utility that decreases rapidly with consumption. This generates a high willingness-to-pay for life-extension following a shock that increases consumption.

III.A. The value of statistical illness

Unlike the deterministic model, the stochastic model permits an investigation not only into the value of preventing death, but also into the value of preventing transitions to other health states. This requires only a slight modification to the analysis presented above, and will result in a more general concept we term the *value of statistical illness*. With a slight abuse of notation, let state $N + 1$ correspond to death, so that $V(t, W(t), N + 1) = 0$. Let $\delta_{ij}(t),\ i, j \leq N$, be a perturbation on the transition intensity $\lambda_{ij}(t)$ and $\delta_{i,N+1}(t)$ be a perturbation on the mortality rate $\overline{\mu}_i(t)$, where $\sum_{j=1, j\neq i}^{N+1} \int_0^T \delta_{ij}(t)dt = 1$, and consider

$$\tilde{S}^{\varepsilon}(i,t) = \exp\left[-\int_0^t \left(\overline{\mu}_i(s) - \varepsilon\delta_{i,N+1}(s)\right) + \sum_{j=1,j\neq i}^{N} \left(\lambda_{ij}(s) - \varepsilon\delta_{ij}(s)\right) ds\right], \text{where } \varepsilon > 0$$

Proposition 7:

The marginal utility of preventing an illness or death is given by:

$$\left.\frac{\partial V}{\partial \varepsilon}\right|_{\varepsilon=0} = \int_0^T e^{-\rho t}\tilde{S}(i,t)\left[\int_0^t \sum_{j\neq i} \delta_{ij}(s)\, ds \left(u(c_i(t), q_i(t)) + \sum_{j\neq i} \lambda_{ij}(t)\, V(t, W(t), j)\right) - \sum_{j\neq i} \delta_{ij}(t)\, V(t, W(t), j)\right] dt$$

Proof of Proposition 7:

From (13), the marginal utility of preventing an illness or death is:

$$\begin{aligned}\left.\frac{\partial V}{\partial \varepsilon}\right|_{\varepsilon=0} &= \frac{\partial}{\partial \varepsilon}\int_0^T e^{-\rho t} \exp\left\{-\int_0^t \left(\overline{\mu}_i(s) - \varepsilon\delta_{i,N+1}(t)\right) + \sum_{j\neq i}\left(\lambda_{ij}(s) - \varepsilon\delta_{ij}(s)\right) ds\right\}\left(u\left(c_i^{\varepsilon}(t), q_i(t)\right)\right. \\ &\qquad \left.\left. + \sum_{j\neq i}\left(\lambda_{ij}(t) - \varepsilon\delta_{ij}(t)\right) V(t, W^{\varepsilon}(t), j)\right) dt\right|_{\varepsilon=0} \\ &= \int_0^T e^{-\rho t}\tilde{S}(i,t)\left[\left(\int_0^t \sum_{j\neq i}\delta_{ij}(s)\, ds\right)\left(u(c_i(t), q_i(t)) + \sum_{j\neq i}\lambda_{ij}(t)\, V(t, W(t), j)\right) - \sum_{j\neq i}\delta_{ij}(t)\, V(t, W(t), j)\right] dt \\ &\qquad + \int_0^T e^{-\rho t}\tilde{S}(i,t)\left(u_c\left(c_i^{\varepsilon}(t), q_i(t)\right)\frac{\partial c_i^{\varepsilon}(t)}{\partial \varepsilon} + \sum_{j\neq i}\lambda_{ij}(t)\frac{\partial V(t, W(t), j)}{\partial W}\frac{\partial W^{\varepsilon}(t)}{\partial \varepsilon}\right) dt\end{aligned}$$

Following the same argument as in the VSL case, the second term in the last equality is equal to 0.

QED

The value of preventing an illness or death is equal to the marginal rate of substitution between the transition perturbation and wealth:

$$\frac{\partial V/\partial \varepsilon}{\partial V/\partial W} = \int_0^T \frac{e^{-\rho t}\tilde{S}(i,t)}{u_c\left(c_i(0), q_i(0)\right)}\left[\left(\int_0^t \sum_{j\neq i}\delta_{ij}(s)\, ds\right)\left(u(c_i(t), q_i(t)) + \sum_{j\neq i}\lambda_{ij}(t)\, V(t, W(t), j)\right) - \sum_{j\neq i}\delta_{ij}(t)\, V(t, W(t), j)\right] dt$$

As before, it is helpful to choose the Dirac delta function for $\delta(\cdot)$, so that the probability is perturbed at $t = 0$ and remains unaffected otherwise. It is also helpful to consider a reduction in the transition probability for only one alternative state, j_0, so that $\delta_{ij}(t) = 0\ \forall j \neq j_0$. Applying these two conditions then yields what we term the value of statistical illness, $VSI(i,j)$:

$$\begin{aligned}VSI(i,j) &= \frac{V(0, W(0), i) - V(0, W(0), j)}{u_c\left(c_i(0), q_i(0)\right)} \\ &= VSL(i) - VSL(j)\frac{u_c\left(c_j(0), q_j(0)\right)}{u_c\left(c_i(0), q_i(0)\right)}\end{aligned} \tag{17}$$

The interpretation of VSI is analogous to VSL: it is the amount that 1,000 individuals would collectively be willing to pay in order to eliminate a current disease risk that is expected to befall one of them. Note that if health state j corresponds to death, so that $VSL(j) = VSL(N + 1) = 0$, then $VSI(i, j) = VSL(i)$. Thus, VSI is a generalization of VSL.

It is instructive to compare (17) to the expression for VSI obtained when the consumer is fully annuitized (derivation available in Appendix C):

$$VSI^*(i, j) = VSL^*(i) - VSL^*(j) \tag{18}$$

Equation (18) provides justification for the common practice of equating the values of prevention and treatment.[13] Conventional cost-effectiveness analysis relies upon the standard fully annuitized framework that assumes the value of a life-year is equal across health states (holding quality of life constant). If the value of a life-year is constant, then equation (18) implies that prevention and treatment are equally valuable, as long as they add the same number of expected life-years. For example, conventional cost-effectiveness frameworks value a treatment that prevents the onset of an illness that lowers life expectancy by 10 years the same as a therapeutic treatment that cures an illness and adds 10 years of life expectancy (Drummond et al. 2005b).

In contrast, equation (17) shows that removing access to annuity markets breaks this equivalence between treatment and prevention. VSI in this case is not equal to the simple difference in VSL between the healthy and sick states, because VSL in the sick state is valued from the perspective of the sick, who have a lower marginal utility of consumption due to a shorter life span. This leads to the natural hypothesis that whenever VSL rises following an illness, the value of treatments (VSL per life-year) will be higher than equivalent preventive care prior to the illness (VSI per life-year). It is simple to show this for the case where the illness reduces life expectancy by one-half or more (proof available upon request). We conjecture that the hypothesis is true for any illness that reduces life expectancy.

To summarize, the stochastic mortality model yields the following implications:

- The values of near-term life-years rise, and distant life-years fall, when an individual transitions to a higher mortality state.
- The value of statistical life may rise or fall when an individual transitions to a higher mortality state; if the individual’s demand is sufficiently inelastic, or insufficiently prudent, , then it will rise.
- Therapies that increase survival by treating sick patients are not the same as, and may even be more valuable than, those that add the same amount of life expectancy by preventing illness in healthy patients.

IV. ESTIMATES OF THE VALUE OF LIFE

IV.A. Framework

We will work with the discrete time analogue of our model. There are n health states. Denote the transition probabilities between health states by:

[13] When the consumer is fully annuitized, the value of her annuity depends on her health state. In particular, if she purchases an annuity in state i and then later transitions to a worse health state j, causing her life expectancy to fall, then the value of her annuity will also fall. This technicality is not reflected in the notation for equation (18); see Appendix D for details and discussion.

$$p_{ij}(t) = \mathbb{P}[Y_{t+1} = j | Y_t = i]$$

As in the continuous time model, the mortality rate at time t, $d(t)$, depends on the individual's health state:

$$d(t) = \sum_{j=1}^{n} \overline{d}^j(t) \mathbf{1}\{Y_t = j\}$$

where $\left\{\overline{d}^j(t)\right\}$ are given and $\mathbf{1}\{Y_t = j\}$ is an indicator variable equal to 1 if the individual is in state j at time t and 0 otherwise. The probability of surviving from time period t to time period s is denoted as $S_t(s)$, where

$$S_t(t) = 1,$$

$$S_t(s) = S_t(s-1)(1 - d_{s-1}), s > t$$

Let $c(t)$, $q(t)$, and $W(t)$ denote consumption, quality of life, and wealth in period t, respectively. Let ρ denote the utility discount rate, and r the interest rate. Assume that in each period the consumer receives an exogenously determined income, $y(t)$, and that the maximum lifespan of a consumer is T (i.e., $d(T) = 1$). Our baseline model assumes there is no bequest motive, although we plan to relax this assumption in a later exercise.

The consumer's maximization problem is

$$\max_{\{c(t)\}} \mathbb{E}_0 \left[\sum_{t=0}^{T} e^{-\rho t} S_0(t) u\big(c(t), q(t)\big) \right]$$

subject to

$$W(0) \text{ given},$$
$$W(t) \geq 0,$$

$$W(t+1) = \big(W(t) + y(t) - c(t)\big)e^r$$

We assume throughout that $r = \rho = 0.03$ (Siegel 1992; Moore and Viscusi 1990). Finally, we assume that utility takes the following CRRA form:

$$u(c,q) = q\frac{c^{1-\gamma}}{1-\gamma} - \frac{\underline{c}^{1-\gamma}}{1-\gamma} \tag{19}$$

As discussed in Section III, there is no consensus regarding the sign or magnitude of health state dependence ($u_{cq}(\cdot)$). Here, we assume a multiplicative relationship, so that the marginal utility of consumption is higher when quality of life is high, and vice versa.

We have normalized the utility of death to zero in (19). The consumer receives positive utility if she consumes an amount greater than $\underline{c}$, which represents a subsistence level of consumption. Consuming an amount less than $\underline{c}$ generates utility that is worse than death. Although adding a constant to the utility function does not affect the solution to the consumer's maximization problem, it matters when calculating the value of life.[14] We are unaware of any empirical evidence on the magnitude of $\underline{c}$, the subsistence level of consumption in the United States. We assume it is equal to \$5,000, which is in line with the parameterization employed in Murphy and Topel (2006).

The parameter γ is the inverse of the elasticity of intertemporal substitution, an important determinant of both the value of life and the value of annuitization. We follow Hall and Jones (2007) and set $\gamma = 2$ in our analyses. As points of reference, Murphy and Topel (2006) set $\gamma = 1.25$ while Brown (2001) uses survey data to estimate a mean value of $\gamma = 3.95$.

We employ dynamic programming techniques to solve for the optimal consumption path (see Appendix C for details). The value function is defined as:

$$V(t,w,i) = \max_{\{c(t)\}} \mathbb{E}\left[\sum_{s=t}^{T} e^{-\rho(s-t)} S_t(s) u(c(s)) \middle| Y_t = i\right]$$

We can use the value function to rewrite the optimization problem as a recursive Bellman equation:

$$V(t,w,i) = \max_{\{c(t)\}}\left[u\big(c(t)\big) + \frac{1-d(t)}{e^{\rho}}\sum_{j=1}^{N} p_{ij}(t) V\big(t+1, \big(W(t)+y(t)-c(t)\big)e^{r}, j\big)\right]$$

Once we have solved for the optimal consumption path, we can use the analytical formulas derived in the previous sections to calculate the value of life.

We are aware that there is significant uncertainty among economists regarding the proper values of many of the parameters in our model. The goal of the subsequent analyses is to illustrate the significance of our insights when our model is applied to real-world data using reasonable parameterizations.

IV.B. Annuitization, retirement policy, and the value of life

In this section, we explore the link between annuitization and retirement policy. We build up to these results by calculating how the value of statistical life varies over the life-cycle under alternative annuitization policies. We then calculate how these alternative policies influence the value of permanent reductions in mortality. All our calculations account for the effect of mortality reduction on net savings, regardless of the degree of annuitization, in order to facilitate comparison across different annuitization scenarios and because this is appropriate when estimating the social value of increased longevity. (See footnote 7.)

We initiate the model at age 20 and assume nobody survives past age 100. We use data on age-specific mortality rates from www.mortality.org. Because these mortality data are not available by health state, in this section we will work with a deterministic model. (This corresponds to specifying $n = 1$ health states in the framework above.) For this particular analysis, we also abstract from the role of quality of life and set $q(t) = 1 \forall t$ because aggregate, nationally representative data on quality-of-life trends are not generally available. (Quality of life is explicitly incorporated into the analysis presented in Section IV.C.) Finally, we choose the individual's labor earnings, $\{m(t)\}$, to fit data on life-cycle earnings as estimated by the Current Population Survey and the Health and Retirement Survey. See Appendix B1 for details.

---

[14] Rosen (1988) was the first to point out that the level of utility is an important determinant of the value of life. See also additional discussion on this point in Hall and Jones (2007).

The individual's period income is equal to $y(t) = (1-\tau)m(t) + a(t)$, where $a(t)$ is nonwage defined-benefit income financed by an earnings tax, τ. We consider three different policy scenarios in the main text. In the first, financial markets are absent and the consumer's income corresponds to labor earnings: $y^1(t) = m(t)$. Thus, her consumption is limited by current period income and savings from prior periods. The second scenario introduces an actuarially fair Social Security program that provides an annuity equal to $16,195 beginning at age 65.[15] In this second scenario, the consumer is partially annuitized, but she still lacks access to financial markets and cannot borrow against her future income. The third scenario increases the size of the Social Security pension by 50 percent. Finally, in the appendix we also present results for the case where the consumer fully annuitizes at age 20 and enjoys a constant annuity stream, $\overline{y} = \overline{a}$, provided by an actuarially fair and complete annuities market. The income streams in all scenarios are related according to the following equation:

$$\sum_{t=1}^{T} \frac{y^1(t)S(t)}{e^{r(t-1)}} = \sum_{t=1}^{T} \frac{y^2(t)S(t)}{e^{r(t-1)}} = \sum_{t=1}^{T} \frac{y^3(t)S(t)}{e^{r(t-1)}} = \overline{y} \sum_{t=1}^{T} \frac{S(t)}{e^{r(t-1)}}$$

Our assumed interest rate of 3 percent and our data on mortality and earnings imply a full annuity value of $\overline{y} = \$38{,}019$.

The life-cycle profiles of consumption for the first two policy scenarios are displayed in Figure 2. Consumption is constrained by the consumer's low income in early life. She saves during middle age when income is high, and then consumes her savings during retirement until eventually her consumption equals her pension (if available). Consumption for an individual with no annuity is "shifted forward" relative to an individual with a Social Security pension. This effect is particularly dramatic in the final 10 years of life, when old consumers outlive their wealth.[16] This is not surprising: a primary benefit of an annuity is its ability to provide income to consumers in their oldest ages.

Figure 3 shows that this difference in consumption generates a corresponding difference in the value of a life-year. Individuals place a low value on life-years in early and late ages, when consumption is low. The drop at age 65 reflects the effect of retirement on the net savings component of the value of life.

Figure 4 displays the corresponding value of statistical life (VSL) for these two scenarios, as calculated by equation (7). At age 40, VSL is equal to $7 million for an individual with no annuity, and $8 million for an individuals with a Social Security potential. These values are within the ranges estimated by empirical studies of VSL for working-age individuals (Viscusi and Aldy 2003). Figure 4 also shows that VSL is greater at older ages for a person with a Social Security pension than it is for a person with no annuity. This suggests that public annuity programs are complementary with retiree healthcare programs and other investments in life-extension for the elderly population.

Finally, we calculate the value of historical reductions in mortality for these different annuitization scenarios, as well as the prospective value of permanent reductions in future mortality for selected diseases. Let δ denote a vector of mortality reductions for different ages. As in Murphy and Topel (2006),

---

[15] This corresponds to the average retirement benefit paid by Social Security to retired workers in 2016 (www.ssa.gov/policy/docs/quickfacts/stat_snapshot/2016-07.pdf).

[16] Hubbard, Skinner, and Zeldes (1995) show that failing to include a "welfare floor" in the budget constraint causes life-cycle models to overestimate savings for low-income households. Our calibration exercises model median-income individuals, however, for whom this issue is less important.

we calculate the total social value of a mortality reduction by aggregating over the age distribution of the 2015 US population:

$$Social\,Value = \sum_{a=0}^{110} VLE(a, \delta) f(a)$$

where $VLE(a, \delta)$ is defined as in equation (5), and $f(a)$ is the count of individuals alive in 2015 at age a.[17]

We report our results in Table 1. Like Murphy and Topel (2006), we find that the social value of past longevity gains are enormous: the post-1940 gains are worth over \$100 trillion today, and the post-1970 gains are worth over \$50 trillion. Comparing results for different annuitization scenarios informs our understanding of the interaction between retirement policies and the value of longevity. For example, consider the introduction of Social Security over the last century. Comparing Column (1) to Column (2) of Table 1 suggests that this increased the value of post-1940 longevity gains by \$11.5 trillion (10.6 percent), and increased the value of post-1970 gains by \$6.2 trillion (11.6 percent). One way to interpret these values is to compare them to the current aggregate insurance value of Social Security, which is approximately \$17 trillion.[18] Thus, the interaction between post-1940 longevity gains and Social Security is worth about half as much as the entire Social Security program itself.

Table 1 also reveals that Social Security has raised the value of a 10 percent cancer mortality reduction by \$446 billion, or 13 percent. Alternatively, it has raised the value of a 10 percent reduction in all-cause mortality by \$1.43 trillion (12 percent). Column (3) reports that increasing the size of Social Security pensions by 50 percent would add \$488 billion more to that value.

To summarize, our model predicts that annuitization raises the value of life for the elderly. This should cause them to spend more on healthcare and invest more in healthy behaviors, which in turn should ultimately manifest in increased life expectancy. This dovetails with the point, made by Philipson and Becker (1998), that the moral hazard effects of retirement programs also increase the willingness to pay for longevity. Philipson and Becker (1998) analyze data from Virga (1996) to show that people with more generous annuities live longer than those with less generous annuities. They interpret this as the effect of endogenous longevity investments, which are encouraged among highly annuitized individuals who do not bear the full cost of an increase in their longevity. In our model, by contrast, annuitization increases the value of statistical life even if annuities are actuarially fair.[19] Given that these effects reinforce each other, it is not surprising that increases in the generosity of public pensions in developed countries have been accompanied by large increases in public spending on retiree healthcare.

---

[17] Specifically, $VLE(a, \delta) = \int_a^{100} e^{-r(t-a)} \left[\int_a^t \delta(s) ds\right] v(t) dt$. We assume $VLE(a, \delta) = VLE(20, \delta)$ for $a < 20$, and equal to $VLE(100, \delta)$ for $a > 100$. Unlike Murphy and Topel (2006), our social value calculation does not account for the value that mortality reductions generate for future (unborn) populations.

[18] This value is calculated using the methodology of Mitchell et al. (1999) and does not account for other potential benefits of Social Security such as protection against inflation risk. See Appendix C1 for details.

[19] As discussed at the beginning of this section, all our calculations account for the effect of mortality reduction on net savings.

IV.C. Heterogeneity in VSL, and the value of prevention versus treatment

The conventional economic theory of life extension conceives of VSL as depending primarily on age and consumption. The general framework with stochastic mortality and incomplete annuitization implies instead a substantial amount of variability in VSL within these categories alone. Individuals who have experienced a recent negative mortality shock have systematically higher VSL, but this VSL premium decays over time. We use real-world data on mortality, quality of life, health shocks, and income to estimate the degree to which VSL varies within the traditional categories, and the factors explaining the variation.

These data are provided by the Future Elderly Model (FEM), a widely published microsimulation model that employs comprehensive, nationally representative data from a wide array of sources (Michaud et al. 2011; Goldman et al. 2005; Lakdawalla, Goldman, and Shang 2005; Goldman et al. 2009; Lakdawalla et al. 2009; Goldman et al. 2013; Michaud et al. 2012; Goldman et al. 2010). The model produces estimates of mortality, disease incidence, and quality of life at the individual level for people over the age of 50 with different comorbid conditions.[20] The FEM accounts for six different chronic conditions (cancer, diabetes, heart disease, hypertension, chronic lung disease, and stroke) and six different impaired activities of daily living (bathing, eating, dressing, walking, getting in or out of bed, and using the toilet).

We divide the health space within the FEM into 20 states. Each state corresponds to the number (0, 1, 2, 3 or more) of impaired activities of daily living (ADL) and the number (0, 1, 2, 3, 4 or more) of chronic conditions, for a total of $4 \times 5 = 20$ health states. States are ordered first by number of ADL's and then by number of chronic diseases. So state 1 corresponds to 0 ADL's and 0 chronic conditions, state 2 corresponds to 0 ADL's and 1 chronic condition, and so on. For each health state and age, the FEM estimates the probability of dying and the probability of transitioning to each of the other health states in the next year. As in the theoretical model, individuals can transition only to higher-numbered states, i.e., $p_{ij}(t) = 0 \forall j < i$. In other words, all ADL's and chronic conditions are permanent. The FEM also estimates quality of life for each health state and age, as measured by the EuroQol five dimensions questionnaire (EQ-5D). The EQ-5D is a well-validated tool that measures quality of life on a scale from zero to one, using answers from five survey questions regarding the extent of a person's problems in mobility, self-care, daily activities, pain, and anxiety/depression.

Table 2 presents basic descriptive statistics for the data provided by the FEM model. Life expectancy at age 50 ranges from 35 years for a healthy individual in state 1 to 15 years for an ill individual in state 20. Quality of life, as measured by the EQ-5D index, ranges 0.54 to 0.88. Columns (7)-(8) of Table 2 report the annual probability that an individual exits her health state but remains alive, i.e., acquires at least one new ADL or chronic condition. Health states are relatively persistent, with exit rates never exceeding 15 percent. State 20 is an absorbing state with an exit rate of 0 percent.

We focus here on a setting where individuals do not have access to annuity markets. We make two simplifying assumptions that allow us to generate exact, analytical solutions to the stochastic mortality model: we assume an individual can borrow against her future income, and that income is not survival contingent. These two assumptions imply an equivalence between income and wealth, allowing us to

[20] Additional details about its methodology are provided in Appendix B2. A complete technical description of the FEM is available at roybalhealthpolicy.usc.edu/fem/technical-specifications/.

ignore income and to work with wealth only.[21] See the appendix for details on how the model is calculated. We set initial wealth equal to \$807,226 at age 50, which corresponds to the net present value of all wealth and future earnings at age 50 as estimated by the deterministic model presented in the prior section. All other parameterizations are the same as before.

If an individual never suffers a health shock, then her consumption and VSL profiles will decline smoothly with age, as shown in Figure 4 for ages over 50. Figure 5 demonstrates a key mechanism for variability in VSL when mortality is stochastic: The arrival of a health shock can increase VSL, sometimes substantially. The figure displays contrasting plots for an initially healthy individual who develops one ADL at age 60, and then two more ADLs plus four chronic conditions at age 80. The first shock reduces her life expectancy by 2.3 years. The second one reduces her life expectancy by 7.4 years. In contrast to the healthy consumer, the sick consumer's consumption exhibits discontinuous jumps at ages 60 and 80 as a result of the negative health shocks. The first shock has little effect on the declining trend in VSL, but the second one nearly doubles her VSL at age 80, from \$1.5 million to \$2.9 million. While this second effect is dramatic, it is rare: the FEM estimates that an 80-year-old with one ADL has less than a one percent chance of developing two more ADLS plus four chronic conditions within the following year.

Nevertheless, the arrival of shocks at the individual level generate substantial variability in VSL in the aggregate. Figure 6 reports results from a Monte Carlo simulation of 10,000 life-cycle modeling exercises. At age 50, all individuals are identical and have a VSL of \$5.7 million. As they age, some begin to suffer health shocks that, at least initially, increase their VSL. By age 60, the VSL inter-vigintile range spans \$4 to \$5 million. This dispersion is compressed towards the end of life, where mortality eventually reaches 100 percent.

The stochastic mortality approach also allows us to calculate the value of a statistical illness (VSI). Column (3) of Table 3 reports VSI at age 50 from the perspective of a healthy individual. Each value represents the healthy individual's willingness to pay for a marginal reduction in the probability of developing an illness corresponding to one of the 19 other health states. The values are increasing functions of life expectancy in the sick state because it is more valuable to prevent the onset of a lethal disease than a mild one. The highest VSI value is \$3.3 million, which corresponds to preventing the onset of a sick state with 3 ADL's and 4 chronic conditions (health state 20). The interpretation of this value is analogous to VSL: it is the amount that 1,000 individuals would collectively be willing to pay in order to reduce their risk of developing this illness by 1/1000. In our framework, VSL can be interpreted as the willingness to pay to avoid the "illness" of dying, which correspond to a state with 0 years of remaining life expectancy. VSI corresponds to the willingness to pay to avoid a lower life expectancy, but still living, state.

How does the value of prevention compare to the value of treatment? We investigate this question by normalizing VSL and VSI by the number of life-years "saved." We report the results of those calculations in Table 3. Our VSL estimate implies that an individual with one chronic condition and no ADL's (health state 2) has a marginal willingness-to-pay of \$203,000 per life-year for a treatment that extends her life. By contrast, our VSI estimate implies that a healthy individual (health state 1) is only willing to pay \$102,000 per life-year for a preventive treatment that reduces her chances of developing 1 chronic condition. The ratio of these values is equal to 1.99. Column (6) of Table 2 shows that this ratio is always

[21] Generalizing the model to allow for partial annuitization is possible but prohibits the calculation of an exact solution. The effect of annuitization on the value of life is illustrated by the deterministic mortality model presented in the previous section.

greater than 1, and sometimes significantly so: therapeutic treatments for very sick individuals with 3+ ADL's and 4+ chronic conditions (health state 20) are worth nearly 10 times more than equally effective preventive care for healthy individuals. This is in stark contrast to the standard cost-effectiveness framework, which values prevention and treatment equally. The model developed earlier provides the intuition behind these differences: all else equal, the marginal utility of consumption is higher for a healthy individual than for someone who has just suffered a health shock that reduces her lifespan. This in turn drives the difference in willingness-to-pay.

Figure 8 shows how the value of statistical life varies with remaining life expectancy, and how it varies between treatment and prevention investments. The x-axis in the figure depicts 19 different health states, ordered by increasing severity. More severe health states involve shorter life expectancy. The solid blue bars demonstrate that the value of life-years gained through treatment is monotonically higher for states with lower remaining life expectancy. An individual with 9.6 years of life expectancy will pay over \$1.3m per life-year gained, while her counterpart with 30 years of life left will only pay \$200K. This suggests that consumers facing greater fatality risk are willing to pay substantially more for a given gain in life expectancy.

The dotted red bars show the value per life-year gained through prevention. For instance, the left-most dotted red bar reports the value of each life-year saved when a perfectly healthy consumer reduces the risk of entering the health state with 30 years of life expectancy. The value of life-years gained through prevention is systematically lower than the value of life-years gained through treatment. This suggests that imperfectly annuitized consumers are willing to pay more for life-extension when they find themselves in a sicker, lower life expectancy state. Note also that the value of prevention is quite insensitive to the severity of the health state being prevented. This is also reasonably intuitive, since the prevention investment is being made in the same (perfectly healthy) state throughout. Thus, there are no major differences in the marginal utility of consumption.

V. CONCLUSION

The economic theory surrounding the value of life has many important applications. Yet, like most theories, it suffers from a few anomalies that appear at odds with intuition, common sense, or empirical facts. We have demonstrated that several of these anomalies are easily explained without abandoning the standard framework, simply by relaxing its strong assumptions around the completeness of annuity markets and deterministic mortality. Moreover, relaxing these assumptions generates new predictions with implications for health policy and behavior. We show that VSL varies with the arrival of mortality shocks and with remaining life expectancy. A given gain in longevity is more valuable to a consumer who has less life remaining, and vice-versa. Even holding wealth and income fixed, VSL may vary by \$1 million or more for a 50-year-old. In addition, we demonstrate an interaction between annuity policy and health policy: Completing the annuity market may significantly increase the value of life-extension, especially for the elderly. For instance, the US Social Security program has increased the value of mortality reductions, adding as much as \$150 billion to the value of a 1 percent mortality decline.

Our findings have several implications for the valuation of health investments and for policy more generally. The value of a life-year will tend to vary across types of risk, not just across types of people. It can be more valuable to add one month of life for a patient facing a highly fatal disease than for one facing a much milder ailment. Thus, health spending should be more targeted towards the severely ill than current economic models of cost-effectiveness suggest.

In addition, public programs that expand the market for annuities might simultaneously boost the demand for life-extending technologies. Intuitively, annuities calm consumer fears about outliving their wealth and thus enable more aggressive investments in life-extension. Viewed differently, our results also show that market failures in annuities affect the value of statistical life, and thus the socially optimal level of health care spending.

Finally, our framework offers a single unified framework for valuing both life-extension and the prevention of illness. This provides a more practical tool for policymakers and decision makers, since many health investments involve preventing the deterioration of health, not a direct and immediate mortality risk. Our result also provides one explanation for why it has proven to be so difficult for policymakers and public health advocates to encourage investments in the prevention of disease. From the private perspective, prevention is often less valuable than treatment, even though there may be public goods – e.g., savings in public health insurance programs – associated with prevention investments. Kremer and Snyder (2015) show that heterogeneity in consumer values distorts R&D incentives by allowing firms extract more consumer surplus from treatments than with preventives. Our results suggest that the deadweight loss associated with this may be less large than previously believed, because treatments generate more private value than preventives.

Our analysis raises a number of important questions for further research. First, how does the value of longevity vary with endogenous demand for quality of life? Elsewhere, we have studied how incomplete health insurance enhances the value of medical technology that improves quality of life, because such technology acts as insurance by compressing the difference in utility between the sick and healthy states (Lakdawalla, Malani, and Reif 2017). Less clear is how demands for the quantity and quality of life interact with financial market incompleteness of various kinds. Second, what does the generalized value of life model mean for the value of different kinds of medical technologies? For instance, the model suggests that short-term survival gains for high-risk diseases are more valuable than previously believed, but very long-term survival gains might actually be less valuable than previously believed. Finally, what are the implications for the empirical literature on the value of statistical life? Empirical analysis has typically proceeded under the assumption that different kinds of mortality risk are all valued the same way, as long as they imply similar changes in the probability of dying (Viscusi and Aldy 2003; Hirth et al. 2000; Mrozek and Taylor 2002). Our framework casts doubt on this assumption and suggests the need for a more nuanced empirical approach. This missing insight may be one reason for the widely disparate estimates in the empirical literature on the value of a statistical life.

VI. REFERENCES

Arthur, W Brian. 1981. 'The economics of risks to life', *The American Economic Review*: 54-64.

Banks, James, Richard Blundell, and Sarah Tanner. 1998. 'Is there a retirement-savings puzzle?', *American Economic Review*: 769-88.

Bertsekas, D.P. 2005. *Dynamic Programming and Optimal Control* (Athena Scientific).

Brown, Jeffrey R. 2001. 'Private pensions, mortality risk, and the decision to annuitize', *Journal of public economics*, 82: 29-62.

Brown, Jeffrey R, Jeffrey R Kling, Sendhil Mullainathan, and Marian V Wrobel. 2008. 'Why Don't People Insure Late-Life Consumption? A Framing Explanation of the Under-Annuitization Puzzle', *The American Economic Review*, 98: 304-09.

Carroll, Christopher D, and Lawrence H Summers. 1991. 'Consumption growth parallels income growth: Some new evidence.' in, *National Saving and Economic Performance* (University of Chicago Press).

Dranitsaris, George, and George Papadopoulos. 2015. 'Health technology assessment of cancer drugs in Canada, the United Kingdom and Australia: should the United States take notice?', *Applied health economics and health policy*, 13: 291-302.

Drummond, Michael F, Mark J Sculpher, George W Torrance, Bernie J O'Brien, and Greg L Stoddart. 2005a. 'Methods for the economic evaluation of health care programmes', *OUP Catalogue*.

Drummond, Michael F., Mark J. Sculpher, George W. Torrance, Bernie J. O'Brien, and Greg L. Stoddart. 2005b. *Methods for the Economic Evaluation of Health Care Programmes*.

Edwards, Ryan D. 2008. 'Health risk and portfolio choice', *Journal of Business & Economic Statistics*, 26: 472-85.

Ehrlich, Isaac. 2000. 'Uncertain lifetime, life protection, and the value of life saving', *Journal of Health Economics*, 19: 341-67.

Evans, William N, and W Kip Viscusi. 1991. 'Estimation of state-dependent utility functions using survey data', *The Review of Economics and Statistics*: 94-104.

Fernandez-Villaverde, Jesus, and Dirk Krueger. 2007. 'Consumption over the life cycle: Facts from consumer expenditure survey data', *The Review of Economics and Statistics*, 89: 552-65.

Finkelstein, Amy, Erzo FP Luttmer, and Matthew J Notowidigdo. 2013. 'What good is wealth without health? The effect of health on the marginal utility of consumption', *Journal of the European Economic Association*, 11: 221-58.

Goldman, Dana, Samuel Nussbaum, and Mark Linthicum. 2016. "Rapid biomedical innovation calls for similar innovation in pricing and value measurement." In *Health Affairs Blog*.

Goldman, Dana P, David Cutler, John W Rowe, Pierre-Carl Michaud, Jeffrey Sullivan, Desi Peneva, and S Jay Olshansky. 2013. 'Substantial health and economic returns from delayed aging may warrant a new focus for medical research', *Health Affairs*, 32: 1698-705.

Goldman, Dana P, Pierre-Carl Michaud, Darius Lakdawalla, Yuhui Zheng, Adam Gailey, and Igor Vaynman. 2010. 'The fiscal consequences of trends in population health', *National Tax Journal*, 63: 307-30.

Goldman, Dana P, Baoping Shang, Jayanta Bhattacharya, Alan M Garber, Michael Hurd, Geoffrey F Joyce, Darius N Lakdawalla, Constantijn Panis, and Paul G Shekelle. 2005. 'Consequences of health trends and medical innovation for the future elderly', *Health Affairs*, 24: W5.

Goldman, Dana P, Yuhui Zheng, Federico Girosi, Pierre-Carl Michaud, S Jay Olshansky, David Cutler, and John W Rowe. 2009. 'The benefits of risk factor prevention in Americans aged 51 years and older', *American Journal of Public Health*, 99: 2096.

Green, C., and K. Gerard. 2009. 'Exploring the social value of health-care interventions: a stated preference discrete choice experiment', *Health Economics*, 18: 951-76.

Hall, Robert E, and Charles I Jones. 2007. 'The Value of Life and the Rise in Health Spending', *The Quarterly Journal of Economics*, 122: 39-72.

Hirth, Richard A, Michael E Chernew, Edward Miller, A Mark Fendrick, and William G Weissert. 2000. 'Willingness to pay for a quality-adjusted life year in search of a standard', *Medical Decision Making*, 20: 332-42.

Hubbard, R Glenn, Jonathan Skinner, and Stephen P Zeldes. 1995. 'Precautionary Saving and Social Insurance', *Journal of Political Economy*, 103.

Johansson, Per-Olov. 2002. 'On the definition and age-dependency of the value of a statistical life', *Journal of Risk and Uncertainty*, 25: 251-63.

Kimball, Miles S. 1990. 'Precautionary Saving in the Small and in the Large', *Econometrica: Journal of the Econometric Society*: 53-73.

Kremer, Michael, and Christopher M Snyder. 2015. 'Preventives versus treatments', *The Quarterly Journal of Economics*, 130: 1167-239.

Lakdawalla, Darius, Anup Malani, and Julian Reif. 2017. 'The Insurance Value of Medical Innovation', *Journal of public economics*, 145: 94-102.

Lakdawalla, Darius N, Dana P Goldman, Pierre-Carl Michaud, Neeraj Sood, Robert Lempert, Ze Cong, Han de Vries, and Italo Gutierrez. 2009. 'US pharmaceutical policy in a global marketplace', *Health Affairs*, 28: w138-w50.

Lakdawalla, Darius N, Dana P Goldman, and Baoping Shang. 2005. 'The health and cost consequences of obesity among the future elderly', *Health Affairs*, 24: W5R30.

Lillard, Lee A, and Yoram Weiss. 1997. 'Uncertain health and survival: Effects on end-of-life consumption', *Journal of Business & Economic Statistics*, 15: 254-68.

Linley, Warren G., and Dyfrig A. Hughes. 2013. 'Societal views on NICE, Cancer Drugs Fund, and Value Based Pricing Criteria for Prioritising Medicines: A Cross-Sectionalk Survey of 4118 Adults in Great Britain', *Health Economics*, 22: 948-64.

Michaud, Pierre-Carl, Dana Goldman, Darius Lakdawalla, Adam Gailey, and Yuhui Zheng. 2011. 'Differences in health between Americans and Western Europeans: Effects on longevity and public finance', *Social Science and Medicine*, 73: 254-63.

Michaud, Pierre-Carl, Dana P Goldman, Darius N Lakdawalla, Yuhui Zheng, and Adam H Gailey. 2012. 'The value of medical and pharmaceutical interventions for reducing obesity', *Journal of Health Economics*, 31: 630-43.

Mitchell, Olivia S, James M Poterba, Mark J Warshawsky, and Jeffrey R Brown. 1999. 'New evidence on the money's worth of individual annuities', *American Economic Review*, 89: 1299-318.

Moore, Michael J, and W Kip Viscusi. 1990. 'Models for estimating discount rates for long-term health risks using labor market data', *Journal of Risk and Uncertainty*, 3: 381-401.

Mrozek, Janusz R, and Laura O Taylor. 2002. 'What determines the value of life? a meta-analysis', *Journal of Policy Analysis and Management*, 21: 253-70.

Murphy, Kevin M, and Robert H Topel. 2006. 'The Value of Health and Longevity', *Journal of Political Economy*, 114: 871-904.

Nord, E., J. Richardson, A. Street, H. Kuhse, and P. Singer. 1995. 'Maximizing health benefits vs egalitarianism: an Australian survey of health issues', *Social Science and Medicine*, 41: 1429-37.

Parpas, Panos, and Mort Webster. 2013. 'A stochastic minimum principle and an adaptive pathwise algorithm for stochastic optimal control', *Automatica*, 49: 1663-71.

Philipson, Tomas J, and Gary S Becker. 1998. 'Old-age longevity and mortality-contingent claims', *Journal of Political Economy*, 106: 551-73.

Reichling, Felix, and Kent Smetters. 2015. 'Optimal Annuitization with Stochastic Mortality and Correlated Medical Costs', *The American Economic Review*, 105: 3273-320.

Rosen, Sherwin. 1988. 'The value of changes in life expectancy', *Journal of Risk and Uncertainty*, 1: 285-304.

Seierstad, Atle, and Knut Sydsaeter. 1977. 'Sufficient conditions in optimal control theory', *International Economic Review*: 367-91.

Shepard, Donald S, and Richard J Zeckhauser. 1984. 'Survival versus consumption', *Management Science*, 30: 423-39.

Siegel, Jeremy J. 1992. 'The real rate of interest from 1800–1990: A study of the US and the UK', *Journal of Monetary Economics*, 29: 227-52.

Sloan, Frank A, W Kip Viscusi, Harrell W Chesson, Christopher J Conover, and Kathryn Whetten-Goldstein. 1998. 'Alternative approaches to valuing intangible health losses: the evidence for multiple sclerosis', *Journal of Health Economics*, 17: 475-97.

Virga, M. 1996. 'Earn More, Live Longer-Variation in Mortality by Income Level', *Pension Section News*: 1-7.

Viscusi, W Kip, and Joseph E Aldy. 2003. 'The value of a statistical life: a critical review of market estimates throughout the world', *Journal of Risk and Uncertainty*, 27: 5-76.

Viscusi, W Kip, and William N Evans. 1990. 'Utility functions that depend on health status: estimates and economic implications', *The American Economic Review*: 353-74.

Yaari, Menahem E. 1965. 'Uncertain lifetime, life insurance, and the theory of the consumer', *The review of economic studies*: 137-50.

VII. TABLES AND FIGURES

Table 1. Aggregate social value of historical and prospective reductions in mortality (billions of dollars)

	(1)	**(2)**	**(3)**
	No annuity	**Social Security**	**Social Security + 50%**
Historical reduction:			
1940-2010	$109,252	$120,799	$126,398
1970-2010	$53,447	$59,651	$62,728
10% reduction, all ages:			
All causes	$11,715	$13,142	$13,630
Cancer	$3,417	$3,863	$3,790
Diabetes	$375	$422	$421
Heart disease	$2,444	$2,772	$2,782
Homicide	$109	$106	$180
Infectious diseases	$165	$188	$192

Notes: Aggregate values were calculated by using the 2015 US population by age. Column (1) presents estimates under the assumption that individuals have no annuities in retirement. Column (2) presents estimates under the assumption that individuals receive typical Social Security benefits that are financed by an earnings tax. Column (3) increases the generosity of Social Security by 50%, financed by an increase in the earnings tax. The net present value of individuals' wealth at age 20 is the same across all three columns.

Table 2. Summary statistics for the twenty health states employed by the stochastic mortality model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Life expectancy		Quality of life		Exit probability	
Health state	ADL's	Chronic conditions	Age 50	Age 70	Age 50	Age 70	Age 50	Age 70
1 (healthy)	0	0	32.8	16.3	0.884	0.874	4.3%	12.6%
2	0	1	30.0	14.6	0.850	0.841	3.6%	10.5%
3	0	2	26.3	12.5	0.812	0.804	3.6%	10.2%
4	0	3	22.1	10.5	0.773	0.764	4.0%	10.3%
5	0	4+	17.0	8.5	0.730	0.719	4.4%	7.6%
6	1	0	28.4	14.2	0.830	0.816	6.4%	15.1%
7	1	1	25.8	12.8	0.795	0.782	5.9%	12.8%
8	1	2	22.0	11.0	0.754	0.745	5.6%	12.2%
9	1	3	17.8	9.1	0.716	0.705	6.5%	11.8%
10	1	4+	14.3	7.4	0.669	0.661	6.5%	8.7%
11	2	0	26.1	13.2	0.781	0.763	7.5%	16.0%
12	2	1	23.2	11.6	0.746	0.731	7.2%	13.9%
13	2	2	19.5	10.0	0.706	0.693	7.6%	13.8%
14	2	3	15.9	8.3	0.669	0.655	7.4%	12.9%
15	2	4+	12.3	6.5	0.630	0.606	8.0%	11.4%
16	3+	0	23.4	11.1	0.700	0.693	3.6%	10.4%
17	3+	1	20.5	10.1	0.664	0.659	3.2%	8.8%
18	3+	2	16.9	8.4	0.622	0.622	2.3%	7.2%
19	3+	3	13.2	6.9	0.584	0.585	1.3%	5.2%
20	3+	4+	9.6	5.3	0.536	0.540	0.0%	0.0%

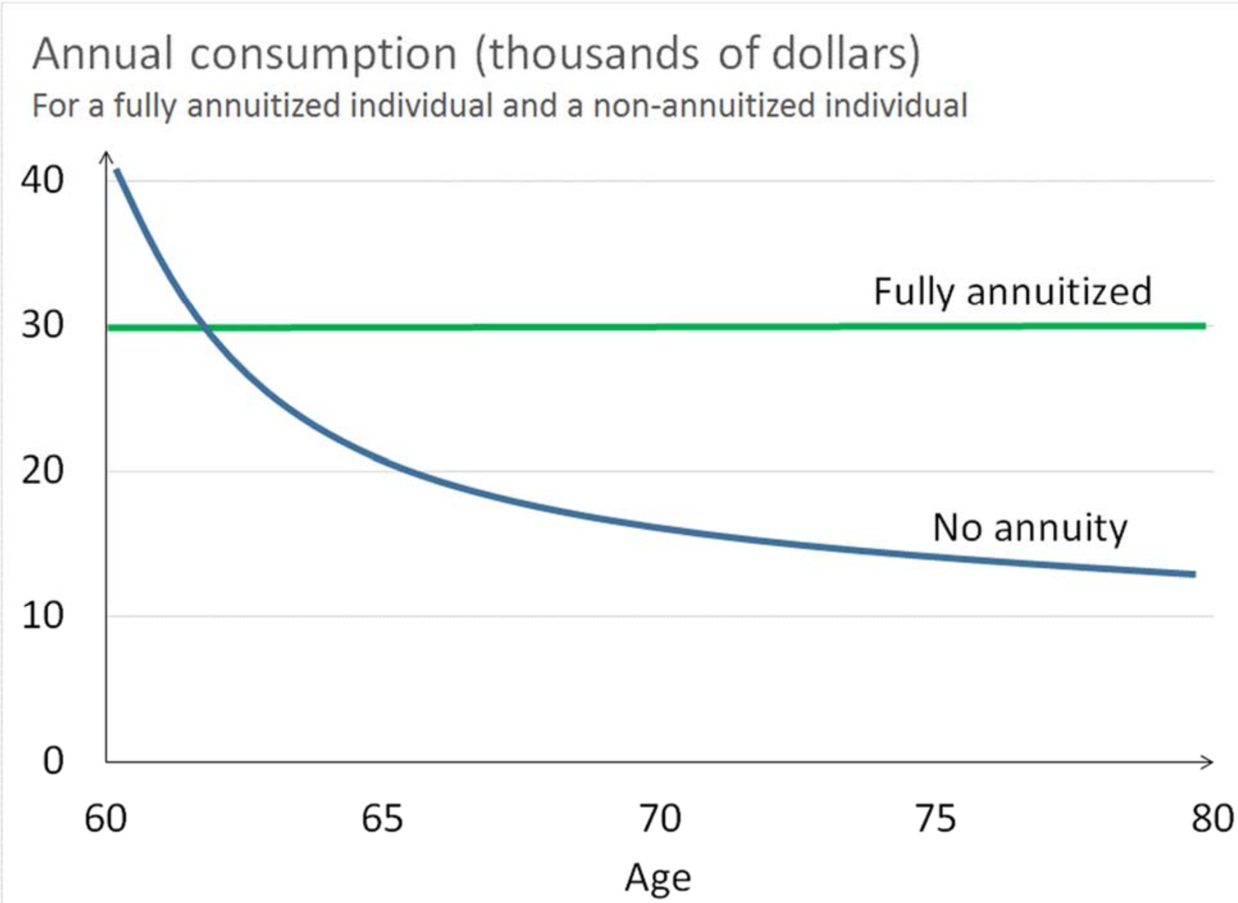
Notes: Columns (1) and (2) report the number of impaired activities of daily living (ADL) and the number of chronic conditions corresponding to each health state. Column (3)-(6) report life expectancy and quality of life for an individual in one of these health states. Quality of life is measured using the EQ-5D index, which ranges from 0 (death) to 1 (perfect healthy). Columns (7)-(8) report the probability that an individual transitions to a different health state in the following year. Source: Future Elderly Model.

Table 3. Value of treatment and prevention (in thousands of dollars) at age 50

	(1)	(2)	(3)	(4)	(5)	(6)
				WTP per life-year		
Health state	**Life expectancy**	**VSL**	**VSI**	**Treatment**	**Prevention**	**Treatment/Prevention**
1 (healthy)	32.8	$5,713	N/A	$174	N/A	N/A
2	30.0	$6,090	$281	$203	$102	1.99
6	28.5	$6,523	$432	$229	$100	2.29
3	26.3	$6,644	$706	$253	$109	2.32
11	26.1	$7,085	$701	$271	$105	2.59
7	25.8	$6,993	$739	$271	$106	2.55
16	23.5	$7,612	$1,063	$324	$115	2.82
12	23.1	$7,662	$1,085	$331	$112	2.95
4	22.1	$7,381	$1,248	$334	$117	2.86
8	21.9	$7,714	$1,250	$352	$115	3.05
17	20.6	$8,252	$1,474	$401	$121	3.33
13	19.5	$8,494	$1,586	$435	$120	3.64
9	17.7	$8,661	$1,872	$488	$124	3.93
5	17.0	$8,429	$2,015	$497	$127	3.90
18	16.9	$9,249	$2,020	$548	$127	4.32
14	16.0	$9,563	$2,121	$598	$126	4.73
10	14.3	$9,955	$2,395	$694	$130	5.34
19	13.2	$10,561	$2,628	$800	$134	5.96
15	12.3	$11,045	$2,749	$895	$135	6.65
20	9.6	$12,720	$3,286	$1,329	$142	9.38

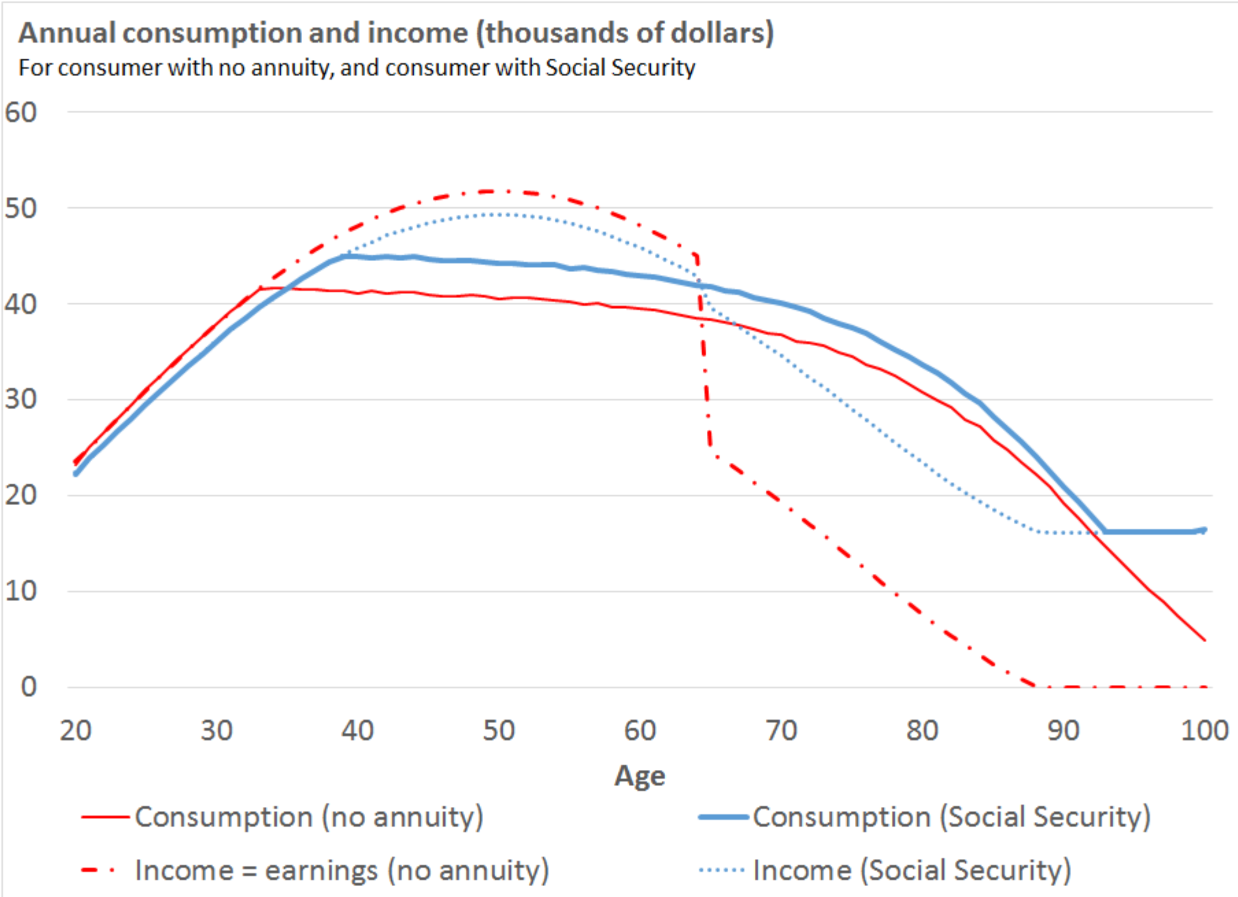
Notes: Table displays values (in thousands of dollars) from a life-cycle modeling exercise where mortality is stochastic. Values are sorted by life expectancy at age 50, as reported in column (1). Column (2) reports value of statistical life (VSL) for each health state. Column (3) reports the values of statistical illness (VSI) for a healthy individual in state 1, i.e., the individual's willingness-to-pay (WTP) to prevent a marginal increase in the probability of transitioning to one of the other 19 health states. Column (4) reports a sick individual's WTP per life-year for a therapeutic treatment, which is equal to the value in column (2) divided by the value in column (1). Column (5) reports the healthy individual's corresponding WTP for a preventive treatment, which is equal to the value in column (3) divided by the difference between 32.8 (life expectancy when healthy) and the value in column (1). Column (6) reports the ratio of the values reported in columns (4) and (5). The twenty health states are defined in Table 2.

Figure 1. Illustrative example: annual consumption for fully annuitized and non-annuitized consumers



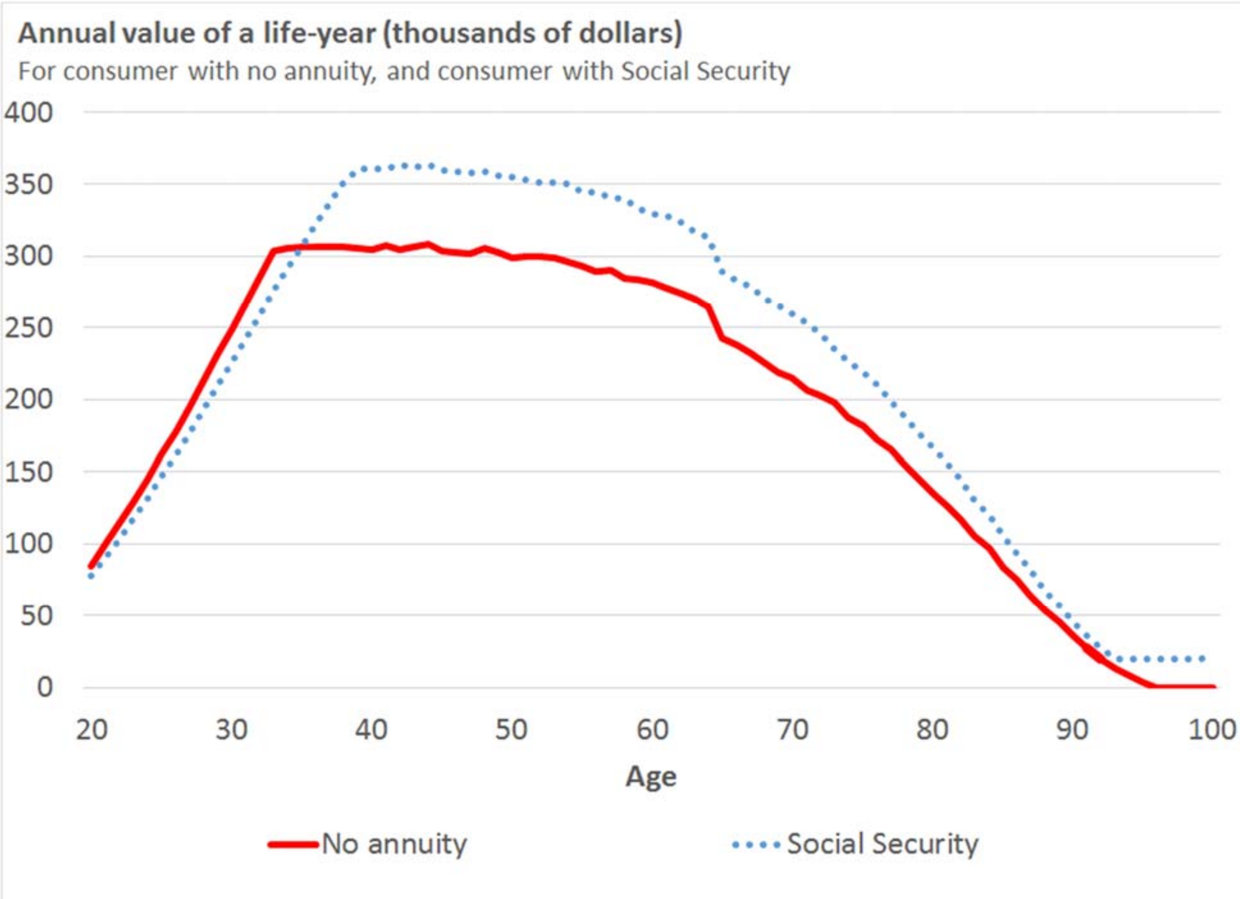
Notes: This figure illustrates the well-known result that it is optimal for a non-annuitized consumer who is exposed to longevity risk to shift her consumption forward in time, relative to a fully annuitized consumer. For the sake of simplicity, this example assumes that the consumption profile of the fully annuitized consumer is flat.

Figure 2. Life-cycle profiles of consumption and income when mortality is deterministic



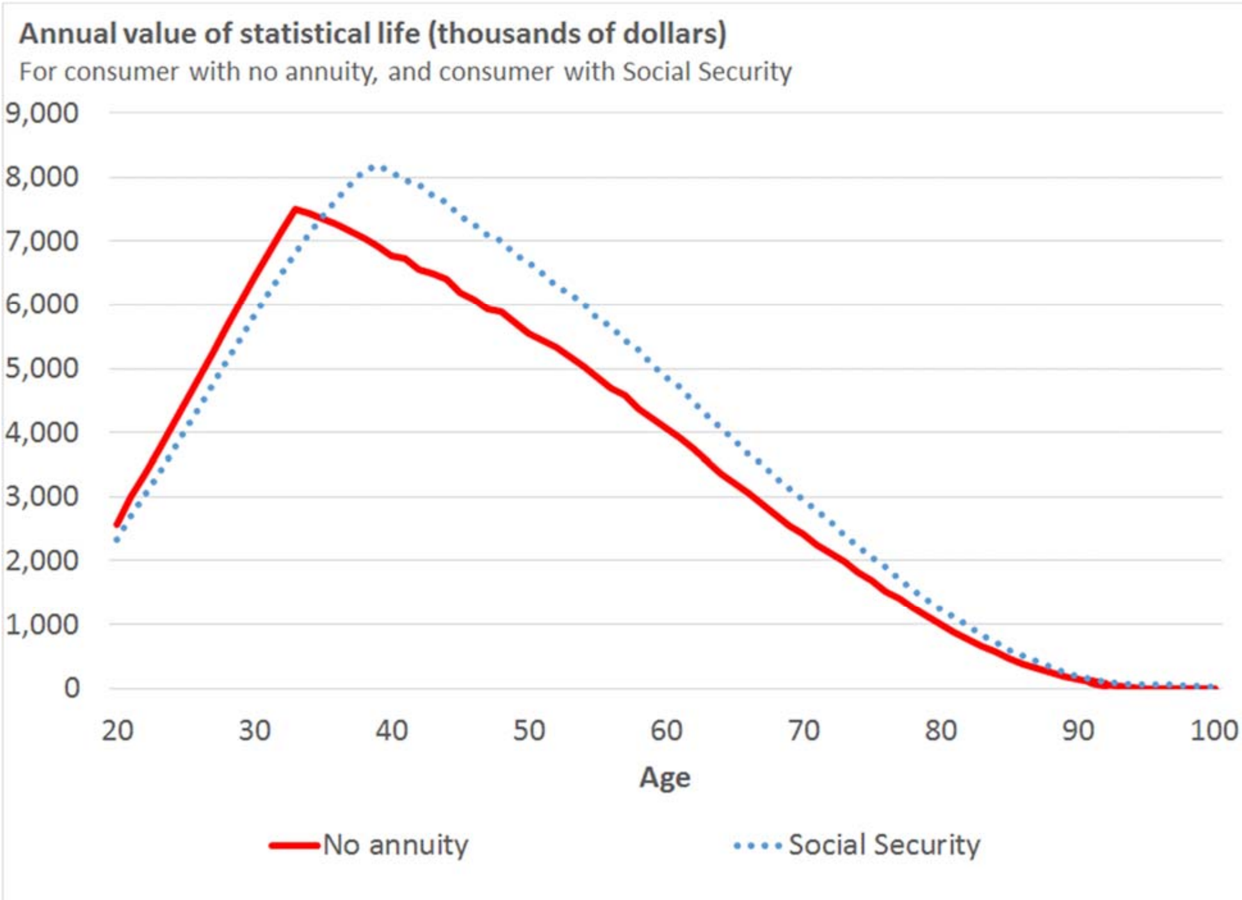
Notes: Figure plots consumption results from a life-cycle modeling exercise where mortality is deterministic. "Consumption (no annuity)" displays consumption for a consumer whose income equals her earnings. "Consumption (Social Security)" displays consumption for a consumer receiving typical Social Security benefits that are financed by an earnings tax. The net present value at age 20 of all future income is the same across both scenarios.

Figure 3. Life-cycle profile of the value of a life-year when mortality is deterministic



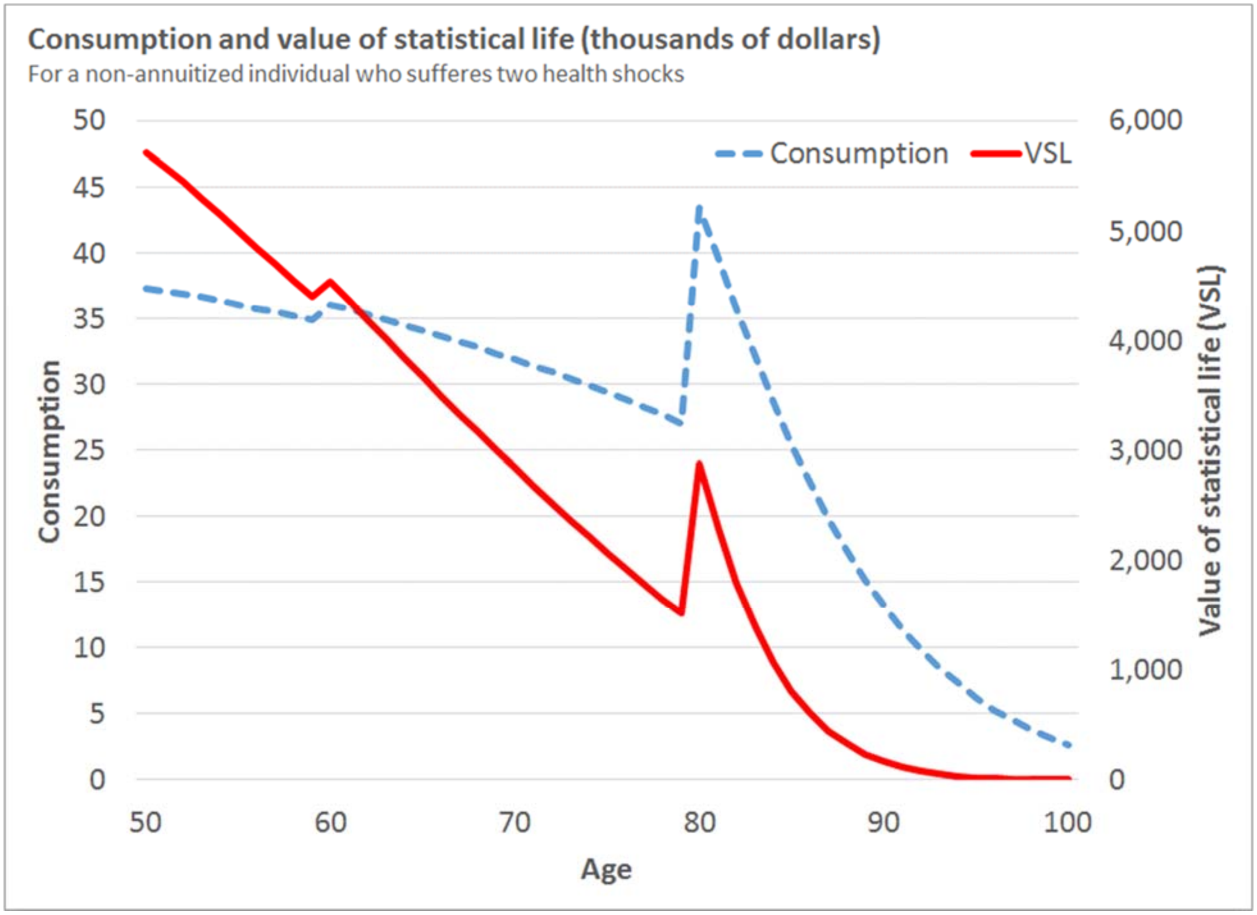
Notes: Figure plots the value of a life-year for the two scenarios displayed in Figure 2. "No annuity" assumes the consumer's income equals her earnings. "Social Security" assumes the consumer receives typical Social Security benefits that are financed by an earnings tax. The net present value at age 20 of all future income is identical in both scenarios.

Figure 4. Life-cycle profile of the value of statistical life when mortality is deterministic



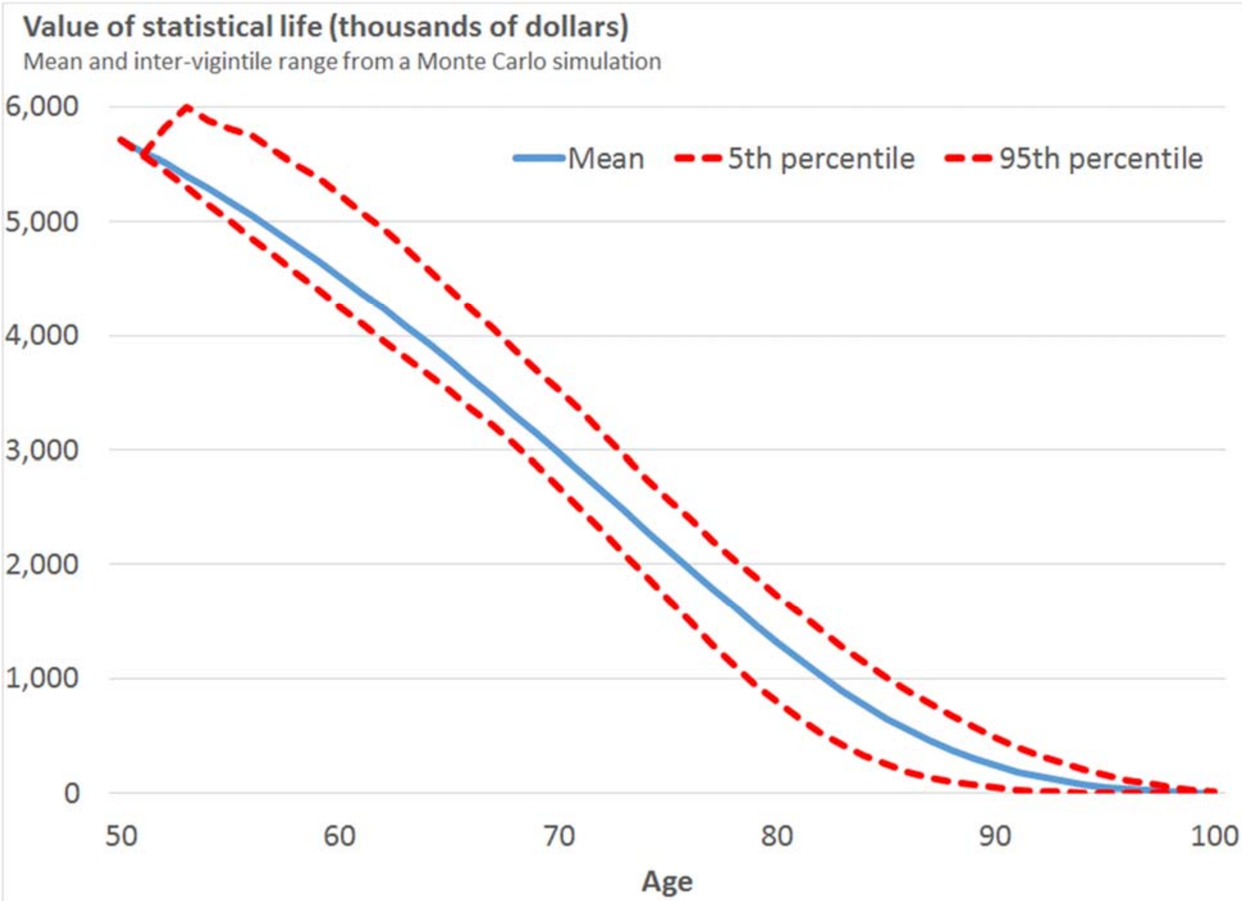
Notes: Figure plots the value of statistical life for the two scenarios displayed in Figure 2. "No annuity" assumes the consumer's income equals her earnings. "Social Security" assumes the consumer receives typical Social Security benefits that are financed by an earnings tax. The net present value at age 20 of all future income is identical in both scenarios.

Figure 5. Consumption and the value of statistical life increase when an individual falls ill



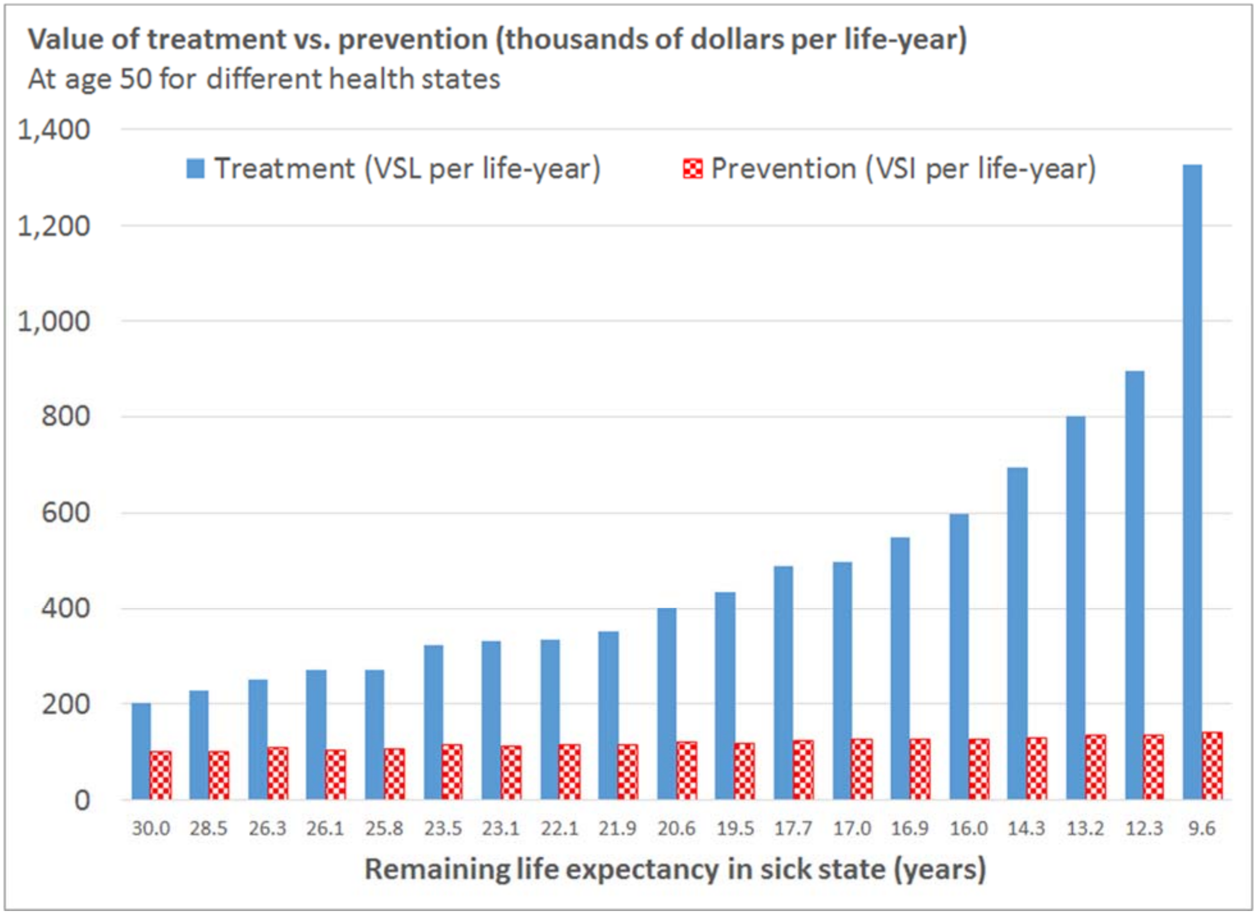
Notes: The figure plots an individual's consumption profile (left axis) and corresponding value of statistical life (right axis) as calculated from a life-cycle modeling exercise where mortality is stochastic. This consumer is healthy at age 50, but then falls ill twice, once at age 60 and then again at age 80. At age 60, the illness causes permanent difficulties with one routine activity of daily living (ADL). At age 80, she is diagnosed with two chronic conditions and subsequently has difficulties with two additional ADL's. The second illness is severe enough that it causes a 50 percent increase in her VSL.

Figure 6. The value of statistical life depends on an individual's health history



Notes: The figure reports the mean, 5th percentile, and 95th percentile of VSL from a Monte Carlo simulation that is repeated 10,000 times. Each individual began the simulation at age 50 in the same healthy state. Stochastic health shocks generate differences in VSL at older ages.

Figure 7. Treatments for an ill patient are worth more than preventive care for a healthy individual



Notes: The blue solid bars report the value of statistical life (VSL) for an individual in one of 19 different sick states, divided by life expectancy in that state. The red dotted bars report the value of statistical illness (VSI) for a healthy individual (life expectancy: 32.8 years) divided by the *reduction* in life expectancy she would experience if she fell ill. These data are also reported in columns (4) and (5) of Table 3.

APPENDIX (FOR ONLINE PUBLICATION ONLY)

Appendix A provides proofs for lemmas and propositions stated in the main text. Appendix B provides supporting details for the data employed in the numerical models presented in Section IV, and Appendix C presents derivations for those models. Finally, Appendix D provides derivations for the value of statistical life and the value of statistical illness for a fully annuitized consumer when mortality is stochastic.

A. Mathematical proofs of results from main text

Proof of Lemma 1:

Let $V(t, W(t), j)$ be taken as given (exogenous). Consider the deterministic optimization problem:

$$V(0, W_0, i) = \max_{c_i(t)} \left\{ \int_0^T e^{-\rho t} \tilde{S}(i,t) \left(u(c_i(t), q_i(t)) + \sum_{j \neq i} \lambda_{ij}(t) V(t, W(t), j) \right) dt \right\}$$

subject to

$$\frac{\partial W(t)}{\partial t} = rW(t) + m_i(t) - c_i(t)$$

Denote the optimal value-to-go as

$$\tilde{V}(u, W(u), i) = \max_{c_i(t)} \left\{ \int_u^T e^{-\rho t} \tilde{S}(i,t) \left(u(c_i(t), q_i(t)) + \sum_{j \neq i} \lambda_{ij}(t) V(t, W(t), j) \right) dt \right\}$$

Setting $\tilde{V}(t, W(t), i) = e^{-\rho t} \tilde{S}(i,t) V(t, W(t), i)$ then demonstrates that $V(\cdot)$ satisfies the HJB (12) for i. See Parpas and Webster (2013) for additional details.

QED

Proof of Lemma 3:

The proof proceeds by induction on $i \leq n$. For the base case $i = n$, in which no state transitions are possible, the solution to the costate equation (given in the main text) simplifies to:[22]

$$\begin{aligned} p_\tau^{(n)} &= \theta^{(n)} e^{-r\tau} = \exp\left\{ -\int_0^\tau \rho + \overline{\mu}_n(s) ds \right\} u_c(c_n(\tau), q_n(\tau)) \\ &= \theta^{(n)} e^{-rt} e^{-r(\tau - t)} \\ &= p_t^{(n)} e^{-r(\tau - t)} \\ &= \exp\left\{ -\int_0^t \rho + \overline{\mu}_n(s) ds \right\} u_c(c_n(t), q_n(t)) e^{-r(\tau - t)} \end{aligned}$$

This then implies that

[22] When no transitions are possible, this reduces to the deterministic model outlined in Section II.B.

$$u_c(c_n(t), q_n(t)) = e^{r(\tau-t)} e^{-\rho(\tau-t)} \exp\left\{-\int_t^\tau \overline{\mu}_n(s) ds\right\} u_c(c_n(\tau), q_n(\tau))$$

which shows that the lemma holds for $i = n$.

For the induction step, suppose the lemma is true for $j > i$, $1 \le i \le n-1$. For any subinterval $[0, \tau]$, the solution of the costate equation can be written as:

$$p_t^{(i)} = \left[\int_t^\tau e^{(r-\rho)s} \exp\left\{-\int_0^s \overline{\mu}_i(u) + \sum_{j\neq i} \lambda_{ij}(u)\, du\right\} \sum_{j\neq i} \lambda_{ij}(s) \frac{\partial V(s, W(s), j)}{\partial W(s)} ds\right] e^{-rt} + \theta(\tau, i) e^{-rt} \tag{A1}$$

where $\theta(\tau, i)$ is a constant that depends on the choice of τ and i. (Take the derivative of $p_t^{(i)}$ with respect to t to verify.) Evaluating equation (A1) at $t = \tau$ and combining with equation (14) from the main text yields:

$$p_\tau^{(i)} = \theta(\tau, i) e^{-r\tau} = \exp\left\{-\int_0^\tau \rho + \overline{\mu}_i(s) + \sum_{j\neq i} \lambda_{ij}(s)\, ds\right\} u_c\big(c_i(\tau), q_i(\tau)\big)$$

which implies

$$\theta(\tau, i) = e^{(r-\rho)\tau} \exp\left\{-\int_0^\tau \overline{\mu}_i(s) + \sum_{j\neq i} \lambda_{ij}(s)\, ds\right\} u_c\big(c_i(\tau), q_i(\tau)\big) \tag{A2}$$

Also, from equation (14) we know that:

$$p_t^{(i)} = \exp\left\{-\int_0^t \rho + \overline{\mu}_i(s) + \sum_{j\neq i} \lambda_{ij}(s)\, ds\right\} u_c\big(c_i(t), q_i(t)\big)$$

Plugging equations (14) and (A2) into equation (A1) yields:

$$\begin{aligned} u_c\big(c_i(t), q_i(t)\big) \exp\left\{-\int_0^t \rho + \overline{\mu}_i(s) + \sum_{j\neq i} \lambda_{ij}(s)\, ds\right\} \\ = \left[\int_t^\tau e^{(r-\rho)s} \exp\left\{-\int_0^s \overline{\mu}_i(u) + \sum_{j\neq i} \lambda_{ij}(u)\, du\right\} \sum_{j\neq i} \lambda_{ij}(s) \frac{\partial V(s, W(s), j)}{\partial W(s)} ds\right] e^{-rt} \\ + e^{-rt} e^{(r-\rho)\tau} \exp\left\{-\int_0^\tau \overline{\mu}_i(s) + \sum_{j\neq i} \lambda_{ij}(s)\, ds\right\} u_c\big(c_i(\tau), q_i(\tau)\big) \end{aligned}$$

Since $\frac{\partial V(s,W(s),j)}{\partial W(s)} = u_c\big(c_j(s), q_j(s)\big)$, we obtain:

$$\begin{aligned} u_c\big(c_i(t), q_i(t)\big) = \int_t^\tau e^{(r-\rho)(s-t)} \exp\left\{-\int_t^s \overline{\mu}_i(u) + \sum_{j\neq i} \lambda_{ij}(u)\, du\right\} \sum_{j\neq i} \lambda_{ij}(s) u_c\big(c_j(s), q_j(s)\big)\, ds \\ + e^{(r-\rho)(\tau-t)} \exp\left\{-\int_t^\tau \overline{\mu}_i(s) + \sum_{j\neq i} \lambda_{ij}(s)\, ds\right\} u_c\big(c_i(\tau), q_i(\tau)\big) \end{aligned}$$

$$
\begin{aligned}
&= \int_t^\tau e^{(r-\rho)(s-t)} \exp\left\{-\int_t^s \overline{\mu}_i(u) + \sum_{j\neq i} \lambda_{ij}(u)\, du\right\} \sum_{j\neq i} \lambda_{ij}(s) \mathbb{E}\left[e^{(r-\rho)(\tau-s)} \exp\left\{-\int_s^\tau \mu(s)ds\right\} u_c\big(c_{y_t}(\tau), q_{y_t}(\tau)\big)\middle| Y_s = j\right] ds \\
&\quad + e^{(r-\rho)(\tau-t)} \exp\left\{-\int_t^\tau \overline{\mu}_i(s) + \sum_{j\neq i} \lambda_{ij}(s)\, ds\right\} u_c\big(c_i(\tau), q_i(\tau)\big) \\
&= \mathbb{E}\left[e^{(r-\rho)(\tau-s)} \exp\left\{-\int_t^\tau \mu(s)ds\right\} u_c\big(c_{y_\tau}(\tau), q_{y_\tau}(\tau)\big)\middle| Y_t = i\right]
\end{aligned}
$$

where the second equality follows from the induction hypothesis.

QED

Proof of Proposition 4:

Choosing once again the Dirac delta function for $\delta(\cdot)$ in **Lemma 2** yields

$$
\begin{aligned}
\left.\frac{\partial \mathbb{E}U}{\partial \varepsilon}\right|_{\varepsilon=0} &= \int_0^T \left[e^{-\rho t} \tilde{S}(i,t) \left(u\big(c_i(t), q_i(t)\big) + \sum_{j\neq i} \lambda_{ij}(t)\, V(t, W(t), j) \right)\right] dt \\
&= \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) u\big(c_{y_t}(t), q_{y_t}(t)\big) dt \middle| Y_0 = i\right]
\end{aligned}
$$

Dividing the result by the marginal utility of wealth at time $t = 0$ then yields the value of statistical life given by equation (15):

$$
VSL(i) = \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) \frac{u\big(c_{y_t}(t), q_{y_t}(t)\big)}{u\big(c_{Y_0}(0), q_{Y_0}(0)\big)} dt \middle| Y_0 = i\right] = \int_0^T e^{-rt} v(i,t) dt
$$

Applying **Lemma 3** for $t = 0$ allows us to rewrite VSL as

$$
\begin{aligned}
VSL(i) &= \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) \frac{u\big(c_{y_t}(t), q_{y_t}(t)\big)}{\mathbb{E}\left[e^{(r-\rho)t} \exp\left\{-\int_0^t \mu(s) ds\right\} u_c\big(c_{Y_t}(t), q_{Y_t}(t)\big)\middle| Y_0\right]} dt \middle| Y_0 = i\right] \\
&= \mathbb{E}\left[\int_0^T e^{-rt} \frac{S(t) u\big(c_{y_t}(t), q_{y_t}(t)\big)}{\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s) ds\right\} u_c\big(c_{Y_t}(t), q_{Y_t}(t)\big)\middle| Y_0\right]} dt \middle| Y_0 = i\right]
\end{aligned}
$$

which by exchanging expectation and integration shows that the value of a life-year, $v(i,t)$, is equal to the expected utility of consumption normalized by the expected marginal utility of consumption:

$$
v(i,t) = \frac{\mathbb{E}\left[S(t) u\left(c_{y_t}(t), q_{y_t}(t)\right)\middle| Y_0 = i\right]}{\mathbb{E}\left[S(t) u_c\left(c_{y_t}(t), q_{y_t}(t)\right)\middle| Y_0 = i\right]}
$$

QED

Proof of Proposition 5:

The proposition assumes there are $n = 2$ states, with $\overline{\mu}_2(s) > \overline{\mu}_1(s) \forall s$. That is, health in state 2 is strictly worse than health in state 1. For simplicity, we abstract from quality of life, $q(t)$. Without loss of generality, we will prove the proposition for the case where the consumer transitions from state 1 to state 2 at time $t = 0$.

For state 2, the solution to the costate equation is:

$$p_t^{(2)} = \theta^{(2)} e^{-rt}$$

and from the first-order condition (14) we obtain:

$$p_t^{(2)} = e^{-\rho t} \exp\left\{-\int_0^t \overline{\mu}_2(s)ds\right\} u_c\big(c_2(t)\big)$$

The two preceding equations imply that

$$u_c\big(c_2(t)\big) = \theta^{(2)} e^{(\rho - r)t} \exp\left\{\int_0^t \overline{\mu}_2(s)ds\right\}$$

For state 1, the costate equation is:

$$\begin{aligned}
\dot{p}_t^{(1)} &= -p_t^{(1)} r - e^{-\rho t} \exp\left\{-\int_0^t \overline{\mu}_1(s) + \lambda_{12}(s)ds\right\} \lambda_{12}(t) \frac{\partial V(t, W(t), 2)}{\partial W(t)} \\
&= -p_t^{(1)} r - e^{-\rho t} \exp\left\{-\int_0^t \overline{\mu}_1(s) + \lambda_{12}(s)ds\right\} \lambda_{12}(t) u_c\big(c_2(t)\big) \\
&= -p_t^{(1)} r - e^{-rt} \exp\left\{-\int_0^t \lambda_{12}(s)ds\right\} \lambda_{12}(t) \theta^{(2)} \exp\left\{\int_0^t \overline{\mu}_2(s) - \overline{\mu}_1(s)ds\right\}
\end{aligned} \tag{A3}$$

Before proceeding, we first prove the following two lemmas.

Appendix Lemma A1:

There exists a $t \in [0, T]$ such that

$$p_t^{(1)} \geq \exp\left\{-\int_0^t \lambda_{12}(s)ds\right\} p_t^{(2)}$$

Proof of Appendix Lemma A1:

Suppose by way of contradiction that $p_t^{(1)} < \exp\left\{-\int_0^t \lambda_{12}(s)ds\right\} p_t^{(2)}$ $\forall t \in [0, T]$. *Then, since* $\overline{\mu}_2(s) > \overline{\mu}_1(s)$ *we have*

$$e^{-\rho t} \exp\left\{-\int_0^t \overline{\mu}_2(s)ds\right\} p_t^{(1)} < e^{-\rho t} \exp\left\{-\int_0^t \overline{\mu}_1(s)ds\right\} \exp\left\{-\int_0^t \lambda_{12}(s)ds\right\} p_t^{(2)}$$

Rearranging then yields

$$u_c(c_1(t)\,) = \frac{p_t^{(1)}}{e^{-\rho t} \exp\left\{-\int_0^t \overline{\mu}_1(s)ds\right\} \exp\left\{-\int_0^t \lambda_{12}(s)ds\right\}} < \frac{p_t^{(2)}}{e^{-\rho t} \exp\left\{-\int_0^t \overline{\mu}_2(s)ds\right\}} = u_c(c_2(t))$$

which implies $c_2(t) < c_1(t) \forall t$. *But then we have a contradiction:* $c_2(t)$ *cannot be an optimal consumption plan because the feasible consumption plan* $c_1(t)$ *strictly dominates* $c_2(t)$.

QED

Appendix Lemma A2:

$$p_0^{(1)} > \theta^{(2)} = p_0^{(2)}$$

Proof of Appendix Lemma A2:

Define

$$g(t) = \exp\left\{-\int_0^t r + \lambda_{12}(s)ds\right\}\theta^{(2)} = \exp\left\{-\int_0^t \lambda_{12}(s)ds\right\}p_t^{(2)}$$

Differentiating with respect to t yields

$$\begin{aligned}\dot{g}(t) &= -g(t)r - \exp\left\{-rt - \int_0^t \lambda_{12}(s)ds\right\}\lambda_{12}(t)\theta^{(2)} \\ &= \phi(g(t), t)\end{aligned}$$

*Combining this result with equation (*A3*) then yields the following inequality:*

$$\dot{p}_t^{(1)} < \phi\left(p_t^{(1)}, t\right)$$

Suppose by way of contradiction that $p_0^{(1)} < \theta^{(2)} = g(0)$. *Then by standard comparison arguments for ordinary differential equations, we have* $p_t^{(1)} < g(t) = \exp\left\{-\int_0^t \lambda_{12}(s)ds\right\}p_t^{(2)}\ \forall t \in [0, T]$, *which is a contradiction to the result from* ***Appendix Lemma A1****.*

QED

Thus, we have

$$u_c\big(c_1(0)\big) = p_0^{(1)} > p_0^{(2)} = u_c(c_2(0))$$

which implies

$$c_2(0) > c_1(0)$$

QED

Proof of Proposition 6:

Without loss of generality, consider the case $t = 0$. From **Proposition 5** and **Appendix Lemmas A1 and A2**, it is clear that $c_1(t)$ and $c_2(t)$ are decreasing, $c_2(0) > c_1(0)$, $c_2(t) \geq c_1(t)$ for $t \leq t_0$, and $c_2(t) \leq c_1(t)$ for $t > t_0$. Making use of the assumption that no state transitions occur for $t > 0$, we have that

$$\begin{aligned}VSL(2,0) &= \int_0^T e^{-rt}\frac{S_2(t)u\big(c_2(t)\big)}{S_2(t)u_c\big(c_2(t)\big)}dt \\ &= \int_0^T e^{-rt}\frac{u\big(c_2(t)\big)}{u_c\big(c_2(t)\big)}dt\end{aligned}$$

and

$$VSL(1,0) = \int_0^T e^{-rt}\frac{u\big(c_1(t)\big)}{u_c\big(c_1(t)\big)}dt$$

Let $Y(x) = \frac{u(x)}{u_c(x)}$. Under the stated assumptions, we have that

$$Y'(x) = 1 - \frac{u(x)u_{cc}(x)}{\big(u_c(x)\big)^2} > 0,$$
$$Y''(x) = \frac{2\big(u_{cc}(x)\big)^2 u(x) - u_c^2(x)u_{cc}(x) - u_c(x)u(x)u_{ccc}(x)}{\big(u_c(x)\big)^3} > 0$$

Employing Taylor's theorem then yields:

$$\begin{aligned}
VSL(2,0) &= \int_0^T e^{-rt} Y\big(c_2(t)\big)\, dt \\
&= \int_0^T e^{-rt} \left[Y\big(c_1(t)\big) + [c_2(t) - c_1(t)]Y'\big(c_1(t)\big) + \underbrace{\frac{1}{2}[c_2(t) - c_1(t)]^2 Y''\big(\xi(t)\big)}_{>0} \right] dt \\
&> \int_0^T e^{-rt}\, Y\big(c_1(t)\big) dt + \int_0^{t_0} e^{-rt}\, Y'\big(c_1(t)\big) \underbrace{[c_2(t) - c_1(t)]}_{\geq 0}\, dt \\
&\qquad + \int_{t_0}^T e^{-rt}\, Y'\big(c_1(t)\big) \underbrace{[c_2(t) - c_1(t)]}_{\leq 0}\, dt \\
&> \int_0^T e^{-rt}\, Y\big(c_1(t)\big) dt + \int_0^{t_0} e^{-rt}\, Y'\big(c_1(t_0)\big)[c_2(t) - c_1(t)] dt \\
&\qquad + \int_0^{t_0} e^{-rt}\, Y'\big(c_1(t_0)\big)[c_2(t) - c_1(t)] dt \\
&= \int_0^T e^{-rt}\, Y\big(c_1(t)\big) dt + Y'\big(c_1(t_0)\big) \underbrace{\left[\int_0^T e^{-rt}\, c_2(t) dt - \int_0^T e^{-rt}\, c_1(t) dt \right]}_{=0} \\
&= \int_0^T e^{-rt}\, Y\big(c_1(t)\big) dt \\
&= VSL(1,0)
\end{aligned}$$

where the final step follows from the budget constraint.

QED

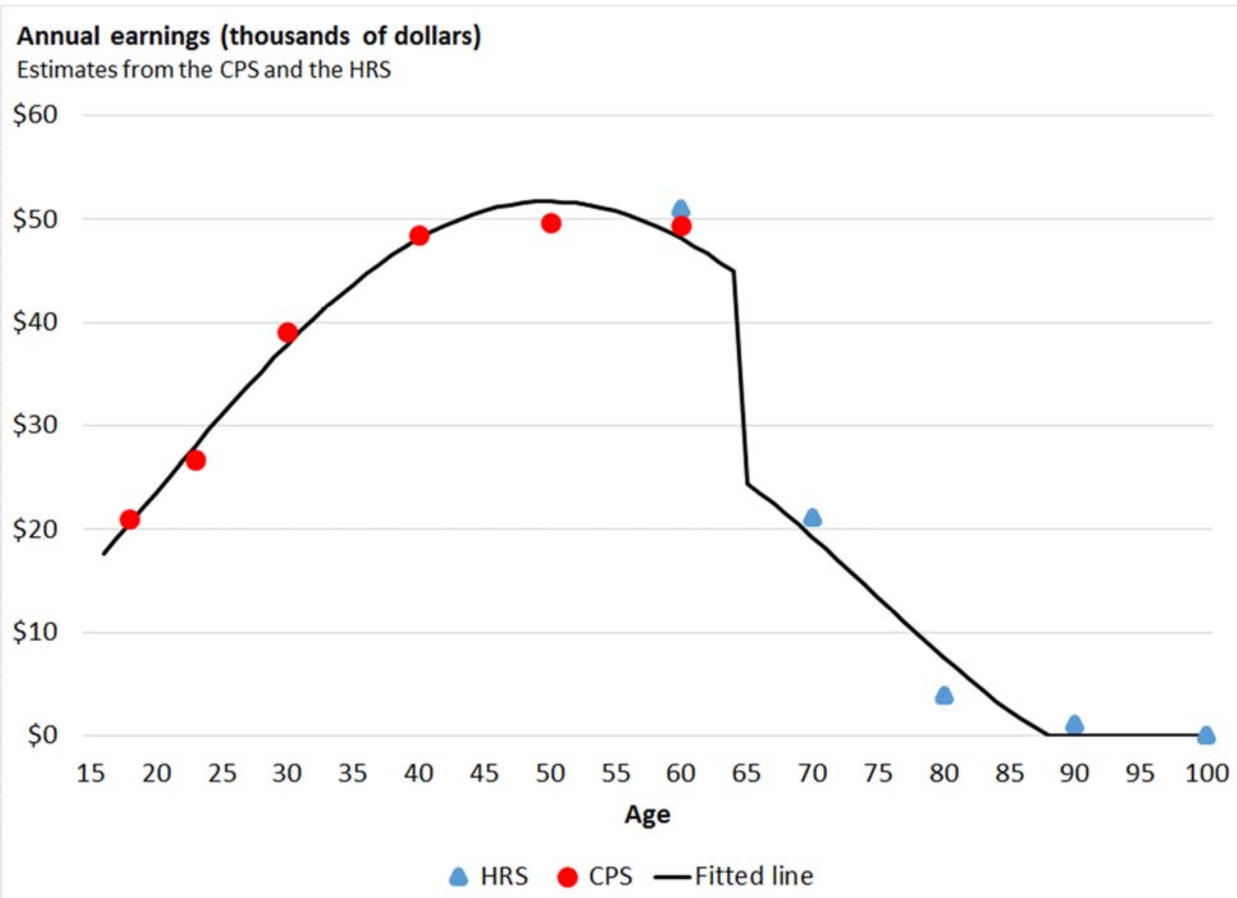
B. Data

B1. Earnings

We obtain earnings data for employed individuals under the age of 65 from the 2016 Current Population Survey (CPS).[23] We also obtain earnings data for respondents over the age of 55 from the 2014 Health and Retirement Survey (HRS). For both surveys, the data represent earnings before taxes and other deductions, and include wages, salaries, and tips. The HRS earnings data also include self-employment income. (The CPS data exclude self-employed individuals.)

The CPS earnings data are binned into the following age groups: 16-19, 20-24, 25-34, 35-44, 45-54, and 55-64. We collapse the HRS earnings data into the following age groups: 55-64, 65-74, 75-84, 85-94, and 95-104. The resulting estimates are plotted in Appendix Figure 1. We smooth the data by fitting it to a quartic polynomial, and include an indicator variable for ages over 65. The dependent variable in the regression is the CPS earnings estimate for ages under 65, and the HRS estimate for ages over 65. Finally, we constrain the fitted prediction to be non-negative.

Appendix Figure B1. Annual earnings estimates from CPS and HRS



Notes: Figure plots annual earnings by midpoint of age group as estimated by the 2016 Current Population Survey (CPS) for respondents under age 65 and the 2014 Health and Retirement Survey (HRS) for respondents over age 55.

---

[23] These data are available at http://data.bls.gov/pdq/querytool.jsp?survey=le.

The fitted line corresponds to a regression of annual earnings on a quartic polynomial in age and an indicator equal to 1 for ages 65 and over. The dependent variable, annual earnings, corresponds to CPS estimates for ages under 65 and HRS estimates for ages over 65.

B2. Future Elderly Model (FEM)

The FEM follows Americans aged 50 years and older and projects their health and medical spending over time. A complete technical document detailing the FEM is available online.[24] The FEM is a microsimulation that follows the evolution of individual-level health trajectories and economic outcomes, rather than the average or aggregate characteristics of a cohort. The FEM has three core modules. The first is the Replenishing Cohorts module, which predicts economic and health outcomes of new cohorts of 50-year-olds with data from the Panel Study of Income Dynamics (PSID), and incorporates trends in disease and trends in other outcomes based on data from external sources, such as National Health Interview Survey and the American Community Survey. This module generates cohorts as the simulation proceeds, so that we can measure outcomes for the age 50+ population in any given year.

The second component is the Health Transition module, which uses the longitudinal structure of the Health and Retirement Survey (HRS) to calculate transition probabilities across various health states, including chronic conditions, functional status, body-mass index and mortality, using linear and nonlinear multivariate regression models. These transition probabilities depend on a battery of predictors: age, sex, education, race, ethnicity, smoking behavior, marital status, employment and health conditions. Baseline factors are also controlled for using a series of initial health variables measured at age 50. FEM transitions produce a large set of simulated outcomes, including diabetes, high-blood pressure, heart disease, cancer (except skin cancer), stroke or transient ischemic attack, and lung disease (either or both chronic bronchitis and emphysema), disability, and body-mass index. Disability is measured by limitations in instrumental activities of daily living, activities of daily living, and residence in a nursing home. This dynamic simulation method has undergone extensive benchmarking and validation.

Finally, the Policy Outcomes module combines individual-level outcomes into aggregate outcomes, such as medical care costs (Medicare, Medicaid and Private), federal, state and property taxes, Social Security expenditures and contributions. Individual health spending is predicted with regard to health status (chronic conditions and functional status), demographics (age, sex, race, ethnicity and education), nursing home status and mortality. Estimates are based on spending data from the Medical Expenditure Panel Survey for individuals aged 64 and younger and the Medicare Current Beneficiary Survey for individuals aged 65 and older, who constitute the bulk of the Medicare population. This module has been comprehensively tested against national aggregates.

An example of how the three modules interact is as follows. For year 2014, the model begins with the population of Americans aged 50 and older based on nationally representative data from the HRS. Individual-level health and economic outcomes for the next two years are predicted using the Policy Outcomes module. The cohort is then aged two years using the Health Transition Module. Aggregate health and functional status outcomes for those years are then calculated. At that point, a new cohort of 50-year-olds is introduced into the 2016 population using the Replenishing Cohort module, and they join those who survived from 2014 to 2016. This forms the age 50+ population for 2016. The transition model is then applied to this population. The same process is repeated until reaching the last year of the simulation.

[24] A complete technical description is available at roybalhealthpolicy.usc.edu/fem/technical-specifications/.

C. Derivations for numerical models

Appendix C1 provides details regarding the implementation of the deterministic mortality model employed in Section IV.B, and explains how it is used to derive the aggregate insurance value of Social Security. This model is estimated numerically using standard dynamic programming methods.

Appendix C2 provides a derivation of the stochastic mortality model employed in Section IV.C. This model is solved analytically and thus provides exact solutions.

C1. Deterministic mortality

The value function is defined as:

$$V\big(t, W(t)\big) = \max_{\{c(t)\}} \sum_{s=t}^{T} e^{-\rho(s-t)} S_t(s) u(c(s))$$

We can use the value function to rewrite the optimization problem as a recursive Bellman equation:

$$V\big(t, W(t)\big) = \max_{\{c(t)\}} u\big(c(t)\big) + \frac{1 - d(t)}{e^{\rho}} V\big(t + 1, W(t + 1)\big)$$

Because the problem is finite, we can work backwards from the final period. We discretize the state space into $N_w = 2{,}000$ points evenly distributed across the interval $[0, W_{max}]$. Let that set of values be $\{W_n\}$. Define $g_t\big(W(t)\big) = W(t + 1)$ as a mapping from the current wealth state, $W(t)$, to the optimal wealth state in the following period, $W(t + 1)$

It is clear that the consumer should consume all her wealth in the final period, i.e., $g_T\big(W(T)\big) = 0$ for all $W(T) \in \{w_n\}$. This implies that $V\big(T, W(T)\big) = u\big(W(T) + y(T)\big)$ for all $W(T) \in \{w_n\}$.

Next, we calculate $V(T - 1, w_{T-1}) = \max_{g(w_{T-1})=w_T} u(W(T - 1) + y(T - 1) - W(T)/e^r) + \frac{1-d(t+1)}{e^{\rho}} V\big(T, W(T)\big)$. In other words, for each $W(T - 1) \in \{w_n\}$, we calculate the optimal $V\big(T - 1, W(T - 1)\big)$ by determining which choice of $g_{T-1}\big(W(T - 1)\big) = W(T) \in \{w_n\}$ will maximize utility. This algorithm is then repeated for $t = T - 2, T - 3, \ldots, 1$.

Given the initial condition, w_1, we can then employ our results to calculate $W(2) = g_1\big(W(1)\big)$, $W(3) = g_2\big(W(2)\big)$,…, $W(T)$. Period consumption, $c(t)$, is then calculated using the equation for the budget constraint. Finally, we the analytical formulas derived in the main text to calculate the value of statistical life.

Insurance value of Social Security

We calculate the insurance value of Social Security at all ages by estimating its wealth equivalence. That is, we follow Mitchell et al. (1999) and estimate the amount of wealth, W^*, required to equalize the utilities of a non-annuitized individual and an individual with Social Security. In other words, we solve for compensating wealth at age t, $W^*(t)$, such that $V\big(t, W(t) + W^*(t)\big) = V^{SS}\big(t, W^{SS}(t)\big)$. Wealth for a non-annuitized individual, $W(t)$, and wealth for an individual with Social Security, $W^{SS}(t)$, are calculated by the deterministic model for the first two policy scenarios discussed in the main text.

We solve for $W^*(t)$ by applying a numerical search algorithm. We estimate that, at age 65, having access to Social Security is equivalent to an increase in wealth of 16.7 percent for a non-annuitized individual. By way of comparison, Mitchell et al. (1999) estimate the before-tax value of *full* (complete) annuitization at age 65 to be 37.4 percent of wealth, using the same parameters for risk aversion, interest rate, and the discount rate.

The aggregate insurance value of Social Security is then calculated by aggregating over the 2015 US population:

$$Aggregate\ Value\ SS = \sum_{a=0}^{110} W^*(a) f(a)$$

C2. Stochastic mortality

We focus on the case where the consumer does not have access to annuities. We ignore income, and assume that all of consumer's wealth is available at time $t = 0$. This will allow us to generate an analytic solution to the consumer's problem, given by:

$$\max_{\{c_t\}} \mathbb{E}_0 \left[\sum_{t=0}^{T} e^{-\rho t} S_0(t) u\left(c(t), q_{Y_t}(t)\right) + e^{-\rho(t+1)} \left(\left(S_0(t) - S_0(t+1)\right) u(W(t+1), b_t) \right) \right]$$

where

$$W(0) \text{ given,}$$

$$W(t) \geq 0,$$

$$W(t+1) = \left(W(t) - c(t)\right) e^{r(t,Y_t)}$$

Here, Y_t denotes the consumer's health state at time t, and we allow the interest rate to depend on it so as to model health-related wealth shocks. Of course, a constant interest rate $r(t,i) = r$ is included as a special case. The utility function is

$$u(c,q) = q \frac{c^{1-\gamma}}{1-\gamma} - \frac{\underline{c}^{1-\gamma}}{1-\gamma}$$

where $\underline{c}$ is the subsistence level of consumption for a healthy person. The parameter b_t measure the bequest motive. Because optimal consumption is unaffected by affine transformations of utility, we will assume $u(c,q) = qc^{1-\gamma}/(1-\gamma)$ when solving the model for consumption.

Define the value function

$$V(t, W(t), Y_t) = \max_{\{c_s\}} \mathbb{E} \left[\sum_{s=t}^{T} e^{-\rho(s-t)} S_t(s) u\left(c(s), q_{Y_s}(s)\right) + e^{-\rho(s+1-t)} \left(S_t(s) - S_t(s+1)\right) u(W(s+1), b_s) \middle| Y_t \right]$$

subject to

$$W(s+1) = \left(W(s) - c(s)\right) e^{r(s,Y_s)}, s > t, W(s) \geq 0$$

Then we obtain the following Bellman equation:

$$V(t,w,i) = \max_{c_t}\left\{u\big(c(t),q_i(t)\big) + e^{-\rho}\overline{d}_i(t)u\left(\big(w-c(t)\big)e^{r(t,i)},b_t\right)\right.$$
$$\left. + e^{-\rho}\left(1-\overline{d}_i(t)\right)\sum_{j=1}^{n} p_{ij}(t)V\big(t+1,\big(w-c(t)\big)e^{r(t,i)},j\big)\right\}$$

Appendix Proposition C1:

The value function and the optimal consumption level satisfy

$$V(t,w,i) = \frac{w^{1-\gamma}}{1-\gamma}K_{t,i},$$
$$c^*(t,w,i) = w \cdot c_{t,i}$$

where

$$c_{t,i} = \left[1 + e^{-r(t,i)}\left(\frac{e^{r(t,i)}\left[\overline{d}_i(t)b_t + \left(1-\overline{d}_i(t)\right)\left(\sum_{j=1}^{n} p_{ij}(t)K_{t+1,j}\right)\right]}{e^{\rho}q_i(t)}\right)^{\frac{1}{\gamma}}\right]^{-1}, t < T,$$
$$c_{T,i} = \left[1 + e^{-r(T,i)}\left(\frac{e^{r(T,i)}b_T}{e^{\rho}q_i(T)}\right)^{\frac{1}{\gamma}}\right]^{-1}$$

and $K_{t,i}$ satisfies the recursion:

$$K_{t,i} = \left[q_i(t)^{\frac{1}{\gamma}} + e^{-r(t,i)}\left[e^{r(t,i)-\rho}\left(\overline{d}_i(t)b_t + \left(1-\overline{d}_i(t)\right)\left(\sum_{j=1}^{n} p_{ij}(t)K_{t+1,j}\right)\right)\right]^{\frac{1}{\gamma}}\right]^{\gamma}, t < T,$$
$$K_{T,i} = \left[q_i(T)^{\frac{1}{\gamma}} + e^{-r(T,i)}\left(e^{r(T,i)-\rho}b_T\right)^{\frac{1}{\gamma}}\right]^{\gamma}$$

Proof of Appendix Proposition C1: see end of appendix C

When calculating VSL, we incorporate subsistence consumption back into the utility function. We then obtain for the value function:

$$V(0,w,i) = \sum_{t=0}^{T} e^{-\rho t}\mathbb{E}_{0,i}\left[\exp\left\{-\int_0^t \mu(s)ds\right\}\left(q_{Y_t}(t)\frac{c(t)^{1-\gamma}}{1-\gamma} - \frac{\underline{c}^{1-\gamma}}{1-\gamma}\right)\right]$$
$$+ \underbrace{e^{-\rho(t+1)}\mathbb{E}_{0,i}\left[\left(\exp\left\{-\int_0^t \mu(s)ds\right\} - \exp\left\{-\int_0^{t+1}\mu(s)ds\right\}\right)\left(b_t\frac{W(t+1)^{1-\gamma}}{1-\gamma} - \frac{\underline{c}^{1-\gamma}}{1-\gamma}\right)\right]}_{*}$$

In specifications without the bequest motive, the second term (*) is dropped. Rearranging yields:

$$V(0,w,i) = \sum_{t=1}^{T} e^{-\rho t}\mathbb{E}_{0,i}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_{Y_t}(t)\frac{c(t)^{1-\gamma}}{1-\gamma}\right]$$
$$+ e^{-\rho(t+1)} b_t \mathbb{E}_{0,i}\left[\left(\exp\left\{-\int_0^t \mu(s)ds\right\} - \exp\left\{-\int_0^{t+1} \mu(s)ds\right\}\right)\frac{W(t+1)^{1-\gamma}}{1-\gamma}\right]$$
$$= \frac{1}{1-\gamma}\left[w^{1-\gamma}K_{0,i} - \underline{c}^{1-\gamma}\left[1 + e^{-\rho} \underbrace{\sum_{t=0}^{T} e^{-\rho t}\,\mathbb{E}_{0,i}\left[\exp\left\{-\int_0^t \mu(s)ds\right\}\right]}_{\text{life expect. in state } i,\ \text{discounted at rate } \rho}\right]\right]$$

We can then calculate VSL in state i using the following formula:

$$VSL_i = \frac{V(0,w,i)}{u_c\big(c_i(0), q_i(0)\big)} = \frac{V(0,w,i)}{V_w(0,w,i)}$$

When bequests are absent and $r(t,i) = r$, we drop the term (*), and the theory presented in the main text yields the following expression for VSL:

$$VSL_i = \mathbb{E}\left[\sum_{t=0}^{T} \exp\left\{-\int_0^t \rho + \mu(s)ds\right\}\frac{u\left(c(t), q_{Y_t}(t)\right)}{u_c\left(c(0), q_{Y_0}(0)\right)}\middle| Y_0 = i\right]$$
$$= \sum_{t=0}^{T} e^{-rt}\frac{\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} u\left(c(t), q_{Y_t}(t)\right)\middle| Y_0\right]}{\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} u_c\left(c(t), q_{Y_t}(t)\right)\middle| Y_0\right]}$$
$$= \sum_{t=0}^{T} e^{-rt}\frac{\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\}\left(q_{Y_t}(t)\frac{c(t)^{1-\gamma}}{1-\gamma} - \frac{\underline{c}^{1-\gamma}}{1-\gamma}\right)\middle| Y_0\right]}{\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_{Y_t}(t)c(t)^{-\gamma}\middle| Y_0\right]}$$

or

$$VSL = \frac{1}{1-\gamma}\sum_{t=0}^{T} e^{-rt}\underbrace{\frac{\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_{Y_t}(t)c(t)^{1-\gamma}\middle| Y_0\right] - \underline{c}^{1-\gamma}\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\}\middle| Y_0\right]}{\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_{Y_t}(t)c(t)^{-\gamma}\middle| Y_0\right]}}_{\overline{v}(t)}$$

To evaluate this expression for VSL, we will make use of the following lemma.

Appendix Lemma C2: Let $W_{t,j}(\Psi) = \mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} W(t)^{\Psi}\mathbf{1}\{Y_t = j\}\middle| Y_0\right]$ for $Y \in (1,\infty)$. Then $W_{t,j}(\Psi)$ satisfies the following recursion:

$$W_{0,Y_0}(\Psi) = w_0^{\Psi}, W_{0,i}(\Psi) = 0, i \neq Y_0,$$
$$W_{t+1,j}(\Psi) = e^{r\Psi}\sum_{k=1}^{n} W_{t,k}(\Psi)\left(1 - c_{t,k}\right)^{\Psi}\left(1 - \overline{d}_k(t)\right)p_{k,j}(t)$$

Proof of Appendix Lemma C2: see end of appendix C

Note that for $\Psi = 0$, the expression $\sum_{j=1}^{n} W_{t,j}(0) = \mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\}\middle| Y_0\right]$ is simply the t-year survival probability. Using this **Appendix Lemma C2**, we obtain:

Appendix Proposition C3:

$$VSL_{Y_0} = \frac{1}{1-\gamma}\sum_{t=0}^{T} e^{-rt} \underbrace{\frac{\sum_{j=1}^{n} q_j(t) c_{t,j}^{1-\gamma} W_{t,j}(1-\gamma) - \underline{c}^{1-\gamma} \sum_{j=1}^{n} W_{t,j}(0)}{\sum_{j=1}^{n} q_j(t) c_{t,j}^{-\gamma} W_{t,j}(-\gamma)}}_{\overline{v}(t)}$$

Proof of Appendix Proposition C3: see end of appendix C

We also immediately obtain the following corollary:

Appendix Corollary C4:

$$\begin{aligned} VSI_{i,j} &= VSL_i - VSL_j \frac{q_j(0) c_{0,j}^{-\gamma}}{q_i(0) c_{0,i}^{-\gamma}} \\ &= VSL_i - \left(\frac{q_j(0)}{q_i(0)}\right)\left(\frac{c_{0,i}}{c_{0,j}}\right)^{\gamma} VSL_j \end{aligned}$$

Proofs for Appendix C
Proof of Appendix Proposition C1:

The proof proceeds by induction on $t \leq T$. For the base case $t = T$, note that $\overline{d}_i(t) = 1$, so that the first-order condition from the Bellman equation gives:

$$q_i(T)c(T)^{-\gamma} = e^{r(T,i)-\rho} b_T \big(w - c(T)\big)^{-\gamma} e^{-r(T,i)\gamma}$$

This implies that

$$c(T) = \frac{w e^{r(T,i)} e^{\frac{(\rho - r(T,i))}{\gamma}} \left(\frac{q_i(T)}{b_T}\right)^{\frac{1}{\gamma}}}{1 + e^{r(T,i)} e^{\frac{(\rho - r(T,i))}{\gamma}} \left(\frac{q_i(T)}{b_T}\right)^{\frac{1}{\gamma}}}$$
$$= w \underbrace{\left[1 + e^{-r(T,i)} \left(\frac{e^{r(T,i)} b_T}{e^{\rho} q_i(T)}\right)^{\frac{1}{\gamma}}\right]^{-1}}_{c_{T,i}}$$

So that:

$$V(T,w,i) = \frac{w^{1-\gamma}}{1-\gamma}\Big(q_i(T) c_{T,i}^{1-\gamma} + e^{-\rho} b_T e^{r(T,i)(1-\gamma)} \big(1 - c_{T,i}\big)^{1-\gamma}\Big)$$
$$= \frac{e^{-\rho} e^{r(T,i)(1-\gamma)}}{\left[b_T^{\frac{1}{\gamma}} + e^{r(T,i)} e^{\frac{(\rho - r(T,i))}{\gamma}} q_i(T)^{\frac{1}{\gamma}}\right]^{-\gamma}}$$
$$= \left[q_i(T)^{\frac{1}{\gamma}} + e^{-r(T,i)} \big(e^{(r(T,i)-\rho)} b_T\big)^{\frac{1}{\gamma}}\right]^{\gamma}$$

For the induction step, suppose the proposition is true for case $t + 1$. We have

$$V(t,w,i) = \max_c \left\{ q_i(t)\frac{c^{1-\gamma}}{1-\gamma} + b_t e^{-\rho} \overline{d}_i(t) \frac{\big((w-c)e^{r(t,i)}\big)^{1-\gamma}}{1-\gamma} + e^{-\rho}\left(1 - \overline{d}_i(t)\right) \sum_{j=1}^{n} p_{ij}(t) \frac{K_{t+1,j}}{1-\gamma} \left[(w-c)e^{r(t,i)}\right]^{1-\gamma} \right\}$$

From the first-order condition we obtain:

$$q_i(t)c^{-\gamma} = b_t e^{r(t,i)-\rho} \overline{d}_i(t) e^{-r(t,i)\gamma} (w-c)^{-\gamma} + e^{r(t,i)-\rho}\left(1 - \overline{d}_i(t)\right) e^{-\gamma r(t,i)} (w-c)^{-\gamma} \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j}$$

Rearranging yields:

$$q_i(t)c^{-\gamma} = (w-c)^{-\gamma} e^{r(t,i)-\rho} e^{-r(t,i)\gamma} \left[\overline{d}_i(t) b_t + \left(1 - \overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j}\right]$$

which implies:

$$q_i(t)^{-1/\gamma} c = (w-c) e^{(\rho - r(t,i))/\gamma} e^{r(t,i)} \left[\overline{d}_i(t) b_t + \left(1-\overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j} \right]^{-1/\gamma}$$

Rearranging further yields:

$$c = w \frac{e^{r(t,i)} \left[e^{r(t,i)} \left[\overline{d}_i(t) b_t + \left(1-\overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j} \right] \right]^{-1/\gamma}}{e^{\rho} q_i(t)^{-1/\gamma} + e^{r(t,i)} \left[e^{r(t,i)} \left[\overline{d}_i(t) b_t + \left(1-\overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j} \right] \right]^{-1/\gamma}}$$

$$= w \underbrace{\left[1 + e^{-r(t,i)} \left(\frac{e^{r(t,i)} \left[\overline{d}_i(t) b_t + \left(1-\overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j} \right]}{e^{\rho} q_i(t)} \right)^{\frac{1}{\gamma}} \right]^{-1}}_{c_{t,i}}$$

Thus we obtain:

$$V(t,w,i) = q_i(t) c_{t,i}^{1-\gamma} \frac{w^{1-\gamma}}{1-\gamma} + b_t e^{-\rho} \overline{d}_i(t) \frac{w^{1-\gamma}}{1-\gamma} \left(1-c_{t,i}\right)^{1-\gamma} e^{r(t,i)(1-\gamma)} + e^{-\rho} \left(1-\overline{d}_i(t)\right) \frac{w^{1-\gamma}}{1-\gamma} \left(1-c_{t,i}\right)^{1-\gamma} e^{r(t,i)(1-\gamma)} \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j}$$

$$= \frac{w^{1-\gamma}}{1-\gamma} \left[q_i(t) c_{t,i}^{1-\gamma} + e^{-\rho} \left(1-c_{t,i}\right)^{1-\gamma} e^{r(t,i)(1-\gamma)} \left[\overline{d}_i(t) b_t + \left(1-\overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j} \right] \right]$$

$$= \frac{w^{1-\gamma}}{1-\gamma} \frac{q_i(t) e^{r(t,i)(1-\gamma)} \left[e^{r(t,i)} \left(\overline{d}_i(t) b_t + \left(1-\overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j} \right) \right]^{1-1/\gamma} + e^{-\rho} e^{r(t,i)(1-\gamma)} \left(e^{\rho} q_i(t)\right)^{1-1/\gamma} \left[\overline{d}_i(t) b_t + \left(1-\overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j} \right]}{\left[\left(e^{\rho} q_i(t)\right)^{-1/\gamma} + e^{r(t,i)} \left[e^{r(t,i)} \left[\overline{d}_i(t) b_t + \left(1-\overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j} \right] \right]^{-\frac{1}{\gamma}} \right]^{1-\gamma}}$$

$$= \frac{w^{1-\gamma}}{1-\gamma} \frac{e^{r(t,i)(1-\gamma)} q_i(t) \left[\overline{d}_i(t) b_t + \left(1-\overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j} \right]}{\left[\left(e^{\rho} q_i(t)\right)^{-1/\gamma} + e^{r(t,i)} \left[e^{r(t,i)} \left[\overline{d}_i(t) b_t + \left(1-\overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j} \right] \right]^{-\frac{1}{\gamma}} \right]^{-\gamma}}$$

$$= \frac{w^{1-\gamma}}{1-\gamma} \underbrace{\left[q_i(t)^{\frac{1}{\gamma}} + e^{-r(t,i)} \left[e^{r(t,i)-\rho} \left(\overline{d}_i(t) b_t + \left(1-\overline{d}_i(t)\right) \sum_{j=i}^{n} p_{ij}(t) K_{t+1,j} \right) \right]^{\frac{1}{\gamma}} \right]^{\gamma}}_{K_{t,i}}$$

QED

Proof of Appendix Lemma C2:

$$W_{t+1,j}(\Psi) = \mathbb{E}\left[\exp\left\{ -\int_0^{t+1} \mu(s) ds \right\} \left(W(t+1)\right)^{\Psi} \mathbf{1}\{Y_{t+1} = j\} \right]$$

$$= \mathbb{E}\left[\exp\left\{ -\int_0^{t} \mu(s) ds \right\} \left(\left(W(t) - c(t)\right) e^{r} \right)^{\Psi} \mathbf{1}\{Y_{t+1} = j\} \exp\left\{ -\int_t^{t+1} \mu(s) ds \right\} \right]$$

$$= \sum_{k=1}^{n} \mathbb{E}\left[\mathbf{1}\{Y_t = k\} \exp\left\{-\int_0^t \mu(s)ds\right\} e^{r\Psi} W(t)^{\Psi} \left(1 - c_{t,k}\right)^{\Psi} \underbrace{\mathbb{E}\left[\mathbf{1}\{Y_{t+1} = j\} \exp\left\{-\int_t^{t+1} \mu(s)ds\right\} \middle| Y_t = k\right]}_{\left(1-\overline{d}_k(t)\right)p_{kj}(t)}\right]$$

$$= e^{r\Psi} \sum_{k=1}^{n} W_{t,k}(\Upsilon)\left(1 - c_{t,k}\right)^{\Psi} \left(1 - \overline{d}_k(t)\right) p_{kj}(t)$$

QED

Proof of Appendix Proposition C3:

Note that we have

$$\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_{Y_t}(t)c(t)^{\Psi}\right] = \sum_{j=1}^{n} \mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_{Y_t}(t)c(t)^{\Psi}\mathbf{1}\{Y_t = j\}\right]$$

$$= \sum_{j=1}^{n} \mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_j(t)c_{t,j}^{\Psi} W(t)^{\Psi}\mathbf{1}\{Y_t = j\}\right]$$

$$= \sum_{j=1}^{n} q_j(t)c_{t,j}^{\Psi}\mathbb{E}\underbrace{\left[\exp\left\{-\int_0^t \mu(s)ds\right\} W(t)^{\Psi}\mathbf{1}\{Y_t = j\}\right]}_{W_{t,j}(\Psi)}$$

The proof follows by setting $\Psi = 1 - \gamma$, 0, and $-\gamma$ in the expression for VSL.

QED

D. The fully annuitized value of life when mortality is stochastic

Even when mortality is stochastic, a complete annuities market allows the consumer to fully insure against mortality risk. We assume a full menu of actuarially fair annuities is available where consumers can choose consumption streams, $c_{Y_t}(t)$, that depend on the health state, Y_t. The consumer's maximization problem is:

$$\max_{c_{Y_t}(t)} \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) u\big(c_{Y_t}(t), q_{Y_t}(t)\big) dt \,\middle|\, Y_0\right] \tag{20}$$

$$\text{s.t.}\, \mathbb{E}\left[\int_0^T e^{-rt} S(t) c_{Y_t}(t) dt \,\middle|\, Y_0\right] = \mathbb{E}\left[W_0 + \int_0^T e^{-rt} S(t) m_{Y_t}(t) dt \,\middle|\, Y_0\right] \equiv \overline{W}(0, Y_0)$$

where the net present value of wealth and future earnings at time t in state i is $\overline{W}(t,i)$, and $S(t)$ is defined as before. Define the consumer's objective function at time u as:

$$J(u,i) = \mathbb{E}\left[\int_0^{T-u} e^{-\rho t} \exp\left\{-\int_0^t \mu(u+s) ds\right\} u\big(c_{Y_{u+t}}(u+t), q_{Y_{u+t}}(u+t)\big) dt \,\middle|\, Y_u = i\right] \tag{21}$$

We can write the objective function (21) recursively as:

$$J(u,i) = \int_0^{T-u} e^{-\rho t} \exp\left\{-\int_0^t \overline{\mu}_i(u+s) + \sum_{j\neq i} \lambda_{ij}(u+s)\, ds\right\} \left(u(c_i(u+t), q_i(u+t)) + \sum_{j \neq i} \lambda_{ij}(u+t)\, J(u+t, j) \right) dt$$

Similarly, current wealth at time u in state i, including the value of future labor income, pays for future consumption such that:

$$\begin{aligned} \overline{W}(u,i) &= \mathbb{E}\left[\int_0^{T-u} e^{-rt} \exp\left\{-\int_0^t \mu(u+s) ds\right\} c_{Y_{u+t}}(u+t) dt \,\middle|\, Y_u = i\right] \\ &= \int_0^{T-u} e^{-rt} \exp\left\{-\int_0^t \overline{\mu}_i(u+s) + \sum_{j\neq i} \lambda_{ij}(u+s)\, ds\right\} \left(c_i(u+t) + \sum_{j\neq i} \lambda_{ij}(u+t)\, \overline{W}(u+t, j) \right) dt \end{aligned}$$

This in turn implies:

$$\frac{\partial \overline{W}(t,i)}{\partial t} = \big(r + \overline{\mu}_i(t)\big) \overline{W}(t,i) - c_i(t) + \sum_{j\neq i} \lambda_{ij}(t) \big[\overline{W}(t,i) - \overline{W}(t,j)\big]$$

Define the optimal value function as:

$$V\big(t, \overline{W}_t, Y_t\big) = \max_{\{c_{Y_s}(s), s\geq t\}} \{J(t, Y_t)\}$$

where $\overline{W}_t = \big(\overline{W}(t,1), \ldots, \overline{W}(t,n)\big)$. Under conventional regularity conditions, we know that if V and its partial derivatives are continuous, then V satisfies the following Hamilton-Jacobi-Bellman (HJB) system of equations:

$$\left(\rho+\bar{\mu}_i(t)\right)V(t,\overline{W}_t,i) \tag{22}$$
$$= \max_{c_i(t)}\left\{u\big(c_i(t),q_i(t)\big) + \sum_{k=1}^{n}\frac{\partial V(t,\overline{W}_t,i)}{\partial \overline{W}(t,k)}\left[\left(r+\bar{\mu}_k(t)\right)\overline{W}(t,k)-c_k(t)+\sum_{l\neq k}\lambda_{kl}(t)\left[\overline{W}(t,k)-\overline{W}(t,l)\right]\right] + \frac{\partial V(t,\overline{W}_t,i)}{\partial t}+\sum_{j\neq i}\lambda_{ij}(t)\left[V(t,\overline{W}_t,j)-V\left(t,\overline{W}_t,i\right)\right]\right\}, 1\le i\le$$

We are interested in understanding how optimal consumption, and thus the value of life, changes over the life-cycle in this problem. Similarly to the uninsured case in the main text, we follow Parpas and Webster (2013), who demonstrate that it is possible to reformulate a stochastic optimization problem as a deterministic problem that takes $V\big(t,\overline{W}_t,j\big), j\neq i$, along with the corresponding optimal policies, as exogenous. This then allows us to apply the maximum principle and derive analytic expressions.

Appendix Lemma D1:

The optimal value function for $Y_0=i$, $V\big(t,\overline{W}_t,i\big)$, for the following deterministic optimization problem also satisfies the HJB given by (22), for each $i\in\{1,\dots,n\}$:

$$V_0\big(0,\overline{W}_0,i\big)=\max_{c_i(t)}\left[\int_0^T e^{-\rho t}\tilde{S}(i,t)\left(u(c_i(t),q_i(t))+\sum_{j\neq i}\lambda_{ij}(t)V\big(t,\overline{W}_t,j\big)\right)dt\right] \tag{23}$$
$$\text{s.t.}\frac{\partial \overline{W}(t,j)}{\partial t}=\left(r+\bar{\mu}_j(t)\right)\overline{W}(t,j)-c_j(t)+\sum_{k\neq j}\lambda_{jk}(t)\left[\overline{W}(t,j)-\overline{W}(t,k)\right], j=1,\dots,n$$

where $V\big(t,\overline{W}_t,j\big)$ and $c_j(t), j\neq i$, are taken as exogenous.

Proof of Appendix Lemma D1: see end of Appendix D

Following Bertsekas (2005), the Hamiltonian for the (deterministic) maximization problem (23) is:

$$H\big(\overline{W}_t,c_i(t),p_t\big)=e^{-\rho t}\tilde{S}(i,t)\left(u(c_i(t),q_i(t))+\sum_{j\neq i}\lambda_{ij}(t)V\big(t,\overline{W}_t,j\big)\right) + \sum_{k=1}^{n}p_t^{(k)}\left[(r+\bar{\mu}_k(t))\overline{W}(t,k)-c_k(t)+\sum_{l\neq k}\lambda_{kl}(t)\left[\overline{W}(t,k)-\overline{W}(t,l)\right]\right] \tag{24}$$

where $p_t^{(k)}$ is the costate variable corresponding to wealth $\overline{W}(t,k)$.

Appendix Lemma D2:

The consumer's first-order condition for the Hamiltonian (24) for $Y_0=i$ is

$$e^{(r-\rho)t}u_c(c_i(t),q_i(t))=\theta \tag{25}$$

where $\theta=\partial V\big(0,\overline{W}_0,i\big)/\partial\overline{W}(0,i)$ is equal to the marginal utility of wealth.

Proof of Appendix Lemma D2: see end of Appendix D

To analyze the value of life, we again let $\delta(t)$ be a perturbation on the mortality rate with $\int_0^T \delta(t)dt = 1$. As in the deterministic case, we will first derive the marginal utility of the life extension associated with this perturbation.

Appendix Proposition D3:

The marginal utility of life extension takes the same form as in the deterministic case:

$$\left.\frac{\partial V/\partial \varepsilon}{\partial V/\partial W}\right|_{\varepsilon=0} = \mathbb{E}\left[\int_0^T \left[e^{-\rho t}u\left(c_{Y_t}(t), q_{Y_t}(t)\right) + e^{-rt}\theta(m_{Y_t}(t) - c_{Y_t}(t))\right]\left(\int_0^t \delta(s)ds\right)S(t)dt \middle| Y_0\right]$$

Proof of Appendix Proposition D3: see end of Appendix D

Choosing again the Dirac delta function for $\delta(\cdot)$ and dividing the result by the marginal utility of wealth, θ, yields the value of statistical life:

$$VSL = \mathbb{E}\left[\int_0^T e^{-rt}S(t)v_{Y_t}(t)dt \middle| Y_0\right] \tag{26}$$

where the value of a statistical life-year is:

$$v_{Y_t}(t) = \frac{u\left(c_{Y_t}(t), q_{Y_t}(t)\right)}{u_c\left(c_{Y_t}(t), q_{Y_t}(t)\right)} + m_{Y_t}(t) - c_{Y_t}(t)$$

Comparing (26) to (3) reveals that stochastic mortality does not alter the basic expression for VSL. Consumers continue to discount future life-years by the rate of interest and by survival. One notable difference is that stochastic mortality generates variance in the value of life, which can now increase or decrease following the transition to a new health state.

We can obtain the life-cycle profile of consumption by differentiating the first-order condition (25) with respect to t. Doing so confirms that, as in the deterministic case, annuitization insulates consumption from mortality risk:[25]

$$\frac{\dot{c}_{Y_t}}{c_{Y_t}} = \frac{\dot{c}}{c} = \sigma(r-\rho) + \sigma\eta\frac{\dot{q}}{q}$$

Our results demonstrate that stochastic mortality, by itself, does not alter the basic insights regarding VSL offered by the prior literature as long as one maintains the assumption of full annuitization.

A novel feature of the stochastic model is that it permits an investigation into the value of prevention, i.e., the value of a reduction in the probability of transitioning to a different health state. This is not possible in a deterministic environment, where there is implicitly only one health state.

To analyze the value of prevention, let $\delta_{ij}(t)$ be a perturbation on $\lambda_{ij}(t)$, where $\sum_{j\neq i}\int_0^T \delta_{ij}(t)dt = 1$. As in the life-extension case, it is helpful to choose the Dirac delta function for $\delta(\cdot)$, so that the probability is

[25] We assume—like all prior studies—that full indemnity healthcare insurance is available, which is equivalent to assuming that $q(t)$ is independent of the health state. Without this assumption, sudden decreases in q could cause the value of life to jump (Lakdawalla, Malani, and Reif 2017).

perturbed at $t = 0$ and remains unaffected otherwise. It is also helpful to consider a reduction in the transition probability for only one alternative state, j_0, so that $\delta_{ij}(t) = 0 \ \forall j \neq j_0$.

Appendix Proposition D4:

Define the value of statistical illness, $VSI(i, j_0)$, to be the value of marginal reduction in the probability of transitioning to state j_0 when in state i. This value is equal to:

$$\begin{aligned} VSI(i, j_0) = \ & \mathbb{E}\left[\int_0^T e^{-rt}\left[\frac{u\left(c_{Y_t}(t), q_{Y_t}(t)\right)}{u_c\left(c_{Y_t}(t), q_{Y_t}(t)\right)} + m(t) - c(t)\right] S(t)dt \,\middle|\, Y_0 = i\right] \\ & - \mathbb{E}\left[\int_0^T e^{-rt}\left[\frac{u\left(c_{Y_t}(t), q_{Y_t}(t)\right)}{u_c\left(c_{Y_t}(t), q_{Y_t}(t)\right)} + m(t) - c(t)\right] S(t)dt \,\middle|\, Y_0 = j_0\right] \\ = \ & VSL(i) - VSL\big(j_0 | \overline{W}(0) = W^*\big) \end{aligned} \tag{27}$$

where W^* is the value of the annuity that was initially purchased in state *i* that promised the state-contingent consumption stream $c^*_{Y_t}(t)$:

$$W^* = \mathbb{E}\left[\int_0^T e^{-rt} S(t) c^*_{Y_t}(t) dt \,\middle|\, Y_0 = j_0\right]$$

Proof of Appendix Proposition D4: see end of Appendix D

The notation in equation (27) indicates that VSL in state j_0 is evaluated under the assumption that the consumer's annuity was purchased when she was in state i. If life expectancy in state j_0 is lower than in state i, the value of the annuity to the consumer falls.

Proofs for Appendix C

Proof of Appendix Lemma D1:

Let $V(t, \overline{W}_t, j)$ and $c_j(t), j \neq i$, be taken as given (exogenous). Consider the deterministic optimization problem:

$$V(0, \overline{W}_0, i) = \max_{c_i(t)} \left\{ \int_0^T e^{-\rho t} \tilde{S}(i,t) \left(u(c_i(t), q_i(t)) + \sum_{j \neq i} \lambda_{ij}(t) V(t, \overline{W}_t, j) \right) dt \right\}$$

$$s.t.\ \frac{\partial \overline{W}(t,j)}{\partial t} = \left(r + \overline{\mu}_j(t)\right) \overline{W}(t,j) - c_j(t) + \sum_{k \neq j} \lambda_{jk}(t) \left[\overline{W}(t,j) - \overline{W}(t,k)\right], j = 1, \ldots, n$$

Denote the optimal value-to-go as

$$\tilde{V}(u, \overline{W}_u, i) = \max_{c_i(t)} \left\{ \int_u^T e^{-\rho t} \tilde{S}(i,t) \left(u(c_i(t), q_i(t)) + \sum_{j \neq i} \lambda_{ij}(t) V(t, \overline{W}_t, j) \right) dt \right\}$$

Setting $\tilde{V}(u, \overline{W}_u, i) = e^{-\rho t} \tilde{S}(i,t) V(t, \overline{W}_t, i)$ then demonstrates that $V(\cdot)$ satisfies the HJB (22) for i.

QED

Proof of Appendix Lemma D2:

The costate equations for the Hamiltonian (24) are:

$$\dot{p}_t^{(i)} = -p_t^{(i)} \left(r + \overline{\mu}_i(t) + \sum_{j \neq i} \lambda_{ij}(t) \right) + \sum_{l \neq i} \lambda_{li}(t) p_t^{(l)} \text{ and}$$

$$\dot{p}_t^{(k)} = e^{-\rho t} \tilde{S}(i,t) \lambda_{ik}(t) \frac{\partial V(t, \overline{W}_t, k)}{\partial \overline{W}(t,k)} - p_t^{(k)} \left(r + \overline{\mu}_k + \sum_{l \neq k} \lambda_{kl}(t) \right) + \sum_{l \neq k} \lambda_{lk}(t) p_t^{(l)}$$

for $k \neq i$. Suppose that $p_t^{(k)} = 0, k \neq i$. (We will verify this at the end of the proof.) This implies:

$$p_t^{(i)} = \theta e^{-rt} \tilde{S}(i,t)$$

where θ is a constant. Note also that the first-order condition of the Hamiltonian with respect to $c_i(t)$ is

$$e^{-\rho t} \tilde{S}(i,t) u_c(c_i(t), q_i(t)) = p_t^{(i)}$$

Setting these last two equations equal to each other then yields the desired result.

To verify that $p_t^{(k)} = 0$, $k \neq i$, note that the previous result implies via the HJB that $\partial V(t, \overline{W}_t, i) / \partial \overline{W}(t,i) = \theta\, e^{-(r-\rho)t}$, so that the costate equation for $k \neq i$ is

$$
\begin{aligned}
\dot{p}_t^{(k)} &= -\overbrace{\theta e^{-rt}\tilde{S}(i,t)}^{p_t^{(i)}}\lambda_{ik}(t) + p_t^{(i)}\lambda_{ik}(t) - p_t^{(k)}\left(r + \overline{\mu}_k + \sum_{l\neq k}\lambda_{kl}(t)\right) + \sum_{l\neq\{k,i\}}\lambda_{lk}(t)p_t^{(l)} \\
&= 0
\end{aligned}
$$

QED

Proof of Appendix Proposition D3:

The marginal utility of life extension is defined as

$$
\left.\frac{\partial V/\partial\varepsilon}{\partial V/\partial W}\right|_{\varepsilon=0} = \frac{\partial}{\partial\varepsilon}\mathbb{E}\left[\left.\left.\int_0^T e^{-\rho t}\exp\left\{-\int_0^t \mu(s) - \varepsilon\delta(s)ds\right\}\left(u\left(c_{Y_t}^{\varepsilon}(t), q_{Y_t}(t)\right)\right)dt\right| Y_0\right]\right|_{\varepsilon=0}
$$

where $c^{\varepsilon}(t)$ represents the equilibrium variation in $c(t)$ caused by this perturbation. Then

$$
\begin{aligned}
\left.\frac{\partial V/\partial\varepsilon}{\partial V/\partial W}\right|_{\varepsilon=0} &= \mathbb{E}\left[\left.\int_0^T e^{-\rho t}\left(\int_0^t \delta(s)ds\right)\exp\left\{-\int_0^t \mu(s)ds\right\}u\big(c_{Y_t}(t), q_{Y_t}(t)\big)dt\right| Y_0\right] \\
&\quad + \mathbb{E}\left[\left.\int_0^T e^{-\rho t}\exp\left\{-\int_0^t \mu(s)ds\right\}u_c\left(c_{Y_t}^{\varepsilon}(t), q_{Y_t}(t)\right)\left.\frac{\partial c_{Y_t}^{\varepsilon}(t)}{\partial\varepsilon}\right|_{\varepsilon=0} dt\right| Y_0\right] \\
&= \mathbb{E}\left[\left.\int_0^T e^{-\rho t}\left(\int_0^t \delta(s)ds\right)\exp\left\{-\int_0^t \mu(s)ds\right\}u\big(c_{Y_t}(t), q_{Y_t}(t)\big)dt\right| Y_0\right] \\
&\quad + \theta\mathbb{E}\left[\left.\int_0^T e^{-rt}\exp\left\{-\int_0^t \mu(s)ds\right\}\left.\frac{\partial c_{Y_t}^{\varepsilon}(t)}{\partial\varepsilon}\right|_{\varepsilon=0} dt\right| Y_0\right]
\end{aligned}
$$

Finally, the budget constraint implies

$$
\begin{aligned}
0 &= \left.\frac{\partial W_0}{\partial\varepsilon}\right|_{\varepsilon=0} \\
&= \frac{\partial}{\partial\varepsilon}\mathbb{E}\left[\left.\left.\int_0^T e^{-rt}\exp\left\{-\int_0^t \mu(s) - \varepsilon\delta(s)ds\right\}\left(c_{Y_t}^{\varepsilon}(t) - m_{Y_t}(t)\right)dt\right| Y_0\right]\right|_{\varepsilon=0} \\
&= \mathbb{E}\left[\left.\int_0^T e^{-rt}\left(\int_0^t \delta(s)ds\right)\exp\left\{-\int_0^t \mu(s)ds\right\}\left(c_{Y_t}(t) - m_{Y_t}(t)\right)dt\right| Y_0\right] \\
&\quad + \mathbb{E}\left[\left.\int_0^T e^{-rt}\exp\left\{-\int_0^t \mu(s)ds\right\}\left.\frac{\partial c_{Y_t}^{\varepsilon}(t)}{\partial\varepsilon}\right|_{\varepsilon=0} dt\right| Y_0\right]
\end{aligned}
$$

Plugging this last result into the expression for $\left.\frac{\partial V/\partial\varepsilon}{\partial V/\partial W}\right|_{\varepsilon=0}$ then yields the desired result.

QED

Proof of Appendix Proposition D4:

Working from equation (23) in the text, the marginal utility of prevention is given by

$$\left.\frac{\partial V/\partial \varepsilon}{\partial V/\partial W}\right|_{\varepsilon=0} = \frac{\partial}{\partial \varepsilon}\int_0^T e^{-\rho t}\exp\left\{-\int_0^t \overline{\mu}_i(s) + \sum_{j\neq i}\left[\lambda_{ij}(s) - \varepsilon\delta_{ij}(s)\right] ds\right\}\left(u\left(c_i^\varepsilon(t), q_i(t)\right)\right.$$
$$\left.\left. + \sum_{j\neq i}\left[\lambda_{ij}(t) - \varepsilon\delta_{ij}(t)\right] V\left(t, \overline{W}_t^\varepsilon, j\right)\right) dt\right|_{\varepsilon=0}$$

where $c_i^\varepsilon(t)$ and $\overline{W}_t^\varepsilon$ represent the equilibrium variations in $c_i(t)$ and $\overline{W}_t$ caused by this perturbation. This yields

$$\left.\frac{\partial V/\partial \varepsilon}{\partial V/\partial W}\right|_{\varepsilon=0} = \int_0^T e^{-\rho t}\left(\int_0^t \sum_{j\neq i}\delta_{ij}(s)\, ds\right)\tilde{S}(i,t)\left(u\left(c_i(t), q_i(t)\right) + \sum_{j\neq i}\lambda_{ij}(t)\, V\left(t, \overline{W}_t, j\right)\right)$$
$$- e^{-\rho t}\tilde{S}(i,t)\sum_{j\neq i}\delta_{ij}(t)\, V\left(t, \overline{W}_t, j\right)$$
$$+ e^{-\rho t}\tilde{S}(i,t)\left(\underbrace{u_c\left(c_i(t), q_i(t)\right)}_{\theta e^{-(r-\rho)t}}\left.\frac{\partial c_i^\varepsilon(t)}{\partial \varepsilon}\right|_{\varepsilon=0} + \sum_{j\neq i}\lambda_{ij}(t)\underbrace{V_W\left(t, \overline{W}_t, j\right)}_{\theta e^{-(r-\rho)t}}\left.\frac{\partial \overline{W}^\varepsilon(t,j)}{\partial \varepsilon}\right|_{\varepsilon=0}\right) dt$$

Next, note that the budget constraint implies

$$0 = \left.\frac{\partial W_0^\varepsilon}{\partial \varepsilon}\right|_{\varepsilon=0} = \frac{\partial}{\partial \varepsilon}\int_0^T e^{-rt}\exp\left\{-\int_0^t \overline{\mu}_i(s) + \sum_{j\neq i}\left[\lambda_{ij}(s) - \varepsilon\delta_{ij}(s)\right] ds\right\}\left(c_i^\varepsilon(t) - m_i(t)\right.$$
$$\left.\left. + \sum_{j\neq i}\left[\lambda_{ij}(t) - \varepsilon\delta_{ij}(t)\right]\overline{W}^\varepsilon(t,j)\right) dt\right|_{\varepsilon=0}$$
$$= \int_0^T e^{-rt}\left(\int_0^t \sum_{j\neq i}\delta_{ij}(s)\, ds\right)\tilde{S}(i,t)\left(c_i(t) - m_i(t) + \sum_{j\neq i}\lambda_{ij}(t)\overline{W}(t,j)\right) dt$$
$$- e^{-rt}\tilde{S}(i,t)\sum_{j\neq i}\delta_{ij}(t)\,\overline{W}(t,j)$$
$$+ e^{-rt}\tilde{S}(i,t)\left(\left.\frac{\partial c_i^\varepsilon(t)}{\partial \varepsilon}\right|_{\varepsilon=0} + \sum_{j\neq i}\lambda_{ij}(t)\left.\frac{\partial \overline{W}^\varepsilon(t,j)}{\partial \varepsilon}\right|_{\varepsilon=0}\right) dt$$

Substituting in then yields the final result for the marginal utility of the reduction in this transition intensity:

$$
\begin{aligned}
\left.\frac{\partial V/\partial \varepsilon}{\partial V/\partial W}\right|_{\varepsilon=0} &= \int_0^T e^{-\rho t}\left(\int_0^t \sum_{j\neq i}\delta_{ij}(s)\,ds\right)\tilde{S}(i,t)\left(u\big(c_i(t),q_i(t)\big)+\sum_{j\neq i}\lambda_{ij}(t)\,V\big(t,\overline{W}_t,j\big)\right)\\
&\quad - e^{-\rho t}\tilde{S}(i,t)\sum_{j\neq i}\delta_{ij}(t)\,V\big(t,\overline{W}_t,j\big)\\
&\quad - \theta e^{-rt}\left(\int_0^t \sum_{j\neq i}\delta_{ij}(s)\,ds\right)\tilde{S}(i,t)\left(c_i(t)-m_i(t)+\sum_{j\neq i}\lambda_{ij}(t)\overline{W}(t,j)\right)\\
&\quad + \theta e^{-rt}\tilde{S}(i,t)\sum_{j\neq i}\delta_{ij}(t)\,\overline{W}(t,j)dt\\
&= \int_0^T \left(e^{-\rho t}u(c_i(t),q_i(t)) + \sum_{j\neq i}\lambda_{ij}(t)\,V\big(t,\overline{W}_t,j\big)\right.\\
&\quad \left. + \theta e^{-rt}\left(m_i(t)-c_i(t)-\sum_{j\neq i}\lambda_{ij}(t)\,\overline{W}(t,j)\right)\right)\left(\int_0^t \sum_{j\neq i}\delta_{ij}(s)\,ds\right)\tilde{S}(i,t)\\
&\quad - \left(e^{-\rho t}\sum_{j\neq i}\delta_{ij}(t)\,V\big(t,\overline{W}_t,j\big) - \theta e^{-rt}\sum_{j\neq i}\delta_{ij}(t)\,\overline{W}(t,j)\right)\tilde{S}(i,t)dt
\end{aligned}
$$

The first term inside the integral of the above expression represents the gain in marginal utility from a reduction in the probability of exiting state $Y_t = i$, and is analogous to the expression for $\left.\frac{\partial V/\partial \varepsilon}{\partial V/\partial W}\right|_{\varepsilon=0}$ for life-extension. The second term represents the loss in marginal utility from the reduction in probability of transitioning to other possible states. If these other states correspond to lower health (utility) than state i, then the net effect on marginal utility is positive.

Next, we choose the Dirac delta function for $\delta(\cdot)$, so that the probability is perturbed at $t = 0$ and remains unaffected otherwise. We also consider a reduction in the transition probability for only one alternative state, j_0, so that $\delta_{ij}(t) = 0\ \forall j \neq j_0$. This simplifies the above expression to

$$
\begin{aligned}
\left.\frac{\partial V/\partial \varepsilon}{\partial V/\partial W}\right|_{\varepsilon=0} &= \int_0^T \left(e^{-\rho t}u\big(c_i(t),q_i(t)\big) + \sum_{j\neq i}\lambda_{ij}(t)\,V\big(t,\overline{W}_t,j\big) + \theta e^{-rt}\left(m_i(t)-c_i(t)-\sum_{j\neq i}\lambda_{ij}(t)\,\overline{W}(t,j)\right)\right)\tilde{S}(i,t)\\
&\quad - \left(e^{-\rho t}V\big(t,\overline{W}_t,j_0\big) - \theta e^{-rt}\overline{W}(t,j_0)\right)\tilde{S}(i,t)dt\\
&= \int_0^T \left(e^{-\rho t}\tilde{S}(i,t)u\big(c_i(t),q_i(t)\big) + \sum_{j\neq i}\lambda_{ij}(t)\,V\big(t,\overline{W}_t,j\big)\right)dt - V\big(0,\overline{W}_t,j_0\big)\\
&\quad + \theta\left[\int_0^T \left(e^{-rt}\left(m_i(t)-c_i(t)-\sum_{j\neq i}\lambda_{ij}(t)\,\overline{W}(t,j)\right)\right)\tilde{S}(i,t)dt + \overline{W}(0,j_0)\right]
\end{aligned}
$$

$$
\begin{aligned}
= \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) u\left(c_{Y_t}(t), q_{Y_t}(t)\right) dt \middle| Y_0 = i\right] - \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) u\left(c_{Y_t}(t), q_{Y_t}(t)\right) dt \middle| Y_0 = j_0, W_0^* = W^{new}\right] \\
+ \theta\left(\mathbb{E}\left[\int_0^T e^{-rt} S(t)\left(m_{Y_t}(t) - c_{Y_t}(t)\right) dt \middle| Y_0 = i\right]\right. \\
\left. - \mathbb{E}\left[\int_0^T e^{-rt} S(t)\left(m_{Y_t}(t) - c_{Y_t}(t)\right) dt \middle| Y_0 = j_0, W_0^* = W^{new}\right]\right)
\end{aligned}
$$

where W^{new} represents the change in value of the annuity menu purchased in state i when immediately jumping to state j_0. Dividing the above expression by the marginal utility of wealth, given by (25), then yields (27), the value of statistical illness (VSI).

QED