APPENDIX (FOR ONLINE PUBLICATION ONLY)

The value of life depends greatly on the elasticity of intertemporal substitution, which under CRRA is equal to the inverse of γ, the coefficient of relative risk aversion. The specification in the main text, which sets $\gamma = 2$, calculated that Social Security raised the aggregate social value of post-1940 reductions by \$11.5 trillion (10.5 percent). Appendix Table A1 and Appendix Table A2, which both replicate Table 4 from the main text, show that varying γ over the range $[1.5, 2.5]$ yields analogous increases that range from 8.3 percent to 14.2 percent.

The first two columns of Appendix Table A3 show that when a strong bequest motive is present, the increase in the aggregate social value of life attributable to Social Security is equal to \$5.5 trillion (5.4 percent). Finally, the third column of Appendix Table A3 shows that fully annuitizing all wealth and future earnings at age 20 increases the aggregate value of life by \$17.7 trillion (16 percent), relative to a setting with no annuity markets.

Appendix Figure A1 reports average out-of-pocket medical spending, by age, for a healthy individual in health state 1 and for a very sick individual in health state 20. These spending data include all inpatient, outpatient, prescription drug, and long-term care payments made by the individual, as estimated by the Future Elderly Model. The large increase in spending that occurs after age 80 is due primarily to the large costs of long-term care.

Appendix A provides proofs for lemmas and propositions stated in the main text. Appendix B provides descriptions of the data employed by the numerical models presented in Section IV, and Appendix C provides supporting calculations 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.

Appendix Tables and Figures

Appendix Table A1. Aggregate social value of historical and prospective reductions in mortality when the degree of relative risk aversion is set equal to $\gamma = 2.5$ (billions of dollars)

	(1)	**(2)**	**(3)**
	No annuity	**Social Security**	**Social Security + 50%**
Historical reduction:			
1940-2010	\$222,046	\$253,546	\$269,951
1970-2010	\$109,580	\$126,291	\$135,146
10% reduction, all ages:			
All causes	\$23,879	\$27,569	\$29,566
Cancer	\$6,943	\$8,081	\$8,697
Diabetes	\$762	\$885	\$951
Heart disease	\$5,068	\$5,910	\$6,374
Homicide	\$189	\$187	\$184
Infectious diseases	\$349	\$408	\$441

Notes: These aggregate values were calculated using the 2015 US population by age. Panel A reports the current value of historical reductions in all-cause mortality. Panel B reports the value of a 10 percent prospective reduction in mortality. 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. The degree of relative risk aversion, γ, is equal to the inverse of the elasticity of intertemporal substitution. In the main text, we assume that $\gamma = 2$.

Appendix Table A2. Aggregate social value of historical and prospective reductions in mortality when the degree of relative risk aversion is set equal to $\gamma = 1.5$ (billions of dollars)

	(1)	**(2)**	**(3)**
	No annuity	**Social Security**	**Social Security + 50%**
Historical reduction:			
1940-2010	\$27,121	\$29,381	\$30,465
1970-2010	\$5,750	\$6,277	\$6,555
10% reduction, all ages:			
All causes	\$1,661	\$1,822	\$1,903
Cancer	\$183	\$200	\$209
Diabetes	\$1,185	\$1,310	\$1,379
Heart disease	\$63	\$61	\$59
Homicide	\$80	\$89	\$94
Infectious diseases	\$0	\$0	\$0

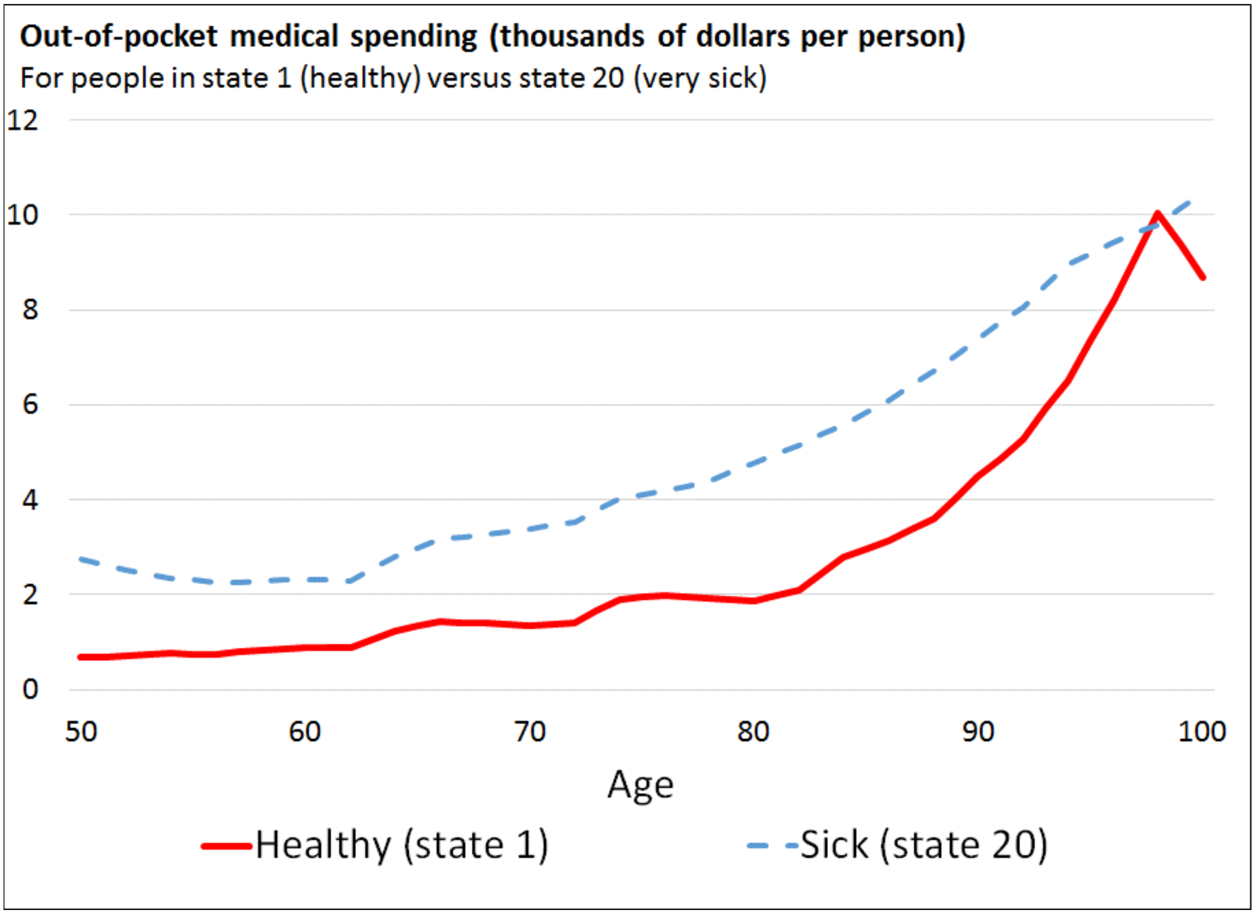
Notes: These aggregate values were calculated using the 2015 US population by age. Panel A reports the current value of historical reductions in all-cause mortality. Panel B reports the value of a 10 percent prospective reduction in mortality. Column (1) presents estimates under the assumption that individuals 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. The degree of relative risk aversion, γ, is equal to the inverse of the elasticity of intertemporal substitution. In the main text, we assume that $\gamma = 2$.

Appendix Table A3. Aggregate social value of historical and prospective reductions in mortality when a bequest motive is present or when consumer is fully annuitized (billions of dollars)

	(1)	(2)	(3)	(4)
	Bequest motive		**No bequest motive**	
	No annuity	**Social Security**	**No annuity**	**Full annuitization**
Historical reduction:				
1940-2010	$102,744	$108,261	$109,356	$127,030
1970-2010	$50,110	$53,081	$53,492	$60,571
10% reduction, all ages:				
All causes	$11,042	$11,758	$11,550	$13,403
Cancer	$3,150	$3,362	$3,348	$3,708
Diabetes	$348	$371	$368	$412
Heart disease	$2,338	$2,512	$2,425	$2,755
Homicide	$99	$95	$105	$173
Infectious diseases	$163	$176	$166	$192

Notes: The bequest motive specification is described at the end of Appendix C1. These aggregate values were calculated using the 2015 US population by age. Panel A reports the current value of historical reductions in all-cause mortality. Panel B reports the value of a 10 percent prospective reduction in mortality. Column (1) presents estimates under the assumption that individuals have no annuities. 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.

Appendix Figure A1. Annual out-of-pocket medical spending for a healthy person versus a very sick patient



Notes: These medical spending estimates include out-of-pocket spending on both health care and nursing homes. State 1 corresponds to a healthy individual with no impaired activities of daily living (ADL) and no chronic conditions. State 20 corresponds to an individual with three or more ADL's and four or more chronic conditions. Additional characteristics for these health states are provided in Table 1. These estimates are provided by the Future Elderly Model, which is described in greater detail in Appendix B2.

A. Mathematical proofs of results from main text

Proof of Lemma 1:

Recall that the transition intensities $\lambda_{ij}(t) = 0\ \forall j < i$. The optimization problem in the absorbing state n is therefore a standard deterministic problem. We can contemplate a successive solution strategy by starting in state n and then moving sequentially to state $n-1$, $n-2$, etc. Thus, we can consider the deterministic optimization problem for an arbitrary state i by taking $V(t, w, j), j > i$, as given (exogenous):

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

subject to:

$$\frac{\partial W_i(t)}{\partial t} = rW_i(t) + m_i(t) - c_i(t), W_i(0) = W_0$$

Optimal consumption and wealth in state i are denoted by $c_i(t)$ and $W_i(t)$, respectively. Denote the optimal value-to-go function as:

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

Setting $\tilde{V}(t, W_i(t), i) = e^{-\rho t} \tilde{S}(i,t) V(t, W_i(t), i)$ then demonstrates that $V(\cdot)$ satisfies the HJB (11) for i. See Theorem 1 and the proof of Theorem 2 in Parpas and Webster (2013) for additional details and intuition behind this result.

QED

Proof of Lemma 2:

From (12), the marginal utility of life-extension is:

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

where $c_i^\varepsilon(t)$ and $W_i^\varepsilon(t)$ represent the equilibrium variations in $c_i(t)$ and $W_i(t)$ 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:

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

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

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

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

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

where, as in the deterministic case, the last equality follows from application of the budget constraint.

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 (13) simplifies to:[29]

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

This then implies:

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

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

[29] When no transitions are possible, the solution reduces to the first-order condition presented in Section II.B.

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

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

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:

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

which implies:

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

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

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

Since $\frac{\partial V(s, W_i(s), j)}{\partial W_i(s)} = u_c\big(c(s, W_i(s), j), q_j(s)\big)$ from the first-order condition in the HJB for state j, we obtain:

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

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:

$$\left.\frac{\partial V}{\partial \varepsilon}\right|_{\varepsilon=0} = \int_0^T \left[e^{-\rho t}\tilde{S}(i,t)\left(u\big(c_i(t), q_i(t)\big) + \sum_{j>i} \lambda_{ij}(t)\, V(t, W_i(t), j) \right)\right] dt$$
$$= \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) u\big(c(t), q_{Y_t}(t)\big) dt \,\middle|\, Y_0 = i, W(0) = W_0 \right]$$

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

$$VSL(i) = \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) \frac{u\Big(c(t), q_{Y_t}(t)\Big)}{u\Big(c(0), q_{Y_0}(0)\Big)} dt \,\middle|\, Y_0 = i, W(0) = W_0 \right]$$

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

$$VSL(i) = \mathbb{E}\left[\int_0^T e^{-\rho t} \frac{S(t) u\big(c(t), q_{Y_t}(t)\big)}{\mathbb{E}\left[e^{(r-\rho)t} \exp\left\{ - \int_0^t \mu(s) ds \right\} u_c\big(c(t), q_{Y_t}(t)\big) \,\middle|\, Y_0 = i, W(0) = W_0 \right]} dt \,\middle|\, Y_0 = i, W(0) = W_0 \right]$$
$$= \mathbb{E}\left[\int_0^T e^{-rt} \frac{S(t) u\big(c(t), q_{Y_t}(t)\big)}{\mathbb{E}\left[\exp\left\{ - \int_0^t \mu(s) ds \right\} u_c\big(c(t), q_{Y_t}(t)\big) \,\middle|\, Y_0 = i, W(0) = W_0 \right]} dt \,\middle|\, Y_0 = i, W(0) = W_0 \right]$$

Exchanging expectation and integration then yields:

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

where 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:

$$v(i,t) = \frac{\mathbb{E}\left[S(t) u\Big(c(t), q_{Y_t}(t)\Big) \,\middle|\, Y_0 = i, W(0) = W_0 \right]}{\mathbb{E}\left[S(t) u_c\Big(c(t), q_{Y_t}(t)\Big) \,\middle|\, Y_0 = i, W(0) = W_0 \right]}$$

QED

Proof of Proposition 5:

Without loss of generality, we will prove the proposition for the case where the consumer transitions from state 1 to state 2 at time $t = 0$. Because we hold quality of life constant, we omit $q_i(t)$ in the notation below in order to keep the presentation concise.

We want to prove that $c_2(0) \geq c_1(0)$. Assume by way of contradiction that $c_2(0) < c_1(0)$. We will show that this implies $c_2(t) < c_1(t)$ for all $t > 0$, which is a contradiction since the feasible consumption plan $c_1(\cdot)$ dominates $c_2(\cdot)$.

We proceed by inductively constructing a sequence $0 < t_1 < t_2$ ... where for each element in the sequence:

$$c_2(t_i) < c_1(t_i)$$
$$W_1(t_i) \leq W_2(t_i)$$

$$p_{t_i}^{(1)} < \exp\left\{-\int_0^{t_i} \lambda_{12}(s)ds\right\} p_{t_i}^{(2)}$$

To construct the sequence, for the base case $i = 1$, we first note that from the first-order condition (14), we obtain:

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

The costate equation (13) then implies:

$$\begin{aligned}
\dot{p}_0^{(1)} &= -p_0^{(1)} r - \lambda_{12}(0) u_c\big(c_2(0)\big) \\
&= -p_0^{(1)} \left[r + \lambda_{12}(0) \underbrace{\frac{u_c\big(c_2(0)\big)}{u_c\big(c_1(0)\big)}}_{>1} \right] \\
&< -p_0^{(1)} [r + \lambda_{12}(0)] = \left.\frac{\partial g(t)}{\partial t}\right|_{t=0}
\end{aligned}$$

where $g(t) = p_0^{(1)} \exp\left\{-\int_0^t r + \lambda_{12}(s)ds\right\}$. Hence, there exists a $t_1 > t_0 = 0$ such that:

$$p_t^{(1)} \le g(t) < p_0^{(2)} \exp\left\{-\int_0^t \big(r + \lambda_{12}(s)\big)ds\right\} = p_t^{(2)} \exp\left\{-\int_0^t \lambda_{12}(s)ds\right\}, 0 \le t \le t_1$$

which together with the first-order condition (14) implies:

$$e^{-\rho t} \exp\left\{-\int_0^t \left(\overline{\mu}_1(s) + \lambda_{12}(s)\right) ds\right\} u_c\big(c_1(t)\big) < e^{-\rho t} \exp\left\{-\int_0^t \left(\overline{\mu}_2(s) + \lambda_{12}(s)\right) ds\right\} u_c\big(c_2(t)\big), 0 \le t \le t_1$$

so that $c_1(t) > c_2(t), 0 \le t \le t_1$. Since $m_1(s) \le m_2(s)\ \forall s$, this in turn implies $W_1(t_1) \le W_2(t_1)$.

For the induction step, suppose that the following properties also hold for $i \ge 1$:

$$\begin{aligned}
c_2(t_i) &< c_1(t_i) \\
W_1(t_i) &\le W_2(t_i) \\
p_{t_i}^{(1)} &< \exp\left\{-\int_0^{t_i} \lambda_{12}(s)ds\right\} p_{t_i}^{(2)}
\end{aligned}$$

The induction hypothesis implies:

$$c(t_i, W_1(t_i), 2) \le c(t_i, W_2(t_i), 2) = c_2(t_i) < c_1(t_i)$$

so that:

$$\begin{aligned}
\dot{p}_{t_i}^{(1)} &= -p_{t_i}^{(1)} r - e^{-\rho t_i} \tilde{S}(1, t_i) \lambda_{12}(t_i) u\big(c(t_i, W_1(t_i), 2)\big) \\
&= -p_{t_i}^{(1)} \left[r + \lambda_{12}(t_i) \underbrace{\frac{u_c\big(c(t_i, W_1(t_i), 2)\big)}{u_c\big(c_1(t_i)\big)}}_{>1} \right] \\
&< -p_{t_i}^{(1)} [r + \lambda_{12}(t_i)] = \left.\frac{\partial \tilde{g}(t_i)}{\partial t}\right|_{t_i=0}
\end{aligned}$$

with $\tilde{g}(t_i) = p_{t_i}^{(1)} \exp\left\{-\int_{t_i}^{t}\big(r + \lambda_{12}(s)\big)ds\right\}$. Hence, there exists a $t_{i+1} > t_i$ such that:

$$p_t^{(1)} \leq \tilde{g}(t)$$
$$< \exp\left\{-\int_0^{t_i}\lambda_{12}(s)ds\right\} p_{t_i}^{(2)} \exp\left\{-\int_{t_i}^{t}\big(r + \lambda_{12}(s)\big)ds\right\} = p_t^{(2)} \exp\left\{-\int_0^{t}\lambda_{12}(s)ds\right\}, t_i \leq t \leq t_{i+1}$$

In particular, again with the first-order condition (14) for all $t_i \leq t \leq t_{i+1}$:

$$\exp\left\{-\int_0^t \left(\overline{\mu}_1(s) + \lambda_{12}(s)\right) ds\right\} u_c\big(c_1(t)\big) < \exp\left\{-\int_0^t \left(\overline{\mu}_2(s) + \lambda_{12}(s)\right) ds\right\} u_c\big(c_2(t)\big)$$

which in turn implies $u_c\big(c_1(t)\big) < u_c\big(c_2(t)\big)$ and $c_2(t) < c_1(t)$. Once again, together with the assumption $m_1(s) \leq m_2(s)$, this implies $W_1(t_{i+1}) \leq W_2(t_{i+1})$.

Thus, we have proven the existence of the sequence. We then obtain $c_2(t) < c_1(t)\ \forall t$ by noting that $\{t_i\}_{i\geq 0}$ strictly increases due to the uniformly boundedness condition on $\lambda_{12}(t)$, which is the desired contradiction.

We note that this proof implies that the consumption paths $c_1(t)$ and $c_2(t)$ cross (at most) once. As soon as $c_1(t)$ exceeds $c_2(t)$ for some time t_0, $c_1(t)$ will exceed $c_2(t)$ for $t > t_0$. However, we have that $c_2(t)$ exceeds $c_1(t)$ prior to t_0. In particular, consumption jumps up at the transition point. See Figure 2 for an illustration.

QED

Proof of Proposition 6:

Without loss of generality, consider the case $t = 0$, as depicted in Figure 2. From **Proposition 5** it is clear that $c_1(t)$ and $c_2(t)$ are decreasing, $c_2(0) \geq c_1(0)$, $c_2(t) \geq c_1(t)$ 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:

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

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 on preferences, we have that:[30]

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

[30] Strictly speaking, this proof requires only that $Y'(x) > 0$ and $Y''(x) > 0$. The stated assumptions on preferences are therefore sufficient, but not necessary.

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

Employing Taylor's theorem then implies that for some $\xi(t)$ that lies in-between $c_1(t)$ and $c_2(t)$:

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

where the final step follows from the budget constraint.

QED

Proof of Proposition 7:

From (12), the marginal utility of preventing an illness or death is:

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

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

QED

B. Data

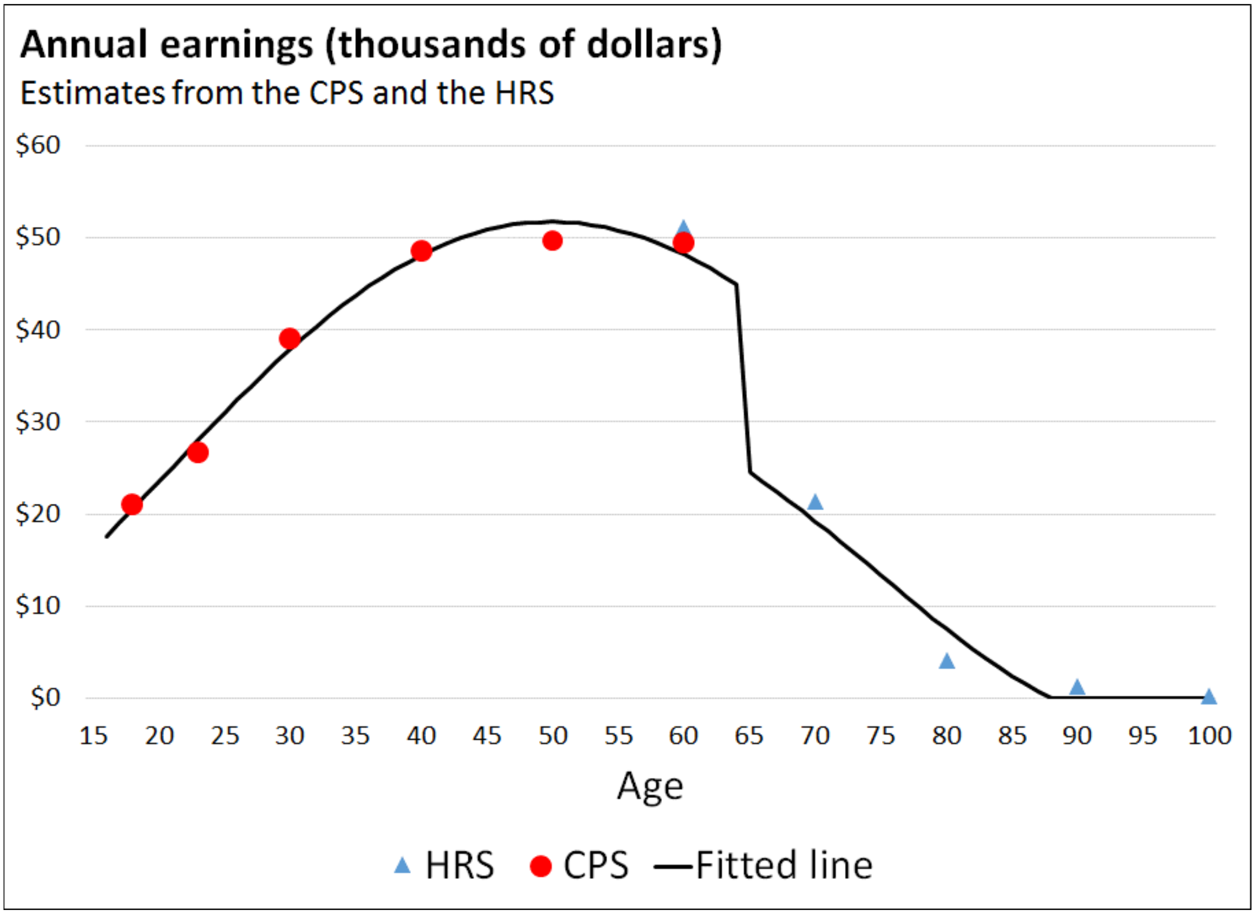
B1. Earnings

We obtain earnings data for employed individuals under the age of 65 from the 2016 Current Population Survey (CPS).[31] 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 B1. 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.

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

Appendix Figure B1. Annual earnings estimates from CPS and HRS



Notes: This figure plots annual earnings by midpoint of age group as estimated by the 2016 Current Population Survey (CPS) for respondents under age 65, and as estimated by the 2014 Health and Retirement Survey (HRS) for respondents over age 55. 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 in that regression, annual earnings, corresponds to CPS estimates for ages under 65 and HRS estimates for ages over 65.

B2. Mortality, quality of life, and medical spending

We obtain data on mortality, quality of life, and medical spending by health state from the 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.[32] 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 the 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.

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

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.

C. Supporting calculations for numerical models

Appendix C1 provides details regarding the implementation of the deterministic mortality model employed in Section IV.C, 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.B. This model is solved analytically and thus provides exact solutions.

C1. Deterministic mortality

In this model, there is only one health state and we abstract from quality of life. The optimal value function then simplifies to:

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

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 $N_w = 3{,}000$ 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 $V(T-1, W_{T-1}) = \max_{g(W_{T-1})=W_T} u(W_{T-1} + y(T-1) - W_T/e^r) + \frac{1-d(T-1)}{e^{\rho}} V(T, W_T)$. 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 $W(2) = g_1\big(W(1)\big)$, $W(3) = g_2\big(W(2)\big)$,…, $W(T) = g_{T-1}\big(W(T-1)\big)$. Period consumption, $c(t)$, is then calculated using the equation for the budget constraint. Finally, we use the analytical formulas derived in Section II to calculate the value of statistical life.

When accounting for a bequest motive, we follow Kopczuk and Lupton (2007) and assume the utility from leaving a bequest is linear in wealth:

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

Kopczuk and Lupton (2007) estimate that the constant $\alpha^{-\gamma}$ is approximately equal to \$50,000, where γ is the coefficient of relative risk aversion from a CRRA utility function. We adopt a (stronger) estimate of \$35,000 when accounting for a bequest motive. This parameterization implies that the marginal utility of consumption is less than the marginal utility of leaving a bequest when consumption in the last year of life is more than \$35,000.

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 16.5 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 assume that the consumer's lifetime wealth is available at time $t = 0$, so that we can abstract away from the income-generating process. This allows us to generate an analytic solution to the consumer's problem, given by:

$$\max_{c(t)} \mathbb{E}\left[\sum_{t=0}^{T} e^{-\rho t} S_0(t) u\left(c(t), q_{Y_t}(t)\right) + e^{-\rho(t+1)}\left(\big(S_0(t) - S_0(t+1)\big) u(W(t+1), b_t)\right) \middle| Y_0, W_0\right]$$

subject to:

$$W(0) = W_0,$$
$$W(t) \geq 0,$$
$$W(t+1) = \big(W(t) - c(t)\big) e^{r(t, Y_t)}$$

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, as described in the main text. Of course, a constant interest rate $r(t,i) = r$ is included as a special case. The parameter b_t measures the bequest motive. 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. 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 as:

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

subject to:

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

Then we obtain the following Bellman equation:

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

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. In this case, the value function is:

$$\begin{aligned} V(0,w,i) = \sum_{t=0}^{T} e^{-\rho t} \mathbb{E} & \left[\exp\left\{-\int_0^t \mu(s)ds\right\}\left(q_{Y_t}(t)\frac{c(t)^{1-\gamma}}{1-\gamma} - \frac{\underline{c}^{1-\gamma}}{1-\gamma}\right) \middle| Y_0 = i, W(0) = w \right] \\ & + \underbrace{e^{-\rho(t+1)} \mathbb{E}\left[\left(\exp\left\{-\int_0^t \mu(s)ds\right\} - \exp\left\{-\int_0^{t+1} \mu(s)ds\right\}\right)\left(b_t \frac{W(t+1)^{1-\gamma}}{1-\gamma} - \frac{\underline{c}^{1-\gamma}}{1-\gamma}\right) \middle| Y_0 = i, W(0) = w \right]}_{*} \end{aligned} \quad \text{(C1)}$$

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

$$
\begin{aligned}
V(0,w,i) = \sum_{t=0}^{T} e^{-\rho t}\mathbb{E}&\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_{Y_t}(t)\frac{c(t)^{1-\gamma}}{1-\gamma}\middle| Y_0 = i, W(0) = w\right] \\
&+ e^{-\rho(t+1)} b_t \mathbb{E}\left[\left(\exp\left\{-\int_0^t \mu(s)ds\right\} - \exp\left\{-\int_0^{t+1} \mu(s)ds\right\}\right)\frac{W(t+1)^{1-\gamma}}{1-\gamma}\middle| Y_0 = i, W(0) = w\right] \\
&- \frac{\underline{c}^{1-\gamma}}{1-\gamma}\left[1 + e^{-\rho}\sum_{t=0}^{T} e^{-\rho t}\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\}\middle| Y_0 = i\right]\right] \\
= \frac{1}{1-\gamma}&\left[w^{1-\gamma}K_{0,i} - \underline{c}^{1-\gamma}\left[1 + e^{-\rho}\underbrace{\sum_{t=0}^{T} e^{-\rho t}\,\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\}\middle| Y_0 = i\right]}_{\text{life expect. in state } i,\ \text{discounted at rate } \rho}\right]\right]
\end{aligned}
$$

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

$$
VSL(i) = \frac{V(0,w,i)}{u_c\left(wc_{0,i}, q_i(0)\right)} = \frac{V(0,w,i)}{V_w(0,w,i)} \tag{C2}
$$

When bequests are absent and $r(t,i) = r$, we drop the term (*) in equation (C1), and the theory presented in the main text then yields the following expression for VSL:

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

which can also be written as:

$$
VSL(i) = \frac{1}{1-\gamma}\sum_{t=0}^{T} e^{-rt}\underbrace{\frac{\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_{Y_t}(t)c(t)^{1-\gamma}\middle| Y_0 = i, W(0) = w\right] - \underline{c}^{1-\gamma}\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\}\middle| Y_0 = i, W(0) = w\right]}{\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_{Y_t}(t)c(t)^{-\gamma}\middle| Y_0 = i, W(0) = w\right]}}_{v(Y_0,t)} \tag{C3}
$$

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

Appendix Lemma C2: Let $W_{t,j}(\Psi) = \mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} W(t)^{\Psi}\mathbf{1}\{Y_t = j\}\middle| Y_0, W_0\right]$ for $\Psi \in (1,\infty)$. Then $W_{t,j}(\Psi)$ satisfies the following recursion:

$$
\begin{aligned}
&W_{0,Y_0}(\Psi) = W_0^{\Psi}, W_{0,i}(\Psi) = 0, i \neq Y_0, \\
&W_{t+1,j}(\Psi) = e^{r\Psi}\sum_{k=1}^{n} W_{t,k}(\Psi)\left(1 - c_{t,k}\right)^{\Psi}\left(1 - \overline{d}_k(t)\right)p_{k,j}(t)
\end{aligned}
$$

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. Applying **Appendix Lemma C2**, we obtain:

Appendix Proposition C3:

VSL in state Y_0 is equal to:

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

Proof of Appendix Proposition C3: see end of appendix C

We also immediately obtain the following corollary.

Appendix Corollary C4:

The value of a marginal reduction in the probability of transitioning from state i to state j is equal to:

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

We have verified in our numerical calculations that **Appendix Proposition C3** and **Appendix Corollary C4** yield the same answer as the direct evaluation via equation (C2) above.

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:

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

Hence, we obtain:

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

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

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

From the first-order condition we obtain:

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

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:

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

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:

$$\begin{aligned} W_{t+1,j}(\Psi) &= \mathbb{E}\left[\exp\left\{-\int_0^{t+1}\mu(s)ds\right\}\left(W(t+1)\right)^{\Psi}\mathbf{1}\{Y_{t+1}=j\}\middle| Y_0, W_0\right] \\ &= \mathbb{E}\left[\exp\left\{-\int_0^{t}\mu(s)ds\right\}\left(\left(W(t)-c(t)\right)e^{r}\right)^{\Psi}\mathbf{1}\{Y_{t+1}=j\}\exp\left\{-\int_t^{t+1}\mu(s)ds\right\}\middle| Y_0, W_0\right] \end{aligned}$$

$$= \sum_{k=1}^{n} \mathbb{E}\left[\mathbf{1}\{Y_t = k\} \exp\left\{-\int_0^t \mu(s)ds\right\} e^{r\Psi} W(t)^{\Psi} \left(1 - c_{t,k}\right)^{\Psi} \underbrace{\mathbb{E}\left[\mathbf{1}\{Y_{t+1} = j\} \exp\left\{-\int_t^{t+1} \mu(s)ds\right\} \middle| Y_t = k \right]}_{\left(1-\overline{d}_k(t)\right) p_{kj}(t)} \middle| Y_0, W_0 \right]$$

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

QED

Proof of Appendix Proposition C3:

Note that we can rewrite one of the terms in equation (C3) as follows:

$$\begin{aligned} \mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_{Y_t}(t) c(t)^{\Psi} \middle| Y_0, W_0 \right] &= \sum_{j=1}^{n} \mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_{Y_t}(t) c(t)^{\Psi} \mathbf{1}\{Y_t = j\} \middle| Y_0, W_0 \right] \\ &= \sum_{j=1}^{n} \mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} q_j(t) c_{t,j}^{\Psi} W(t)^{\Psi} \mathbf{1}\{Y_t = j\} \middle| Y_0, W_0 \right] \\ &= \sum_{j=1}^{n} q_j(t) c_{t,j}^{\Psi} \underbrace{\mathbb{E}\left[\exp\left\{-\int_0^t \mu(s)ds\right\} W(t)^{\Psi} \mathbf{1}\{Y_t = j\} \middle| Y_0, W_0 \right]}_{W_{t,j}(\Psi)} \end{aligned}$$

The proof follows by setting $\Psi = 1 - \gamma$, 0, and $-\gamma$ and then plugging those results into equation (C3) as appropriate.

QED

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

We assume a full menu of actuarially fair annuities is available, where consumers can choose consumption streams, $c(t)$, that depend on the evolution of their health state. Thus, the consumer is able to fully insure against consumption risk. The consumer's maximization problem is:

$$\max_{c(t)} \mathbb{E}\left[\int_0^T e^{-\rho t} S(t) u\big(c(t), q_{Y_t}(t)\big) dt \,\middle|\, Y_0\right] \tag{D1}$$

subject to:

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

where $\overline{W}(0, Y_0)$ is the net present value of wealth and future earnings.

The consumer chooses the consumption profile at time t based on her health state, $Y_t = i$, and on her available wealth, $\overline{W}(t, i)$. Her available wealth finances future consumption such that:

$$\overline{W}(t, i) = \mathbb{E}\left[\int_t^T e^{-r(u-t)} \exp\left\{-\int_t^u \mu(s) ds\right\} c(u) du \,\middle|\, Y_t, \overline{W}(t, i)\right]$$

Appendix Lemma D1:

The law of motion for wealth is:

$$\frac{\partial \overline{W}(t, i)}{\partial t} = \big(r + \bar{\mu}_i(t)\big) \overline{W}(t, i) - c\big(t, \overline{W}(t, i), i\big) + \sum_{j>i} \lambda_{ij}(t) \big[\overline{W}(t, i) - \overline{W}(t, j)\big], i = 1, \ldots, n$$

Proof of Appendix Lemma D1: see end of Appendix D

Note that the dynamics for $\overline{W}(t, i)$ will depend on $\overline{W}(t, j)$, $j > i$, so that $\big(Y_t, \overline{W}(t, Y_t)\big)$ is not Markov, but $\big(Y_t, \overline{W}(t)\big)$, where we define the wealth vector $\overline{W}(t) \equiv \big(\overline{W}(t, 1), \ldots, \overline{W}(t, n)\big)$, is Markov.

Define the optimal value-to-go function as:

$$V\big(t, \overline{W}(t), Y_t\big) = \max_{c(u)} \mathbb{E}\left[\int_t^T e^{-\rho(u-t)} \exp\left\{-\int_t^u \mu(s) ds\right\} u\big(c(u), q_{Y_u}(u)\big) du \,\middle|\, Y_t, \overline{W}(t)\right]$$

subject to the law of motion for wealth given above. As a stochastic dynamic programming problem, $V(\cdot)$ satisfies the following Hamilton-Jacobi-Bellman (HJB) system of equations:

$$\begin{aligned} \big(\rho + \bar{\mu}_i(t)\big) V\big(t, \overline{W}(t), i\big) = \frac{\partial V\big(t, \overline{W}(t), i\big)}{\partial t} + \max_{c(t)} \Bigg\{ & u\big(c(t), q_i(t)\big) + \sum_{j>i} \lambda_{ij}(t) \big[V\big(t, \overline{W}(t), j\big) - V\big(t, \overline{W}(t), i\big)\big] \\ & + \sum_{k \geq i} \frac{\partial V\big(t, \overline{W}(t), i\big)}{\partial \overline{W}(t, k)} \Bigg[\big(r + \bar{\mu}_k(t)\big) \overline{W}(t, k) - c\big(t, \overline{W}(t, k), k\big) \\ & + \sum_{l>k} \lambda_{kl}(t) \big[\overline{W}(t, k) - \overline{W}(t, l)\big] \Bigg] \Bigg\}, 1 \leq i \leq n \end{aligned} \tag{D2}$$

Similarly to the uninsured case presented in the main text, we follow Parpas and Webster (2013) and focus on the path of Y that begins in i and remains in i until time t, with $c_i(t)$ and $\overline{W}_i(t)$ denoting the corresponding optimal consumption and wealth paths. We take optimal consumption rules and value functions from other states as exogenous. As in the uninsured case, this approach will allow us to apply the standard Pontryagin maximum principle and derive analytic expressions.

Appendix Lemma D2:

The optimal value function for $Y_0 = i$, $V\left(0, \overline{W}(0,i), i\right)$, for the following deterministic optimization problem also satisfies the HJB given by (D2), for each $i \in \{1, \dots, n\}$:

$$V\left(0, \overline{W}(0,i), i\right) = \max_{c_i(t)} \left[\int_0^T e^{-\rho t} \tilde{S}(i,t) \left(u(c_i(t), q_i(t)) + \sum_{j>i} \lambda_{ij}(t) V\left(t, \overline{W}_i(t), j\right) \right) dt \right] \tag{D3}$$

subject to:

$$\frac{\partial \overline{W}_i(t,j)}{\partial t} = \left(r + \overline{\mu}_j(t)\right) \overline{W}_i(t,j) - c\left(t, \overline{W}_i(t), j\right) + \sum_{k>j} \lambda_{jk}(t) \left[\overline{W}_i(t,j) - \overline{W}_i(t,k)\right], j \neq i$$

$$\frac{\partial \overline{W}_i(t,i)}{\partial t} = \left(r + \overline{\mu}_j(t)\right) \overline{W}_i(t,i) - c_i(t) + \sum_{k>i} \lambda_{ik}(t) \left[\overline{W}_i(t,i) - \overline{W}_i(t,k)\right]$$

where $V\left(t, \overline{W}_i(t), j\right)$ and $c\left(t, \overline{W}_i(t), j\right), j > i$, are taken as exogenous.

Proof of Appendix Lemma D2: see end of Appendix D

Following Bertsekas (2005), the Hamiltonian for the (deterministic) maximization problem (D3) is:

$$\begin{aligned} H\left(\overline{W}_i(t), c_i(t), p_i(t)\right) = {} & e^{-\rho t} \tilde{S}(i,t) \left(u(c_i(t), q_i(t)) + \sum_{j>i} \lambda_{ij}(t) V\left(t, \overline{W}_i(t), j\right) \right) \\ & + \sum_{k>i} p_i(t,k) \left[(r + \overline{\mu}_k(t)) \overline{W}_i(t,k) - c\left(t, \overline{W}_i(t), k\right) + \sum_{l>k} \lambda_{kl}(t) \left[\overline{W}_i(t,k) - \overline{W}_i(t,l)\right] \right] \\ & + p_i(t,i) \left[(r + \overline{\mu}_i(t)) \overline{W}_i(t,k) - c_i(t) + \sum_{l>i} \lambda_{il}(t) \left[\overline{W}_i(t,i) - \overline{W}_i(t,l)\right] \right] \end{aligned} \tag{D4}$$

where $p_i(t) = \left(p_i(t,1), \dots, p_i(t,n)\right)$ is the vector of costate variables corresponding to wealth $\overline{W}_i(t)$.

Appendix Lemma D3:

We have that $p_i(t,i) = \theta e^{-\rho t} \tilde{S}(i,t)$ for θ independent of i, and $p_i(t,k) = 0, k \neq i$. The necessary first-order condition for consumption is:

$$e^{(r-\rho)t} u_c(c_i(t), q_i(t)) = \theta \tag{D5}$$

where $\theta = p_i(0,i) = \partial V\left(0, \overline{W}_i(0), i\right) / \partial \overline{W}(0,i)$ is the marginal utility of wealth.

Proof of Appendix Lemma D3: see end of Appendix D

To analyze the values of life and illness, let $\delta_{ij}(t), i,j \leq N$, be a perturbation on the transition intensity $\lambda_{ij}(t)$, and let $\delta_{i,N+1}(t)$ be a perturbation on the mortality rate, $\overline{\mu}_i(t)$, where $\sum_{j=i+1}^{N+1} \int_0^T \delta_{ij}(t) dt = 1$, and consider:

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

Appendix Proposition D4:

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

$$\left.\frac{\partial V}{\partial \varepsilon}\right|_{\varepsilon=0} = \int_0^T \left(\tilde{S}(i,t)\left\{e^{-\rho t}\left[u\left(c_i(t), q_i(t)\right) + \sum_{j>i}\lambda_{ij}(t)V\left(t, \overline{W}_i(t), j\right)\right] + \theta e^{-rt}\left[m_i(t) - c_i(t) - \sum_{j>i}\lambda_{ij}(t)\overline{W}_i(t,j)\right]\right\}\right.$$
$$\left. - \tilde{S}(i,t)\sum_{j=i+1}^{N}\delta_{ij}(t)\left\{e^{-\rho t}V\left(t, \overline{W}_i(t), j\right) - \theta e^{-rt}\overline{W}_i(t,j)\right\}\right) dt \quad \text{(D6)}$$

Proof of Appendix Proposition D4: see end of Appendix D

To obtain the value of statistical life (VSL), we first set $\delta_{i,N+1}$ equal to the Dirac delta function, and set all other perturbations equal to 0. Dividing the result by the marginal utility of wealth, θ, then yields:

$$VSL = \int_0^T \tilde{S}(i,t)e^{-rt}\left\{\left[\frac{u(c_i(t), q_i(t))}{u_c(c_i(t), q_i(t))} + \sum_{j>i}\lambda_{ij}(t)\frac{V\left(t, \overline{W}_i(t), j\right)}{\partial V\left(t, \overline{W}_i(t), j\right)/\partial \overline{W}_i(t,j)}\right]\right. \quad \text{(D7)}$$
$$\left. + \left[m_i(t) - c_i(t) - \sum_{j>i}\lambda_{ij}(t)\overline{W}_i(t,j)\right]\right\} dt$$
$$= \mathbb{E}\left[\int_0^T e^{-rt}S(t)v(t)dt\right]$$

where the value of a statistical life-year is:

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

Comparing (D7) to (3) reveals that generalizing the standard model to account for stochastic mortality alone does not alter the basic expression for VSL. Consumers continue to discount future life-years by the rate of interest and by survival. We can obtain the life-cycle profile of consumption in state i by differentiating the first-order condition (D5) with respect to t. Doing so confirms that, as in the deterministic case, annuitization insulates consumption from mortality risk:

$$\frac{\dot{c}_i(t)}{c_i(t)} = \sigma(r - \rho) + \sigma\eta\frac{\dot{q}}{q}$$

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.

However, a novel feature of the stochastic model is that it permits an investigation into the value of prevention. Inspecting the expression for the marginal utility of life extension (D6), the first term inside the integral represents the gain in marginal utility from a reduction in the probability of exiting state i. The second term represents the loss in marginal utility from the reduction in probability of transitioning to other possible states. The net effect depends on the consumer's marginal utility in the different states.

To analyze the value of prevention, consider a reduction in the transition probability for only one alternative state, j, so that $\delta_{ik}(t) = 0\ \forall k \neq j$. The value of avoiding illness j is then equal to:

$$
\begin{aligned}
VSI(i,j) = \int_0^T \tilde{S}(i,t)e^{-rt} \Bigg\{ &\left[\frac{u(c_i(t), q_i(t))}{u_c(c_i(t), q_i(t))} + \sum_{j>i} \lambda_{ij}(t) \frac{V\big(t, \overline{W}_i(t), j\big)}{\partial V\big(t, \overline{W}_i(t), j\big) / \partial \overline{W}_i(t,j)} \right] \\
&+ \left[m_i(t) - c_i(t) - \sum_{j>i} \lambda_{ij}(t) \overline{W}_i(t,j) \right] \Bigg\} dt - \left[\frac{V\big(t, \overline{W}_i(t), j\big)}{\theta} - \overline{W}_i(0,j) \right] \\
= VSL(i) - VSL\big(j \big| \overline{W}(0,j) = \overline{W}_i(0,j)\big)&
\end{aligned}
\tag{D8}
$$

Thus, equation (D8) demonstrates that $VSI(i,j)$ is equal to the difference in VSL for states i and j, with the caveat that VSL in state j uses a measure of total wealth evaluated from the perspective of a person in state i. This technicality arises because the value of the consumer's annuity depends on her expected survival. For example, an annuity is worth more to a healthy 65-year-old than it is to a 65-year-old who was just diagnosed with lung cancer.

Proofs for Appendix D
Proof of Appendix Lemma D1:

Available wealth can be written as:

$$\overline{W}(t,i) = \int_t^T \exp\left\{-\int_t^u r + \overline{\mu}_i(s) + \sum_{j>i} \lambda_{ij}(s)\, ds\right\}\left[c_i(t,u) + \sum_{j>i} \lambda_{ij}(u)\overline{W}_i(u,t,j)\right] du$$

where with a slight abuse of notation, $c_i(t,u)$ and $\overline{W}_i(u,t,j)$ denote the consumption and wealth paths for an individual who is in state i at time t and remains in state i until time u. The result then follows by taking the derivative with respect to t.

Proof of Appendix Lemma D2:

This proof follows the same logic as the proof of **Lemma 1** in Appendix A. Consider the deterministic optimization problem (D3). Denote the optimal value-to-go function as:

$$\overline{V}\left(t,\overline{W}_i(t),i\right) = \max_{c_i(t)}\left\{\int_t^T e^{-\rho u}\tilde{S}(i,u)\left(u(c_i(u),q_i(u)) + \sum_{j>i} \lambda_{ij}(u)V\left(u,\overline{W}_i(u),j\right)\right) du\right\}$$

Setting $\overline{V}\left(t,\overline{W}_i(t),i\right) = e^{-\rho t}\tilde{S}(i,t)V\left(t,\overline{W}_i(t),i\right)$ then demonstrates that $V(\cdot)$ satisfies the HJB (D2) for i.

QED

Proof of Appendix Lemma D3:

The costate equations for the Hamiltonian (D4) are:

$$\begin{aligned}
\dot{p}_i(t,i) &= -\left[\left(r+\overline{\mu}_i(t)\right) + \sum_{l>i}\lambda_{il}(t)\right]p_i(t,i),\\
\dot{p}_i(t,k) &= -\sum_{j>k}\lambda_{ij}(t)\frac{\partial V\left(t,\overline{W}_i(t),j\right)}{\partial \overline{W}_i(t,k)} + \sum_{k\geq j>i} p_i(t,j)\left(\frac{\partial c\left(t,\overline{W}_i(t),j\right)}{\partial \overline{W}_i(t,k)} + \lambda_{jk}(t)\right)\\
&\quad - p_i(t,k)\left[\left(r+\overline{\mu}_k(t)\right) + \sum_{l>k}\lambda_{kl}(t)\right] + p_i(t,i)\lambda_{ik}(t), \text{for } k>i
\end{aligned}$$

From the first costate equation, we obtain:

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

Taking first-order conditions in the Hamiltonian (D4) and plugging this in then yields:

$$u_c\left(c_i(t),q_i(t)\right) = \frac{\partial V\left(t,\overline{W}_i(t),i\right)}{\partial \overline{W}_i(t,i)} = e^{(\rho-r)t}\theta$$

To see that this solution works, let θ be constant across states, and set $p_i(t,k) = 0 = \frac{\partial V(t,\overline{W}_i(t),i)}{\partial \overline{W}_i(t,k)}$. This then satisfies the costate equation system across i, k, and t. In particular, for the second equation we obtain

$$\dot{p}_i(t,k) = -e^{-\rho t}\tilde{S}(i,t)\lambda_{ik}(t)\underbrace{\frac{\partial V\left(t,\overline{W}_i(t),k\right)}{\partial \overline{W}_i(t,k)}}_{e^{(\rho-r)t}\theta} + \lambda_{ik}(t)p_i(t,i)$$
$$= 0$$

QED

Proof of Appendix Proposition D4:

Starting from equation (D3), we have:

$$V^{\varepsilon}\left(0,\overline{W}_i(0,i),i\right) = \int_0^T e^{-\rho t}\exp\left\{-\int_0^t \overline{\mu}_i(s) + \sum_{j>i}\lambda_{ij}(s) - \varepsilon\sum_{j=i+1}^{N+1}\delta_{ij}(s)\,ds\right\}\left[u\left(c_i^{\varepsilon}(t),q_i(t)\right) + \sum_{j=i+1}^{N}\left[\lambda_{ij}(t)-\varepsilon\delta_{ij}(t)\right]V\left(t,\overline{W}_i^{\varepsilon}(t),j\right)\right]dt,$$

where $c_i^{\varepsilon}(t)$ and $\overline{W}_i^{\varepsilon}(t)$ represent the equilibrium variations in $c_i(t)$ and $\overline{W}_i(t)$ caused by the perturbation, $\delta_{ij}(t)$. Differentiating then yields:

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

Next, note that the budget constraint implies:

$$0 = \left.\frac{\partial W_0}{\partial \varepsilon}\right|_{\varepsilon=0}$$
$$= \frac{\partial}{\partial \varepsilon}\int_0^T e^{-rt}\exp\left\{-\int_0^t \overline{\mu}_i(s) + \sum_{j>i}\lambda_{ij}(s) - \varepsilon\sum_{j=i+1}^{N+1}\delta_{ij}(s)\,ds\right\}\left(c_i^{\varepsilon}(t) - m_i(t) + \sum_{j=i+1}^{N}\left[\lambda_{ij}(t)-\varepsilon\delta_{ij}(t)\right]\overline{W}_i^{\varepsilon}(t,j)\right)dt\Bigg|_{\varepsilon=0}$$
$$= \int_0^T\left(e^{-rt}\tilde{S}(i,t)\left[c_i(t) - m_i(t) + \sum_{j=i+1}^{N}\lambda_{ij}(t)\overline{W}_i(t,j)\right] - e^{-rt}\tilde{S}(i,t)\sum_{j=i+1}^{N}\delta_{ij}(t)\overline{W}_i(t,j)\right.$$
$$\left. + e^{-rt}\tilde{S}(i,t)\left[\left.\frac{\partial c_i^{\varepsilon}(t)}{\partial \varepsilon}\right|_{\varepsilon=0} + \sum_{j=i+1}^{N}\lambda_{ij}(t)\left.\frac{\partial \overline{W}_i^{\varepsilon}(t,j)}{\partial \varepsilon}\right|_{\varepsilon=0}\right]\right)dt$$

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

$$\left.\frac{\partial V}{\partial \varepsilon}\right|_{\varepsilon=0} = \int_0^T \left(\tilde{S}(i,t)\left\{ e^{-\rho t}\left[u\big(c_i(t), q_i(t)\big) + \sum_{j=i+1}^{N} \lambda_{ij}(t) V\big(t, \overline{W}_i(t,j), j\big) \right] + \theta e^{-rt}\left[m_i(t) - c_i(t) - \sum_{j>i} \lambda_{ij}(t)\overline{W}_i(t,j) \right] \right\} - \tilde{S}(i,t)\left\{ e^{-\rho t}\sum_{j=1}^{N} \delta_{ij}(t) V\big(t, \overline{W}_i(t), j\big) - \theta e^{-rt} \sum_{j=i+1}^{N} \delta_{ij}(t)\overline{W}_i(t,j) \right\} \right) dt$$

QED