Supplement to "The distribution of wealth and the marginal propensity to consume"

CHRISTOPHER CARROLL
Department of Economics, Johns Hopkins University

JIRI SLACALEK
DG Research, European Central Bank

KIICHI TOKUOKA
International Bureau, Ministry of Finance, Tokyo, Japan

MATTHEW N. WHITE
Department of Economics, University of Delaware

S1. AGGREGATE MPC IN A LIFE-CYCLE MODEL—FULL EXPOSITION

For ease of exposition and tractability of the aggregate shock processes, the models used in the main paper assume that households have infinite horizons, with no difference between "old" and "young" agents. Our qualitative results hold even when households are instead assumed to live out a finite life cycle, with more realistic assumptions about changes in the income process and mortality as the household ages. This section discusses the assumptions used in an overlapping generations life-cycle specification, and presents analogous results corresponding to the analysis in Section 4 of the paper by reestimating the β-point and β-dist models. In this environment, wealth heterogeneity emerges not only from shocks to permanent and transitory income and differences in discount factors, but also through demographic differences in age and education, via differential mortality and income growth expectations. While these latter factors were abstracted into time preference heterogeneity in our benchmark model, here we model them explicitly to demonstrate the robustness of our results to the simplifying assumptions.

S1.1 *Life cycle of a household*

The economy consists of a continuum of expected utility maximizing households with a common CRRA utility function over consumption, $u(\bullet) = \bullet^{1-\rho}/(1-\rho)$; each household

Christopher Carroll: ccarroll@jhu.edu
Jiri Slacalek: jiri.slacalek@ecb.europa.eu
Kiichi Tokuoka: kiichi.tokuoka@mof.go.jp
Matthew N. White: mnwecon@udel.edu

has a time discount factor β. A household enters the economy at time t aged 24 years, endowed with an education level $e \in \{D, HS, C\}$ (for dropout, high school, and college, respectively), an initial permanent income level $\boldsymbol{p}_0$, and a stock of capital k_0. Each quarter, the household receives (after tax) income, chooses how much of their market resources to consume and how much to save, and then transitions to the next quarter by facing shocks to mortality and income.

A household receives a permanent shock to income when transitioning into period t, denoted by ψ_t (along with the age–education-specific average growth factor $\overline{\psi}_{es}$), as well as a transitory shock ξ_t. The income process can be summarized by

$$\boldsymbol{y}_t = (1 - \tau)\xi_t \boldsymbol{p}_t,$$

$$\boldsymbol{p}_t = \psi_t \overline{\psi}_{es} \boldsymbol{p}_{t-1}.$$

Households that have already lived for s periods have permanent shocks drawn from a log-normal distribution with mean 1 and variance $\sigma^2_{\psi s}$, and transitory shocks drawn from a log-normal distribution with mean $1/\cancel{\mu}$ and variance $\sigma^2_{\xi s}$ with probability $\cancel{\mu} = (1 - u)$ and a degenerate distribution at μ with probability u. The prospect of unemployment (at rate u) is a completely transitory event: unemployment in period t has no effect on the probability of unemployment in period $t + 1$. The nonzero transitory shock when unemployed represents a welfare benefit funded by income taxes, as discussed below. When transitioning from one period to the next, a household with education e that has already lived for s periods faces a D_{es} probability of death. In the main specification, the assets of a household that dies are completely taxed by the government to fund activities outside the model.[1]

The household's permanent income level will be factored out from the problem, so that the only state variable that affects the choice of optimal consumption is normalized market resources m_t. After this normalization, the household's budget transition functions can be described by

$$a_t = m_t - c_t, \tag{S1}$$

$$k_{t+1} = a_t / \psi_{t+1}, \tag{S2}$$

$$m_{t+1} = (\daleth + r)k_{t+1} + (1 - \tau)\xi_{t+1}, \tag{S3}$$

$$a_t \geq 0. \tag{S4}$$

These transition constraints are identical to the infinite-horizon model except that capital owned by surviving households does not grow with the inverse survival probability, and income is taxed at a marginal rate τ depending on the household's age and employment status.

Starting from some terminal age $\overline{s}$ at which $\mathsf{D}_{e\overline{s}} = 1$, a household's problem can be solved by backward induction until $s = 0$. At age $\overline{s}$, the household will consume all market resources, generating a consumption function of $\boldsymbol{c}_{e\overline{s}}(\boldsymbol{m}_t, \boldsymbol{p}_t) = \boldsymbol{m}_t = m_t \boldsymbol{p}_t$ and a value

---

[1] As a further robustness check, we also estimate versions in which the assets of the newly deceased are distributed to a random household, with varying preferences for bequests.

function of $V_{e\bar{s}}(\boldsymbol{m}_t, \boldsymbol{p}_t) = \mathrm{u}(\boldsymbol{m}_t) = \boldsymbol{p}_t^{1-\rho}\mathrm{u}(m_t)$. At any earlier age, the value function is recursively defined by

$$V_{es}(\boldsymbol{m}_t, \boldsymbol{p}_t) = \max_{\boldsymbol{c}_t} \mathrm{u}(\boldsymbol{c}_t) + \beta\, \cancel{D}_{es}\mathbb{E}_t\big[V_{es+1}(\boldsymbol{m}_{t+1}, \boldsymbol{p}_{t+1})\big] \quad \text{s.t. (S1)–(S4).} \tag{S5}$$

To eliminate the permanent income level as a state variable, further define the normalized consumption function as $\mathrm{c}_{es}(m_t) = \boldsymbol{c}_{es}(\boldsymbol{m}_t, \boldsymbol{p}_t)/\boldsymbol{p}_t$ and the normalized value function as $\mathrm{v}_{es}(m_t) = V_{es}(\boldsymbol{m}_t, \boldsymbol{p}_t)/\boldsymbol{p}_t^{1-\rho}$. Dividing (S5) by $\boldsymbol{p}_t^{1-\rho}$, the problem is reduced to a single state dimension and can be expressed as

$$\mathrm{v}_{es}(m_t) = \max_{c_t} \mathrm{u}(c_t) + \beta\, \cancel{D}_{es}\mathbb{E}_t\big[\psi_{t+1}^{1-\rho}\mathrm{v}_{es+1}(m_{t+1})\big] \quad \text{s.t. (S1)–(S4),} \tag{S6}$$

$$\mathrm{c}_{es}(m_t) = \arg\max_{c_t} \mathrm{u}(c_t) + \beta\, \cancel{D}_{es}\mathbb{E}_t\big[\psi_{t+1}^{1-\rho}\mathrm{v}_{es+1}(m_{t+1})\big] \quad \text{s.t. (S1)–(S4).}$$

A standard envelope condition applies in this model, so that $\mathrm{v}'_{es}(m_t) = \mathrm{u}'(\mathrm{c}_{es}(m_t))$, and the first-order condition for the solution to (S6) is

$$c_t^{-\rho} = (\mathbf{1}+r)\beta\, \cancel{D}\,\mathbb{E}_t\big[(\psi_{t+1}c_{t+1})^{-\rho}\big]. \tag{S7}$$

In this way, the value function need not be tracked or recorded during the solution process, as the age-dependent consumption functions are sufficient.[2]

S1.2 *Macroeconomic dynamics*

The analysis in Section 4.3 demonstrated that while there is considerable variation in the marginal propensity to consume across income, wealth, and employment status, the MPC does not appreciably change depending on the structure of aggregate shocks to the economy or to the current macroeconomic state. Moreover, for reasons previously discussed, it is fairly difficult to account for macroeconomic state variables in an overlapping generations model. Rather than expend significant energy on a feature that would yield little of interest, we do not model aggregate shocks in this section but instead focus on the effects of idiosyncratic shocks and household-level dynamics. However, there are some additional macroeconomic features of the model that warrant discussion.

Unlike the infinite-horizon perpetual youth model, the economy is now perpetually growing, with each new cohort larger than the last and ongoing technological progress. The expected permanent income growth for a household $\overline{\psi}_{es}$ comprises the household's own effective labor supply growth plus technological growth. When aggregating wealth, the contribution of a household that has already lived for s quarters is thus discounted by a factor of $(1 + \Gamma)^{-s}$ relative to the youngest cohort, where Γ is the technological growth rate. Moreover, older households were born into smaller cohorts relative to the newest generation, so our population weighting scheme scales their contribution by the population growth rate N.

---

[2]In practice, we use the method of endogenous grid points, as originally described in Carroll (2006), to discretize the state space and approximate consumption functions at each age and education level.

As mentioned in Section S1.1, households are subject to a tax rate of τ depending on their age and employment status. Households are assumed to retire at age 65 (i.e., when $s = 164$), captured in the model with an expected permanent growth factor well below 1 at this age.[3] Income before retirement is earned through labor, while income after retirement is provided by a pay-as-you-go social security system funded by taxes on the employed. The social security tax rate is calculated as the rate that balances outlays to retired households and tax revenues from the working population:

$$\tau_{SS} = \frac{\displaystyle\sum_{e\in\{D,\mathrm{HS},C\}}\left[\theta_e\overline{\boldsymbol{p}}_{e0}\sum_{t=164}^{384}\left(\left((1+\Gamma)(1+N)\right)^{-t}\prod_{s=0}^{t}(\overline{\psi}_{es}\cancel{\mathcal{D}}_{es})\right)\right]}{\displaystyle\sum_{e\in\{D,\mathrm{HS},C\}}\left[\theta_e\overline{\boldsymbol{p}}_{e0}\sum_{t=0}^{163}\left(\left((1+\Gamma)(1+N)\right)^{-t}\prod_{s=0}^{t}(\overline{\psi}_{es}\cancel{\mathcal{D}}_{es})\right)\right]}.$$

Here, θ_e is the proportion of each new generation with education level e, and $\overline{\boldsymbol{p}}_{e0}$ is the average permanent income of that education type when they enter the economy at age 24. Note that neither permanent nor transitory shocks are relevant, as they simply average to 1 across a cohort. The tax to fund unemployment benefits is simply the product of the unemployment rate and the benefit replacement rate: $\tau_U = u\mu$. Employed households pay a total income tax rate of $\tau = \tau_{SS} + \tau_U$, while unemployed and retired households have $\tau = 0$.

S1.3 *Simulation*

When the economy is simulated, each household represents a population mass rather than an atomic agent. A household is simulated for the full course of life, with a pre-determined set of permanent and transitory shocks to apply in each quarter.[4] Mortality itself is not simulated, but instead accounted for through the population weight used for that mass during aggregation. That is, the relative population weight of a household with education type e that has lived for t periods is $w = \theta_e(1+N)^{-t}\prod_{s=0}^{t}\cancel{\mathcal{D}}_{es}$. Every period of a simulated household's life is included in the aggregation of total capital, so that the life history of one population mass is reinterpreted as the present state of a population mass from each preceding generation with positive probability of survival.

Indexing the simulated households by i, total capital in the economy is calculated as

$$\boldsymbol{K} = \sum_{i}\left[\theta_i\boldsymbol{p}_{i0}\sum_{t=0}^{384}\left(k_{it}\left((1+\Gamma)(1+N)\right)^{-t}\prod_{s=0}^{t}(\psi_{is}\overline{\psi}_{is}\cancel{\mathcal{D}}_{is})\right)\right].$$

Assuming that our simulated population size is sufficiently large that the actual draws of ψ_{is} and ξ_{is} simply average to unity, the economy's income can be approximated as the

---

[3]The drop in permanent income at retirement depends on the household's education: dropouts' income falls by 44%, high school graduates by 56%, and college graduates by 69%.

[4]Estimation of the model requires solving the model simulating the economy thousands of times. Both to reduce computational burden and to ensure convergence of the estimation procedure, we apply the same set of shocks on each simulation.

denominator of the social security tax rate τ_{SS}:

$$Y \approx \sum_{e \in \{D, \text{HS}, C\}} \left[\theta_e \overline{P}_{e0} \sum_{t=0}^{163} \left(((1+\Gamma)(1+N))^{-t} \prod_{s=0}^{t} (\overline{\psi}_{es} \mathcal{D}_{es}) \right) \right].$$

Given that we simulate 2000 households for 384 quarters for each education level (for a total of over 2.3 million population masses), the approximation is nearly perfect.[5]

S1.4 *Calibration*

Calibrations of the distributional parameters are taken from related estimates in the literature. Average permanent income growth rates $\overline{\psi}_{es}$ are calculated using the same trajectories as in Cagetti (2003) for those who have less than a high school education, a high school degree, and four or more years of college. The permanent and transitory shock variances are approximated from the results of Sabelhaus and Song (2010), with extrapolation for ages 55–64.[6] Households are assumed to retire at age 65, withdrawing from the labor force and only receiving income from a pay-as-you-go social security system financed by taxes on the working population. Baseline mortality rates at each age are taken from the Social Security Administration's (2010) Actuarial Life Table,[7] then adjusted by education level using estimates by Brown, Liebman, and Pollett (2002) and converted to quarterly probabilities;[8] households die with certainty if they reach age 120. The unemployment benefit μ is set to 0.15 to match Cagetti (2003), while the unemployment probability is $u = 7\%$, the average rate in the infinite-horizon model.

We assume that the population grows at a rate of 1% annually, while total factor productivity grows at a 1.5% annual rate; these approximately match long run rates in the United States. Educational attainment rates are set to be fairly consistent with U.S. educational rates over the past 20 years, and average initial permanent (quarterly) incomes at age 24 for each educational group are roughly calibrated to recent data.[9] Each simulated household is given an initial log-normal shock to permanent income with standard deviation 0.4, approximately matching the total variance of income among young households in the SCF 2004 data. Households begin with a very low wealth-to-permanent income ratio, drawn uniformly from {0.17, 0.50, 0.83}. Other basic parameters are set to match the values used in the infinite-horizon model. A summary of the model parameters is provided in Table S1.

---

[5]In the β-dist version of the overlapping generations model, we use 10 different discount factors, so that there are 30 types of households and over 23 million population masses.

[6]We assume that $\sigma_{\psi}^2 = \sigma_{\xi}^2 = 0$ in retirement, so there is no income risk.

[7]Following the bulk of related literature, we use women's mortality rates to allow households to simulate households living past the husband's death.

[8]For ages 101–120, we use the adjustment for age 100, as this table does not extend to very late ages to which very few people live.

[9]Precision here is unimportant: after several years of simulation, the initial permanent income differences between types matters much less than their income growth trajectories and idiosyncratic shocks.

TABLE S1. Parameter values in life-cycle model.

Description	Parameter	Value
Coeff of relative risk aversion	ρ	2
Effective interest rate	$(r - \delta)$	0.01
Population growth rate	N	0.0025
Technological growth rate	Γ	0.0037
Rate of high school dropouts	θ_D	0.11
Rate of high school graduates	θ_{HS}	0.55
Rate of college graduates	θ_C	0.34
Average initial permanent income, dropout	$\overline{p}_{D0}$	5000
Average initial permanent income, high school	$\overline{p}_{HS0}$	7500
Average initial permanent income, college	$\overline{p}_{D0}$	12,000
Unemployment insurance payment	μ	0.15
Unemployment rate	u	0.07
Labor income tax rate	τ	0.0942

S1.5 *Aggregate marginal propensity to consume*

Following the same procedure as in the benchmark infinite-horizon model, we first assume that all households have the same time preference factor $\grave{\beta}$ as in the β-point model. Seeking the value of $\grave{\beta}$ at which the aggregate capital-to-income ratio K/Y matches that of the perfect foresight version of the infinite-horizon model the aggregate capital-to-income ratio K/Y of 10.26 matches, we find $\grave{\beta} = 0.9911$. As in the infinite-horizon model, the simulated distribution of wealth in the β-point life-cycle model matches the empirical distribution from the 2004 Survey of Consumer Finances considerably better than the KS-JEDC model; indeed, the life-cycle model has a somewhat better fit than the infinite-horizon model, moving over 40% of the way from the KS-JEDC's Lorenz curve to the empirical distribution, rather than one-third. The additional wealth heterogeneity arises through differences in households' expectations of the future that were suppressed in the infinite-horizon model: income growth rates vary with both education and age (particularly the timing of retirement), while the increasing probability of death plays a key role in older households' target wealth-to-income ratio.

To better fit the empirical distribution of wealth, we again estimate the β-dist model by assuming that each household has its own discount factor drawn from a uniform distribution, $\beta \in [\grave{\beta} - \nabla, \grave{\beta} + \nabla]$, approximated by a 10 point discrete distribution.[10] As in Section 4, the object of the estimation is to find parameters $\grave{\beta}$ and ∇ that minimize the distance between empirical and simulated shares of wealth held by the richest 20%, 40%, 60%, and 80%, subject to the constraint that the simulated K/Y ratio equals the target ratio of 10.26, as in the β-point model. Estimation reveals that the optimal parameters are $\{\grave{\beta}, \nabla\} = \{0.9775, 0.0391\}$, quite similar to the infinite-horizon es-

---

[10]We implicitly assume that the distribution of discount factors is independent from education type. More realistically, individuals with higher discount factors are more likely to remain in school longer. As this would tend to make high income types retain even more assets for the future while low income types will save even less, we would need a narrower range of discount factors to match observed wealth inequality. For this reason and for the sake of simplicity, we ignore this complicating factor.
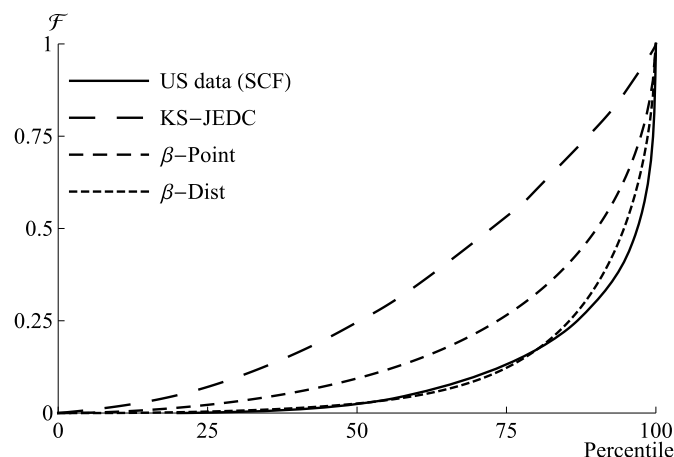
FIGURE S1. Distribution of net worth (Lorenz curve): Life -cycle model.

timates of {0.9751, 0.0238} in the infinite-horizon model with Krusell–Smith macroeconomic shocks. Figure S1 presents the empirical Lorenz curve based on SCF 2004 data alongside the simulated curves for the β-point and β-dist models (and the original KS-JEDC wealth distribution).

The β-dist model is able to match the empirical Lorenz curve extremely well for the bottom 85% of the wealth distribution: the average difference between simulated and actual wealth shares at the levels of interest is less than 0.4%. Indeed, the life-cycle model is able to match the extremely low asset holdings of the bottom half of the population significantly better than the infinite-horizon model. However, the simulated share of wealth held by the very richest individuals is somewhat lower than the SCF data show; in contrast, the infinite-horizon β-dist model is able to match the Lorenz curve fairly well even in the top 10%. We surmise that this is due to differences in the distribution of income at the upper tail between the models. While about 8% of households in the infinite-horizon model have working lives that exceed 100 years, allowing some of these households to attain extremely high permanent incomes through a series of fortuitous shocks,[11] individuals in the overlapping generations/finite-horizon model retire with certainty after 41 working years, receiving a large negative shock to permanent income and experiencing no income shocks thereafter. Without the ability to generate extremely high income households, even the richest do not accumulate enough assets to match the top end of the empirical Lorenz curve. As the simulated distribution matches fairly well and we are concerned with the aggregate marginal propensity to consume, particularly among nonwealthy households, this is not a serious deficiency.

Across all households, the aggregate (annual) marginal propensity to consume in both the β-point (0.08) and β-dist (0.29) models is extremely similar to the correspond-

---

[11]While our results do not critically depend on the inclusion of these very long-lived households, the model requires a somewhat wider distribution of β when they are removed from the aggregation.

ing averages in the infinite-horizon model.[12] Table S2 presents average MPCs for sub-populations.[13] The relationship between wealth-to-permanent income and the MPC is nearly identical to the pattern in the infinite-horizon case, with the MPC slowly rising with lower incomes among the wealthier half of the population, and spiking rapidly among the bottom half. However, the gradient of income to MPC is much shallower in the finite-horizon model, with the wealthiest 1% of households' MPC only 20% less than the poorest half, rather than 50% less in the benchmark model. This is likely due to confounding effects from life-cycle dynamics: income-poor households in the overlapping generations model are made up of both the young (who have not had time to accumulate income growth) and the retired (whose cohorts began with lower initial permanent income and have experienced the large negative wage growth from retirement).

While the ratio of wealth to permanent income is a very strong determinant of the marginal propensity to consume in our model, the wide distribution of household incomes allows for even wealthy households to have high MPCs. Confirming a similar exercise in the benchmark model, we again find that among the one-third of households with the highest MPCs, 50% are in the lowest wealth quintile, 32% are in the second wealth quintile, and 15% are in the middle wealth quintile. Even in a life-cycle model in which wealth is highly correlated with both age and the marginal propensity to consume, there is still a significant fraction of "wealthy hand-to-mouth" households as found in Kaplan, Violante, and Weidner (2014).

Figure S2 shows cumulative distribution functions for the KS-JEDC, KS-hetero, and 2 beta-dist models; the first matching the empirical distribution of net worth, the second matching the distribution of liquid financial and retirement assets.

Figure S3 presents the aggregate marginal propensity to consume by age for the entire population, as well as for the most patient and least patient types in the β-dist model. After an initial drop as households build up a minimum buffer stock, the life-cycle profile of the MPC takes an inverted U-shape for most β types: rising during the rapid income growth ages of 30–50 before falling as households anticipate their retirement and seek to retain assets to consume in old age. Post retirement, the MPC steadily grows as agents experience an ever increasing mortality risk. The most impatient households, with a quarterly discount factor of about $\beta = 0.9384$, have a significantly higher MPC throughout life as they disfavor saving: they begin saving for retirement less than 10 years prior and quickly deplete their assets if they live beyond age 75 (as evidenced by MPCs approaching 1 at these ages). In contrast, the most patient households show an increasing marginal propensity to consume for their entire lives, though beginning from very low levels.

---

[12]When the annual marginal propensity to consume is calculated by simulating the change in consumption over four quarters resulting from an unexpected $1000 payment to each household, we find an aggregate value of 0.28, substantially the same and confirming the same exercise in the infinite-horizon model.

[13]Without much comment, we also present estimates of the β-dist model when matching the empirical distribution of liquid financial and retirement assets rather than net worth, along with subpopulation average MPCs for these models. In each case, the results of the life-cycle model align very well with the earlier findings in the infinite-horizon setting.

TABLE S2. Average (aggregate) marginal propensity to consume in annual terms, finite horizon life cycle

Specification	Bequests Fully Taxed, no Bequest Motive			Bequests Distributed, no Bequest Motive			Bequests Distributed, With Bequest Motive		
	β-Point	β-Dist	β-Dist	β-Point	β-Dist	β-Dist	β-Point	β-Dist	β-Dist
Model Wealth Measure	Net Worth	Net Worth	Liquid Assets	Net Worth	Net Worth	Liquid Assets	Net Worth	Net Worth	Liquid Assets
Overall average	0.08	0.29	0.44	0.09	0.28	0.43	0.09	0.28	0.43
By wealth-to-permanent income ratio									
Top 1%	0.09	0.06	0.06	0.09	0.06	0.07	0.09	0.06	0.07
Top 10%	0.08	0.07	0.06	0.09	0.07	0.07	0.09	0.07	0.07
Top 20%	0.09	0.06	0.07	0.09	0.07	0.07	0.09	0.07	0.08
Top 30%	0.08	0.06	0.08	0.09	0.07	0.09	0.09	0.07	0.10
Top 40%	0.08	0.07	0.12	0.08	0.08	0.13	0.08	0.08	0.14
Top 50%	0.08	0.08	0.17	0.08	0.10	0.18	0.08	0.10	0.18
Top 60%	0.08	0.10	0.22	0.08	0.12	0.23	0.08	0.12	0.23
Bottom 50%	0.09	0.50	0.71	0.10	0.47	0.68	0.10	0.46	0.67
By income									
Top 1%	0.07	0.24	0.38	0.08	0.29	0.42	0.08	0.29	0.42
Top 10%	0.07	0.26	0.39	0.09	0.30	0.43	0.09	0.30	0.43
Top 20%	0.08	0.26	0.40	0.09	0.30	0.43	0.09	0.30	0.44
Top 30%	0.08	0.27	0.40	0.09	0.30	0.43	0.09	0.30	0.44
Top 40%	0.08	0.27	0.41	0.09	0.30	0.43	0.09	0.30	0.44
Top 50%	0.08	0.27	0.41	0.09	0.30	0.44	0.09	0.30	0.44
Top 60%	0.08	0.27	0.41	0.09	0.30	0.44	0.09	0.30	0.44
Bottom 50%	0.09	0.31	0.47	0.09	0.27	0.43	0.09	0.26	0.42
By employment status									
Employed	0.08	0.29	0.43	0.09	0.28	0.42	0.09	0.27	0.42
Unemployed	0.10	0.39	0.57	0.11	0.39	0.57	0.12	0.40	0.58
Time preference parameters									
$\grave{\beta}$	0.9911	0.9775	0.9576	0.9899	0.9709	0.9491	0.9898	0.9701	0.9468
∇		0.0391	0.0606		0.0483	0.0717		0.0494	0.0748

Note: "Liquid Assets" refers to liquid financial plus retirement assets, as in the text. The "with bequest motive" specification uses bequest parameters of $\alpha = 1$, $\nu = \rho$, and $\gamma = 0$.
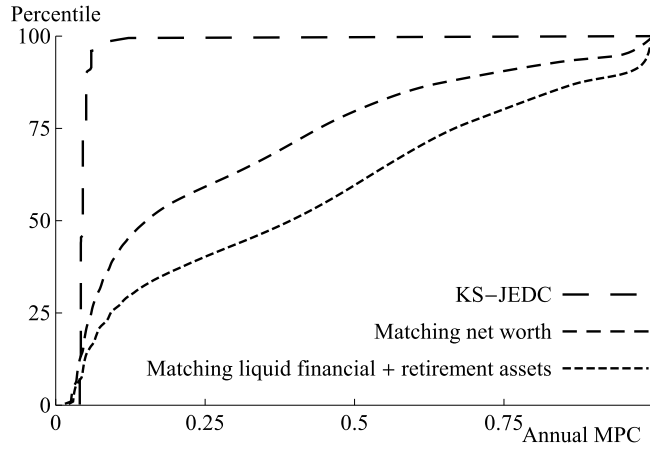
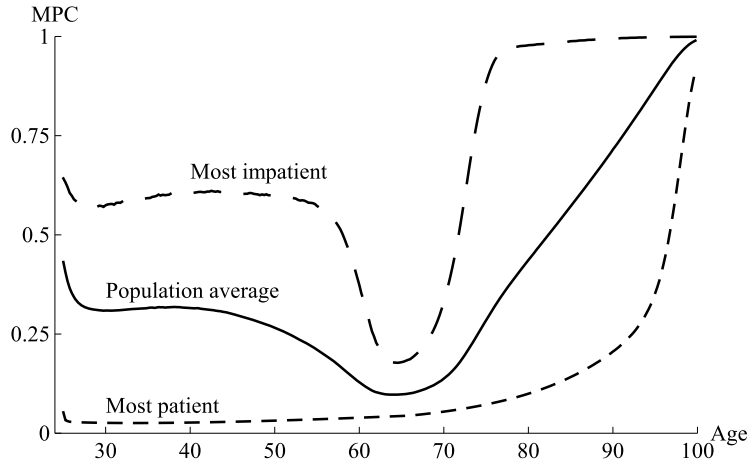Figure S2. Distribution of MPC's across households: Life-cycle model.

Figure S3. Aggregate marginal propensity to consume by age.

S1.6 *Distribution of bequests*

The assumption that the assets of the newly deceased are fully taxed away can be relaxed without disturbing our main results. In the following analyses, we instead assume that these estates are distributed randomly among the population as bequests to other households. Under a wide range of parameters governing preferences over bequests, both the overall aggregate marginal propensity to consume and its decompositions by wealth and income are little changed from the original specification.

When a household dies at the beginning of time t, it receives a one time utility flow[14] of

$$\boldsymbol{p}_t^{1-\rho}\grave{\mathrm{u}}(k_t) = \alpha\frac{(k_t+\gamma)^{1-\nu}}{1-\nu}.$$

---

[14]With apologies to the French, our notation is selected to denote "grave utility."

This form combines the approaches of Cagetti (2003) and Carroll (2000) for maximum flexibility: the factor α controls the absolute magnitude of the bequest motive,[15] the shifter γ allows the bequest motive to be completely inactive for lower wealth households, and the curvature parameter ν determines the extent to which bequests are a luxury good (through the ratio ρ/ν). More complete discussion of the parametrization can be found in the respective sources. With the bequest motive, the household's consumption Euler equation can be updated to

$$c_t^{-\rho} = (\rceil + r)\beta\big(\cancel{D}\mathbb{E}_t\big[(\psi_{t+1}c_{t+1})^{-\rho}\big] + \mathsf{D}\alpha(k_{t+1} + \gamma)^{-\nu}\big) \quad \text{s.t. (S1)–(S4).}$$

The household's problem can be solved by backward induction as before.

As before, the model economy is simulated using a large number of population masses, with each period of a household's life interpreted as the current state of a similar household born in an earlier cohort. Population masses do not "die" during the simulation—they are not discontinued subject to a mortality shock—but rather merely dwindle in population weight as the probability of death accumulates (i.e., the mass represents a smaller collection of similar households). Each age of each simulated mass is assigned a random mass-age (weighting by population mass size) as the beneficiary of bequests given at the end of that period. If mass i at age s is assigned to bequest to mass j at age n, then the ratio of the bequest to mass j's permanent income at age n is

$$b_{is} = \frac{\theta_i \boldsymbol{p}_{i0} \prod\limits_{t=0}^{s}(\cancel{D}_{it}\psi_{it})}{\theta_j \boldsymbol{p}_{j0} \prod\limits_{t=0}^{n}(\cancel{D}_{jt}\psi_{jt})}\big((1+N)(1+\Gamma)\big)^{s-n}k_{is}. \tag{S8}$$

A household might receive multiple bequests at a single age (not uncommon at earlier ages, when population weights are relatively high) or receive none at all. Bequests are received as a second transitory shock to income, but households do not anticipate this possibility and thus do not account for it in their maximization problem.

As our procedure simulates all population masses simultaneously, the bequest that should be received by a household might be unknown when that age is reached, as it is coming from a population mass at a later age that has not yet been simulated. To ensure a consistent economy in which the bequests received equal the bequests actually given, we take an iterative approach and estimate the model several times, updating bequests between estimations. To estimate the extended model, we first set all bequests received to zero and then use the following procedure.[16]

---

[15]An individual who is certain to die will have an MPC of approximately $1/(1 + \alpha^{1/\rho})$ when $\gamma = 0$ and $\nu = \rho$.

[16]In practice, simulating the model and updating bequests in Steps 2 and 3 is orders of magnitude faster than checking a convergence condition for bequests. Rather than evaluating a traditional convergence metric, we instead repeat Steps 2 and 3 20 times, more than enough for convergence to occur. The entire estimation process converges within economically relevant bounds after 6 applications of Step 1, and is thus not overly burdensome.

Step 1. Estimate the model assuming the current set of bequests.

Step 2. Using the simulated values of k_{is} at the estimated parameters, update the bequests given by each mass-age using (S8); then sum bequests received by each mass-age.

Step 3. Simulate the model at the estimated parameters and current set of bequests. Return to Step 1 if bequests have not converged to consistent levels.

Step 4. If the new parameter estimates are sufficiently close to the previous estimates (if any), stop; else return to Step 1.

We estimate the model at several sets of bequest motive parameters,[17] matching simulated to empirical distributions of both net worth and liquid financial plus retirement assets; the results are reported in additional columns of Table S2. First, we estimate a specification in which wealth is distributed to survivors, but households attain no utility from providing these bequests: α is set to zero, suppressing the bequest motive. Next, we set $\alpha = 1$, $\nu = \rho$, and $\gamma = 0$ as simple choices for a standard bequest motive. Finally, we test "extreme" alternatives by turning the scale of the bequest motive very high ($\alpha = 9$, $\nu = \rho$, $\gamma = 0$) and by specifying bequests as a strong luxury good ($\alpha = 1$, $\nu = \rho/2$, $\gamma = 4$); results of these alternatives are not presented due to space considerations, but are available by request.

As seen in the latter columns of Table S2, the distribution of the marginal propensity to consume in the simulated population is substantially similar whether the assets of the dead are taxed or distributed as bequests, or whether households receive any utility benefit from bequeathed wealth, with only small differences. Notably, the presence of bequests pushes the estimation to find both a lower median discount factor $\grave{\beta}$ (0.9775 vs. 0.9709 in the case of net worth) and a wider interval ∇ (0.0391 vs. 0.0483). The bottom half of households in the SCF 2004 data hold a very small share of wealth, but in the simulated economy even the poorest households will occasionally receive a substantial bequest. This tends to boost the share of wealth held by the poorest 20% and 40% above the empirical value, requiring the range of β to reach lower for the most impatient households to hold sufficiently low assets to offset this effect.

The variation in average MPC by income is considerably smaller in the presence of bequests, and is almost perfectly flat in some cases. If a household has received a nontrivial bequest in recent periods, it is less apt to have its consumption habits affected by transitory shocks to income. More strongly, the random distribution mechanism for bequests greatly weakens the link between income and assets, as income-poor households are equally likely to receive large bequests (and when they do, the bequest will be larger relative to their permanent income). As the household's consumption problem is homothetic in permanent income, this weakened relationship reduces the predictive power of income on the MPC.

We report the results of only the "simple" parameters of the bequest motive in Table S2, omitting the more extreme variants. Just as the distribution of the marginal

---

[17]These parameters cannot be meaningfully estimated using our present estimation. Identification of the three parameters would require matching the distribution of wealth at each age, while we are only concerned with the overall distribution. The example parametrizations in this section are meant to demonstrate our results' robustness, not as an economic claim.

propensity to consume varies little between the "no bequest motive" and "with bequest motive" specifications, there is essentially no difference in the MPC as the parameters are shifted about; estimates of the range of discount factors are barely affected. In short, our key point—that the inclusion of an income process with both permanent and transitory shocks allows a dynamic model to better match microeconomic evidence on the aggregate marginal propensity to consume—relies on neither the infinite-horizon assumptions, the specification of macroeconomic shocks, nor the particular treatment of bequests. With a realistic income process and sufficient discount factor heterogeneity to match the empirical distribution of wealth, our results are robust to structural assumptions.

S2. Differences between β-dist and KS-hetero models

This section clarifies the channels underlying the differences in MPCs between our preferred model (β-dist) and the KS-hetero model: While in our model the aggregate MPC is quite high (0.21 when calibrated to the net worth distribution), in the KS-hetero model the aggregate MPC is only 0.09.

We address two points:

(i) We analyze in more detail how sensitive the aggregate MPC is to shutting down the permanent (ψ) and transitory (θ) income shocks.

(ii) We investigate the assumption of the uniform distribution we impose on the discount factor β, and quantify to what extent the assumption drives the results relative to allowing for the permanent–transitory (FBS) income process.

To address these points, we estimated additional specifications, which are summarized in Table S3. Specifications 1–8 turn off the permanent and transitory income shocks, and the borrowing constraint; specifications 9–12 assume that β is log-normally distributed (rather than uniformly).

To preview our results, we find that turning off transitory shocks does not noticeably affect the MPC. Turning off the permanent shocks *and* allowing for borrowing reduces the aggregate MPC from 0.21 to 0.14, each of these two items contributing roughly the same to the decline, compared with the KS-hetero value of 0.09. Finally, assuming a log-normal distribution for β does not affect the results on MPC much (compared to the baseline, uniform distribution).

S2.1 *The FBS income process: A summary*

The FBS household income process $\boldsymbol{y}$ includes the permanent component ψ and the transitory shock ξ_t:

$$\boldsymbol{y}_t = p_t \xi_t \mathsf{W}_t.$$

The permanent component follows a geometric random walk,

$$p_t = p_{t-1} \psi_t, \tag{S9}$$

TABLE S3. Marginal Propensity to Consume, Alternative Specifications

	Baseline		Alternative Income Shocks & Borrowing Constraint						Alternative Distribution of β			
Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Perm shocks ψ?	Y	Y	N	N	Y	Y	N	N	Y	Y	N	Y
Trans shocks θ?	Y	N	Y	N	Y	N	Y	N	Y	N	Y	Y
Borrowing $a_t \geq -2$?	N	N	N	N	Y	Y	Y	Y	N	N	N	Y
β Distribution†	U	U	U	U	U	U	U	U	L	L	L	L
Overall Average	0.21	0.21	0.18	0.18	0.18	0.17	0.14	0.14	0.19	0.19	0.17	0.16
By wealth–permanent income ratio												
Top 1%	0.06	0.06	0.05	0.05	0.06	0.06	0.05	0.05	0.06	0.06	0.05	0.06
Top 10%	0.06	0.06	0.05	0.05	0.06	0.06	0.05	0.05	0.06	0.06	0.05	0.06
Top 20%	0.06	0.06	0.06	0.06	0.07	0.06	0.06	0.06	0.07	0.07	0.06	0.07
Top 20−−40%	0.08	0.08	0.07	0.06	0.08	0.07	0.07	0.06	0.08	0.08	0.08	0.08
Top 40−−60%	0.15	0.14	0.13	0.13	0.11	0.11	0.11	0.10	0.13	0.12	0.12	0.11
Top 60−−80%	0.27	0.28	0.21	0.22	0.21	0.22	0.16	0.17	0.23	0.23	0.20	0.19
Bottom 20%	0.50	0.49	0.43	0.42	0.41	0.40	0.33	0.32	0.47	0.46	0.41	0.38
By income												
Top 1%	0.19	0.19	0.16	0.20	0.16	0.16	0.13	0.15	0.17	0.18	0.16	0.15
Top 10%	0.19	0.20	0.15	0.19	0.16	0.16	0.13	0.14	0.17	0.18	0.15	0.15
Top 20%	0.20	0.20	0.15	0.18	0.16	0.17	0.12	0.14	0.18	0.18	0.14	0.15
Top 20−−40%	0.22	0.23	0.17	0.17	0.17	0.18	0.13	0.14	0.20	0.22	0.16	0.16
Top 40−−60%	0.22	0.20	0.18	0.17	0.18	0.17	0.14	0.14	0.20	0.19	0.17	0.17
Top 60−−80%	0.19	0.18	0.19	0.16	0.17	0.16	0.15	0.13	0.17	0.16	0.18	0.15
Bottom 20%	0.23	0.23	0.22	0.20	0.20	0.20	0.18	0.16	0.21	0.21	0.21	0.18
By employment status												
Employed	0.20	0.20	0.17	0.17	0.17	0.17	0.14	0.13	0.19	0.18	0.17	0.16
Unemployed	0.33	0.34	0.29	0.29	0.27	0.27	0.22	0.23	0.30	0.31	0.27	0.24
Estimated parameters‡												
$\grave{\beta}$	0.9866	0.9869	0.9894	0.9897	0.9874	0.9877	0.9898	0.9901	0.9869	0.9873	0.9895	0.9877
∇	0.0088	0.0083	0.0048	0.0043	0.0077	0.0073	0.0040	0.0037	0.0048	0.0046	0.0027	0.0043
Lorenz distance⋆	1.46	1.34	2.93	2.59	0.88	1.19	1.36	1.06	2.06	1.79	3.47	0.67

Note: Annual MPC is calculated by $1 - (1 - \text{quarterly MPC})^4$. The specification of "Borrowing" assumes alternatively: (i) No—$a_t \geq 0$, (ii) Yes—$a_t \geq -2$. † U denotes a uniform distribution of β, and L denotes a (truncated) log-normal distribution. ‡: Discount factors are uniformly distributed over the interval $[\grave{\beta} - \nabla, \grave{\beta} + \nabla]$. For the log-normal-$\beta$ specification, the "center" $\grave{\beta}$ is the mean of the distribution, while the "spread" ∇ is the log standard deviation. ⋆ The Lorenz distance is defined as: $\min_{\{\beta, \nabla\}} (\sum_{i = 20, 40, 60, 80} (w_i(\beta, \nabla) - \omega_i)^2)^{1/2}$, where w_i and ω_i are the proportions of total aggregate net worth held by the top i percent in the model and in the data, respectively.

where ψ denotes the white noise permanent shock to income. The transitory component is

$$\xi_t = \mu \quad \text{with probability } \mho_t \tag{S10}$$

$$= (1 - \tau_t)\ell\theta_t \quad \text{with probability } 1 - \mho_t, \tag{S11}$$

where $\mu > 0$ is the unemployment insurance payment when unemployed, τ_t is the rate of tax collected to pay unemployment benefits, ℓ is time worked per employee, and θ_t is white noise.

S2.2 *Alternative specifications: Table 1*

Table S3 considers the implications of switching off the permanent component ψ, the transitory component θ, and the borrowing constraint:

$$a_t \geq 0.$$

Specifications 4 and 8 turn off both permanent and transitory income shocks, but keep the unemployment shock $\mho_t$ on; this replicates the idiosyncratic income process used in the KS-hetero model (other than the dependence of $\mho_t$ on the macroeconomic state).

- Specification 1 is the baseline FBS β-dist model reproduced from the seventh column of Table 3 in the paper, with the aggregate MPC of 0.21.

- Specification 2 confirms the referee's point that the MPC has very little sensitivity to shutting down the transitory shock θ.

- Under specification 3, when the permanent shock is turned off, the aggregate MPC declines to 0.18. The fit of the model worsens, as the Lorenz distance roughly doubles (compared to the baseline).
  When transitory and, in particular, permanent shocks are turned off, the width of the band for the discount rate ∇ is smaller. When risk is higher (particularly permanent risk), households want to hold more wealth, so we need to stretch to lower discount factors to get them to hold very little wealth and match the data. Extra risk has little effect on the wealth holdings of the patient, so the top end of the band for β does not move.

- Specification 4 turns off both permanent (ψ) and transitory (θ) shocks. The implications are similar to those of specification 3.

- Specification 5 assumes the baseline FBS income process but allows for borrowing by easing the constraint $a_t \geq 0$ to $a_t \geq -2$. Assets are thus allowed to be negative as long as they do not decline below -2 times quarterly permanent income, which is the same extent as in the KS-hetero model. This results in a decline of aggregate MPC from 0.21 to 0.18. The width ∇ reduces somewhat and the fit of the model improves.

- Specifications 6, 7, and 8 investigate *the interactions between borrowing and in-come risk*. If households can borrow, they need to be a bit more patient on the whole to fit the aggregate capital–income target, so they are operating a bit higher on their consumption function, where MPC is lower (specification 5). In the presence of permanent income risk, a good fraction of them are still on the highly curved portion of the consumption function (which is larger when there is lots of risk), so the MPC is only somewhat lowered. Conversely, when borrowing is not allowed and there is no permanent income risk (specifications 3 and 4), the consumption function flattens out quickly, but matching the distribution of wealth means some agents are operating very close to the constraint, so even transitory risk induces precautionary behavior and the MPC again stays relatively high. When *borrowing is allowed and there is no permanent risk* (specifications 7 and 8), the consumption function becomes flat very quickly and agents are operating above the curved portion, resulting in an appreciably lower MPC.

  The MPC in specification 6 (borrowing and no transitory risk) is essentially the same as in specification 5 (borrowing and transitory risk).

Specification 8 is closest to the KS-hetero model both in its structure (simple idiosyncratic income process and with borrowing) and its estimate of the aggregate MPC (0.14 vs. 0.09). The remaining differences between specification 8 and KS-hetero are the specification of aggregate shocks and the nature of β heterogeneity. As Table 3 in our paper shows that the MPC is similar under the FBS or KS aggregate process (0.21 vs. 0.23), we believe the residual difference in MPCs is mostly or entirely accounted for by the vastly different assumptions about the distribution of β. Note that the lowest, middle, and highest discrete values of β in specification 8 (0.9869, 0.9901, 0.9933) are *very* close to the three KS β values (0.9858, 0.9894, 0.9930). However, in specification 8 about 29% (or 2/7) of simulated agents have an intermediate β between the lowest and the central types; the corresponding percentiles of β in the KS-hetero model (as well as the 4% below that) all have $\beta = 0.9894$. As the average MPC is convex in β, this relative dispersion of the central mass results in an increased MPC in specification 8 relative to KS-hetero.

To further explore how the structure of β heterogeneity affects the aggregate MPC, specifications 9–12 assume that β has a *truncated log-normal* distribution (rather than uniform).[18] Qualitatively, having a log-normal distribution of β generates roughly the same estimated results as a uniform distribution, but with a slightly lower MPC. The log-normal distribution is more centrally concentrated than the uniform distribution, but not as concentrated as the 10–80–10 assumption in KS-hetero; thus the average MPC is lower than in the uniform specifications but still higher than KS-hetero. For specifications 9–11, the log-normal model has a somewhat lower fit in terms of the Lorenz distance than the corresponding uniform models (specifications 1–3), while specification 12 (which allows for borrowing) implies an even better fit than the respective uniform specification (5).

---

[18]The distribution is approximated with seven types of households with different βs. The discretization of the log-normal distribution was done in the same manner as for the income shock process.

To summarize, turning off the permanent shocks and allowing for borrowing reduces the aggregate MPC from 0.21 to 0.18 each. Both features together reduce the aggregate MPC from 0.21 to 0.14 (compared with the KS-hetero value of 0.09). Likewise, assuming log-normal β also reduces the aggregate MPC by about 0.02. In contrast, turning off transitory shocks does not noticeably affect the MPC.

References

Brown, J., J. B. Liebman, and J. Pollett (2002), "Estimating life tables that reflect socioeconomic differences in mortality." In *The Distributional Aspects of Social Security and Social Security Reform* (M. Feldstein and J. B. Liebman, eds.), 447–457, University of Chicago Press. [5]

Cagetti, M. (2003), "Wealth accumulation over the life cycle and precautionary savings." *Journal of Business and Economic Statistics*, 21 (3), 339–353. [5, 11]

Carroll, C. D. (2000), "Why do the rich save so much?" In *Does Atlas Shrug? The Economic Consequences of Taxing the Rich* (J. B. Slemrod, ed.). Harvard University Press. [11]

Carroll, C. D. (2006), "The method of endogenous gridpoints for solving dynamic stochastic optimization problems." *Economics Letters*, 91, 312–320. [3]

Kaplan, G., G. Violante, and J. Weidner (2014), "The wealthy hand-to-mouth." *Brookings Papers on Economic Activity*, Spring, 77–138. [8]

Sabelhaus, J. and J. Song (2010), "The great moderation in micro labor earnings." *Journal of Monetary Economics*, 57 (4), 391–403. [5]

Social Security Administration (2010), "Actuarial life table." Available at http://www.ssa.gov/oact/STATS/table4c6.html. [5]