Digestible Microfoundations: Buffer Stock Saving in a Krusell-Smith World

Christopher D. Carroll Kiichi Tokuoka

October 9, 2010

Abstract

Krusell and Smith (1998) pioneered a technique that permits construction of macroeconomic models with serious microfoundations. We argue that three modifications to their model are required to fulfill the technique's promise. First, we replace their assumption about household income dynamics with a process that matches microeconomic data. Second, our agents have finite lifetimes *a la* Blanchard (1985). Finally, we calibrate heterogeneity in time preference rates so that the model matches the observed degree of inequality in the wealth distribution. Our model has substantially different, and considerably more plausible, implications for macroeconomic questions like aggregate marginal propensity to consume out of an economic 'stimulus' program.

Keywords: Income Shocks, Heterogeneity in Time Preference Rates, Distribution of Wealth, Marginal Propensity to Consume

JEL Classification Codes: D12, D31, D91, E21

Carroll: Department of Economics, Johns Hopkins University, Baltimore, MD, http://econ.jhu.edu/people/ccarroll/, ccarroll@jhu.edu;
Tokuoka: Department of Economics, Johns Hopkins University, ktokuoka@gate01.com

This paper was written as a complement to the papers included in the special issue of the *Journal of Economic Dynamics and Control* (2010, Volume 3, Issue 1) on solution methods for representative agent models, edited by Wouter den Haan, Ken Judd, and Michel Julliard.

1 Introduction

Macroeconomists have sought credible microfoundations since the inception of our discipline. Keynes, his critics, and subsequent generations through Lucas (1976) and beyond have agreed on this, if little else.

Since Keynes's time, consumption modeling has been a battleground beween two microfoundational camps. "Bottom up" modelers (e.g. Modigliani and Brumberg (1954); Friedman (1957)) drew wisdom from microeconomic data and argued that macro models shoudl be constructed by aggregating models that matched robust micro facts. "Top down" modelers (e.g. Samuelson (1958); Diamond (1965); Hall (1978)), on the other hand, treated aggregate consumption as reflecting the optimizing decisions of representative agents; with only one such agent (or, at most, one per generation), these models had "microfoundations" under a generous definition of the term.

The tractability of representative agent models has made them appealing to business cycle modelers. But such models have never been easy to reconcile with either macroeconomic[1] or microeconomic[2] empirical evidence, nor with microeconomic theory which implies that heterogeneity (in age, preferences, wealth, liquidity constraints, taxes, and other dimensions) means that different people should respond differently to any given shock (a proposition supported by empirical evidence too extensive to cite). If any of these differences matter (and it is hard to see how they could *fail* to matter),[3] the aggregate size of a shock is not a sufficient statistic to calculate the aggregate response; information about how the shock is distributed across households is also required.

The bottom-up approach, however, has also has had serious drawbacks. Even judged by a sympathetic standard that asks whether such models can match measured heterogeneity, the bottom-up approach has not been fully successful. For example, bottom-up models calibrated to match the wealth holdings of the median household generally fail to match the large size of the aggregate capital stock, because they seriously underpredict wealth in the upper parts of the wealth distribution. Alternatively, models calibrated to match the aggregate level of wealth greatly overpredict wealth at the median (Hubbard, Skinner, and Zeldes (1994); Carroll (2000b)). A further problem is that (at least until Krusell and Smith (1998)), there has been no common answer to the question of how to analyze

[1]See, e.g., Campbell and Mankiw (1991) and the vast related literature.

[2]A large microeconomic literature, for example, has found average values of the marginal propensity to consume much greater than the 3-5 percent implied by representative agent models; see, e.g., Parker (1999) or Souleles (1999).

[3]See Solow (2003) for an eloquent statement of the deficiencies of representative agent models.

systematic macroeconomic fluctuations (business cycles) in bottom-up models.

Our ambition in this paper is to reconcile the two camps. Specifically, we argue that a satisfactory model can be constructed by making three modifications to the well-known Krusell-Smith ('KS') framework. First, we replace KS's highly stylized assumptions about the nature of idiosyncratic income shocks with a microeconomic labor income process that captures the essentials of the empirical consensus from the labor economics literature (with microeconomically credible transitory and permanent shocks).[4] Second, agents in our model have finite lifetimes *a la* Blanchard (1985), permitting a kind of primitive life cycle analysis. Finally, we obtain a necessary extra boost to wealth inequality by calibrating a simple measure of heterogeneity in 'impatience.'[5]

The resulting model differs sharply from the KS model in its implications for important microeconomic and macroeconomic questions. Given recent fiscal policy debates, a timely example of such a macroeconomic question is how aggregate consumption will respond to a temporary tax cut. In response to a $1-per-capita lump sum transfer, the benchmark KS model implies that the annual marginal propensity to consume (MPC) is about 0.05,[6] almost irrespective of how the tax cut is distributed across households. In contrast, the preferred version of our model implies that if the entire tax cut is directed at households in the bottom half in the wealth-to-income distribution, the MPC will be about 0.22. Since empirical evidence does seem to confirm the prediction from theory (Carroll and Kimball (1996)) that MPC's are higher for lower-wealth households,[7] and since an aggregate shock of any given size might be distributed across the population in a wide variety of different ways, this is an improvement in realism that may matter for important questions of macroeconomic dynamics as well as public policy.

Section 2 of the paper begins building the structure of the model by starting with a perfect foresight representative agent model and then adding the microeconomic modeling elements. Using this model (without macroeconomic dynamics), the section closes by estimating the degree of heterogeneity in impatience necessary to match the U.S. wealth distribution; we find that relatively small differences

[4]See, e.g., Hryshko (2010) for a recent overview of the empirical literature.

[5]The word is in quotes because we refer here not to the pure time preference rate but instead to a relation between time preference rate, the interest rate, relative risk aversion, the magnitude of risk, and expected income growth. All of these parameters surely vary in the population, but our view is that few if any important macroeconomic questions depend on which particular kinds of heterogeneity are most responsible for the heterogeneity in wealth/income ratios. See Subsection 2.4 for a fuller discussion.

[6]That is, if a dollar were given to every household in the economy, over the subsequent year household spending would be higher by about $0.05.

[7]See, e.g., Souleles (1999) and references therein and thereto.

in impatience make a large difference in the fit of the model to the wealth data. Section 3 builds up the full version of the model by adding aggregate shocks of the KS type to the model, and presents what we view as the key comparisons of our model with the KS model. Section 4 improves the model by introducing an aggregate income process that is analytically simpler than the KS process, that we believe is more empirically plausible as well, and that simplifies model solution and simulation considerably. We offer this final, simpler version of the model as our preferred jumping-off point for future macroeconomic research.

2 The Model without Aggregate Uncertainty

2.1 The Perfect Foresight Representative Agent Model

To establish notation and a transparent benchmark for comparison purposes, we begin by briefly setting out a standard perfect foresight representative agent model.

The aggregate production function is

$$\Psi_t \boldsymbol{K}_t^{\alpha} (\bar{l} \boldsymbol{L}_t)^{1-\alpha},$$

where Ψ_t is aggregate productivity in period t, $\boldsymbol{K}_t$ is capital, $\bar{l}$ is time worked per capita, and $\boldsymbol{L}_t$ is employment. The representative agent's goal is to maximize discounted utility from consumption

$$\max \sum_{n=0}^{\infty} \beta^n \mathrm{u}(\boldsymbol{C}_{t+n})$$

for a CRRA utility function $\mathrm{u}(\bullet) = \bullet^{1-\rho}/(1-\rho)$.[8] The representative agent's state at the time of the consumption decision is defined by two variables: $\boldsymbol{M}_t$ is market resources, and Ψ_t is aggregate productivity.

The transition process for $\boldsymbol{M}_t$ is broken up, for clarity of analysis and consistency with later notation, into three steps. Assets at the end of the period are market resources minus consumption, equal to

$$\boldsymbol{A}_t = \boldsymbol{M}_t - \boldsymbol{C}_t,$$

while next period's capital is determined from this period's assets via

$$\boldsymbol{K}_{t+1} = \boldsymbol{A}_t.$$

The final step can be thought of as the transition from the beginning of period $t+1$ when capital has not yet been used to produce output, to the middle of that

[8]Substitute $\mathrm{u}(\bullet) = \log \bullet$ for the case where $\rho = 1$.

period, when output has been produced and incorporated into resources but has not yet been consumed:

$$\begin{aligned}\boldsymbol{M}_{t+1} &= \daleth\boldsymbol{K}_{t+1} + \boldsymbol{K}_{t+1}\mathsf{r}_{t+1} + (\bar{l}\boldsymbol{L}_{t+1})\mathsf{W}_{t+1} \\ &= \daleth\boldsymbol{K}_{t+1} + \Psi_{t+1}\boldsymbol{K}_{t+1}^{\alpha}(\bar{l}\boldsymbol{L}_{t+1})^{1-\alpha},\end{aligned}$$

where r_{t+1} is the interest rate,[9] W_{t+1} is the wage rate,[10] and $\daleth = (1-\delta)$ is the depreciation factor for capital.

After normalizing by the productivity factor $\mathrm{Z}_t = \Psi_t^{1/(1-\alpha)}(\bar{l}\boldsymbol{L}_t)$,[11] the representative agent's problem is

$$\begin{aligned}\mathrm{V}(M_t, \Psi_t) &= \max_{\{C_t\}} \ \mathrm{u}(C_t) + \beta\mathbb{E}_t\left[\Gamma_{t+1}^{1-\rho}\mathrm{V}(M_{t+1}, \Psi_{t+1})\right] &(1)\\ &\text{s.t.} \\ A_t &= M_t - C_t &(2)\\ K_{t+1} &= A_t/\Gamma_{t+1} &(3)\\ M_{t+1} &= \daleth K_{t+1} + K_{t+1}^{\alpha}, &(4)\end{aligned}$$

where the non-bold variables are the corresponding bold variables divided by Z_t (e.g., $A_t = \boldsymbol{A}_t/\mathrm{Z}_t$, $M_t = \boldsymbol{M}_t/\mathrm{Z}_t$); $\Gamma_{t+1} = \mathrm{Z}_{t+1}/\mathrm{Z}_t$; and the expectations operator $\mathbb{E}_t$ here signifies the perfection of the agent's foresight (but will have the usual interpretation when uncertainty is introduced below).

Except where otherwise noted, our parametric assumptions match those of the papers in the special issue of the *Journal of Economic Dynamics and Control* (2010, Volume 34, Issue 1, edited by den Haan, Judd, and Julliard) devoted to comparing solution methods for the KS model (the parameters are reproduced for convenience in the top panel of Table 1).[12] The model is calibrated at the quarterly frequency. When aggregate shocks are shut down ($\Psi_t = 1$ and $\boldsymbol{L}_t = \boldsymbol{L}$), the model has a steady-state solution with a constant ratio of capital to output and constant interest and wage factors, which we write without time subscript as r and W and which are reflected in the lower panel of Table 1.[13]

2.2 The Household Income Process

For our purposes, the principal conclusion of the large literature on microeconomic labor income dynamics is that household income can be reasonably well described

[9] Equal to the marginal product of capital, $\alpha\Psi_{t+1}\boldsymbol{K}_{t+1}^{\alpha-1}(\bar{l}\boldsymbol{L}_{t+1})^{1-\alpha}$.

[10] Equal to the marginal product of labor, $(1-\alpha)\Psi_{t+1}\boldsymbol{K}_{t+1}^{\alpha}(\bar{l}\boldsymbol{L}_{t+1})^{-\alpha}$.

[11] Details of this normalization are discussed in Carroll (2000a).

[12] Examples of such authors include Young (2007) and Algan, Allais, and Haan (2008).

[13] In the steady state, $\boldsymbol{K}_t/(\bar{l}\boldsymbol{L}_t) = \bar{k} = (\alpha\beta/(1-\beta\daleth))^{1/(1-\alpha)} = 38.0$, $\mathsf{r} = \alpha\bar{k}^{\alpha-1}$, and $\mathsf{W} = (1-\alpha)\bar{k}^{\alpha}$.

Table 1: Aggregate Parameter Values

Calibration		
Param	Value	Source
β	0.99	Krusell and Smith (1998)
ρ	1	Krusell and Smith (1998)
α	0.36	Krusell and Smith (1998)
δ	0.025	Krusell and Smith (1998)
$\bar{l}$	1/0.9	Den Haan, Judd, and Julliard (2007)
Steady State		
$\boldsymbol{K}/\boldsymbol{Y}$	10.3	
r	0.035	
W	2.37	

as follows. The idiosyncratic permanent component of labor income p evolves according to

$$p_{t+1} = \mathsf{G}_{t+1} p_t \psi_{t+1}$$

where G_{t+1} captures the predictable low-frequency (e.g., life-cycle and demographic) components of income growth, and the Greek letter psi mnemonically indicates the *p*ermanent *s*hock to *i*ncome. Actual income is equal to a product of the permanent component of income, a mean-one transitory shock, and the wage rate:

$$\boldsymbol{y}_{t+1} = p_{t+1} \xi_{t+1} \mathsf{W}_{t+1}.$$

Table 2 summarizes the annual variances of log permanent shocks (σ_ψ^2) and log transitory shocks (σ_ξ^2) estimated by a selection of papers from the extensive literature.[14] Some authors have used a process of this kind to describe the labor income process for an individual worker (top panel)[15] while others have used it to

[14]All the authors cited above used U.S. data. Nielsen and Vissing-Jorgensen (2006) used Danish data and estimated $\sigma_\psi^2 = 0.005$ and $\sigma_\xi^2 = 0.015$. It would be reasonable to interpret their estimates as the lower bounds for the U.S., given that their administrative data is well-measured and but that Danish welfare is more generous than the U.S. system.

[15]MaCurdy (1982) did not explicitly separate ψ_t and ξ_t, but we have extracted σ_ψ^2 and σ_ξ^2 as implications of statistics that his paper reports. First, we calculate $\text{var}(\log \boldsymbol{y}_{t+d} - \log \boldsymbol{y}_{t+d-1})$ and $\text{var}(\log \boldsymbol{y}_{t+d-1} - \log \boldsymbol{y}_{t+d-2})$ using his estimate (we set $d = 5$). Then, following Carroll and Samwick (1997) we obtain the values of σ_ψ^2 and σ_ξ^2 which can match these statistics, assuming that the income process is $\boldsymbol{y}_t = p_t \xi_t$ and $p_t = p_{t-1}\psi_t$ (i.e., we solve $\text{var}(\log \boldsymbol{y}_{t+d} - \log \boldsymbol{y}_{t+d-1}) = d\sigma_\psi^2 + 2\sigma_\xi^2$ and $\text{var}(\log \boldsymbol{y}_{t+d-1} - \log \boldsymbol{y}_{t+d-2}) = (d-1)\sigma_\psi^2 + 2\sigma_\xi^2$).

describe the process for overall household income.[16]

The second-to-last line of the table shows what labor economists would have found, when estimating a process like the one above, if the empirical data were generated by households who experienced an income process like the one assumed by Krusell and Smith (1998).[17] This row of the table makes our point forcefully: The empirical procedures that have been applied to empirical micro data, if used to measure the income process households experience in a KS economy, would have produced estimates of σ_ψ^2 and σ_ξ^2 that are orders of magnitude different from what the actual empirical literature finds in actual data. This discrepancy naturally prompts the question (answered below) of whether the baseline KS model's well-known difficulty in matching the degree of wealth inequality is largely explained by their unrealistic assumption about the income process.[18]

2.3 Finite Lifetimes and the Finite Variance of Permanent Income

One might wish to use the permanent/transitory income process specified in Subsection 2.2 as a complete characterization of household income dynamics, but that idea has a problem: Since each household accumulates a permanent shock in every period, the cross-sectional distribution of idiosyncratic permanent income becomes wider and wider indefinitely as the simulation progresses; that is, there is no ergodic distribution of permanent income in the population.

This problem and several others can be addressed by assuming that the model's agents have finite lifetimes *a la* Blanchard (1985). Death follows a Poisson process, so that every agent who is part of the population at date t has an equal probability

[16] Meghir and Pistaferri (2004), Jensen and Shore (2008), Hryshko (2009), and Blundell, Pistaferri, and Preston (2008) assume that the transitory component is serially correlated (an MA process), and report the variance of a subelement of the transitory component. σ_ξ^2 for these articles reported in Table 2 are calculated using their estimates.

[17] First, we generated income draws according to the income process in the baseline KS model, using parameter values in Den Haan, Judd, and Julliard (2007). Then, following the method in Carroll and Samwick (1997), we estimated the variances under the assumption that these income draws were produced by the process $\boldsymbol{y}_t = p_t \xi_t$ where $p_t = p_{t-1}\psi_t$. In doing so, as in Carroll and Samwick (1997), the draws of $\boldsymbol{y}_t$ are excluded when $\boldsymbol{y}_t$ is very low relative to its mean (see Carroll and Samwick (1997) for details about this restriction).

[18] The final line reports the variances estimated using income draws generated by the process assumed in Castaneda, Diaz-Gimenez, and Rios-Rull (2003), who were able to reproduce the skewness of the U.S. wealth distribution by reverse-engineering the income-process assumptions required to allow a Markov income process to generate the observed degree of wealth inequality. This process, too, bears little resemblance to the observable micro data on income dynamics.

Table 2: Estimates of Annual Variances of Log Income Shocks

	σ_ψ^2	σ_ξ^2
Using individual data		
MaCurdy (1982)	0.013	0.045
Topel (1990)	0.013	0.017
Topel and Ward (1992)	0.017	0.013
Meghir and Pistaferri (2004)	0.031	0.032
Low, Meghir, and Pistaferri (2005)	0.011	–
Jensen and Shore (2008)	0.054	0.171
Hryshko (2009)	0.038	0.118
Guvenen (2009)	0.015	0.061
Using household data		
Carroll (1992)	0.016	0.027
Carroll and Samwick (1997)	0.022	0.044
Storesletten, Telmer, and Yaron (2004a)	0.017	0.063
Storesletten, Telmer, and Yaron (2004b)	0.008 – 0.026	0.316
Blundell, Pistaferri, and Preston (2008)	0.010 – 0.030	0.029 – 0.055
Implied by Krusell and Smith (1998)	0.000	0.039
Implied by Castaneda et al. (2003)	0.030	0.005

D of dying before the beginning of period $t+1$. Agents are assumed to engage in a Blanchardian mutual insurance scheme: Those who survive receive a proportion of the estates of those who die. Since we assume that there is no bequest motive, the entire estates of the dying households are available for apportionment among the survivors, which means (assuming a zero profit condition for the insurance industry) that the insurance scheme boosts the rate of return (for survivors) by an amount exactly corresponding to the mortality rate.

In order to maintain a steady population, we assume that dying households are replaced by an equal number of newborns. Newborns, however, begin life with a level of idiosyncratic permanent income equal to the mean level of idiosyncratic permanent income $\bar{p}$ in the population as a whole. Conveniently, our definition of the permanent shock implies that in a large population, mean idiosyncratic permanent income will remain fixed at $\bar{p} = 1$ forever.

Making the usual assumption that the population is uniformly distributed on the unit interval with total mass of 1, the mean of p^2 is given by

$$\mathbb{M}[p^2] = \left(\frac{\mathsf{D}}{1 - \not{\mathsf{D}}\mathbb{E}[\psi^2]} \right), \tag{5}$$

where $\mathbb{M}[\bullet_t] = \int_0^1 \bullet_{t,\iota} d\iota$ is the mean operator (see Appendix A for the derivation).

The relation between p^2 and the variance of p is

$$\begin{aligned} \sigma_p^2 &= \mathbb{M}[(p - \mathbb{M}[p])^2] \\ &= \mathbb{M}[(p^2 - 2p\mathbb{M}[p] + (\mathbb{M}[p])^2)] \\ &= \mathbb{M}[p^2] - 1 \end{aligned} \tag{6}$$

where the last line follows because under the other assumptions we have made, $\mathbb{M}[p] = 1$.

Of course for the preceding derivations to be valid, it is necessary to impose the parameter restriction $\not{\mathsf{D}}\mathbb{E}[\psi^2] < 1$ (a requirement that does not do violence to the data, as we shall see). Intuitively, the requirement is that, among surviving consumers, income does not spread out so quickly as to overcome the compression of the permanent income distribution that arises because of the equalizing force of death and replacement.

Since our goal here is to produce a realistic distribution of permanent income across the members of the (simulated) population, we measure the empirical distribution of permanent income in the cross section using data from the Survey of Consumer Finances (SCF), which conveniently includes a question asking respondents whether their income in the survey year was about 'normal' for them, and if not, asks the level of 'normal' income.[19] This corresponds well with our defini-

[19] SCF1992 only asked whether the income level was about 'normal' or not.

Table 3: Variance of Permanent Income

	$\text{var}(\boldsymbol{p}/\mathbb{M}[\boldsymbol{p}])$	$\mathbb{E}[\psi^2]$	$\text{var}(\psi)$
SCF1992	2.5	1.015	0.015
SCF1995	7.5	1.018	0.018
SCF1998	3.1	1.015	0.015
SCF2001	3.6	1.016	0.016
SCF2004	5.2	1.017	0.017
Krusell-Smith	0	1	0

tion of permanent income (level) $\boldsymbol{p}$ (and Kennickell (1995) shows that the answers people give to this question can be reasonably interpreted as reflecting their perceptions of their permanent income), so we calculate the variance of $\boldsymbol{p}_i/\mathbb{M}[\boldsymbol{p}_i]$ among such households.[20]

The results from this exercise are reported in Table 3. Substituting these estimates for σ_p^2 into (5) and (6), we obtain estimates of the variance of $\psi(= \mathbb{E}[\psi^2]-1)$. Reassuringly, we can interpret the variances of ψ thus obtained as being easily in the range of the estimated variances of $\log(\psi) = \sigma_\psi^2$ in Table 2.[21] Such a correspondence, across two quite different methods of measurment, suggests there is considerable robustness to the measurement of the size of permanent shocks. (Below, we will choose a calibration for σ_ψ^2 that is in the middle range of estimates from either method.)

2.4 Heterogeneous Impatience

We now examine how wealth would be distributed in the steady-state equilibrium of an economy with wage rates and interest rates fixed at the steady state values calibrated in Table 1 of Subsection 2.1, an income process like the one described in Subsection 2.2, and finite lifetimes per Subsection 2.3.

The process of noncapital income of each household follows

$$\boldsymbol{y}_t = p_t \xi_t \mathsf{W}_t \tag{7}$$

$$p_t = p_{t-1}\psi_t \tag{8}$$

$$\mathsf{W}_t = (1-\alpha)\Psi_t(\boldsymbol{K}_t/\bar{l}\boldsymbol{L}_t)^\alpha, \tag{9}$$

where $\boldsymbol{y}_t$ is noncapital income for the household in period t, equal to the permanent

[20] We restrict the sample to households between the ages of 25 and 60, because the interpretation of the question becomes problematic for retired households.

[21] So long as the variance of the permanent shocks is small, these two measures should be approximately the same.

component of noncapital income p_t multiplied by a mean-one iid transitory income shock factor ξ_t (from the perspective of period t, all future transitory shocks are assumed to satisfy $\mathbb{E}_t[\xi_{t+n}] = 1$ for all $n \geq 1$) and wage rate W_t; the permanent component of noncapital income in period t is equal to its previous value, multiplied by a mean-one iid shock ψ_t, $\mathbb{E}_t[\psi_{t+n}] = 1$ for all $n \geq 1$. Lastly, $\boldsymbol{K}_t$ is *per capita* capital and $\boldsymbol{L}_t = 1 - u_t$ is the employment rate, where u_t is the unemployment rate. Since there is no aggregate shock, Ψ_t, $\boldsymbol{K}_t$, $\boldsymbol{L}_t$, and W_t are constant ($\Psi_t = \Psi = 1$, $\boldsymbol{K}_t = \boldsymbol{K}$, $\boldsymbol{L}_t = \boldsymbol{L}$, and $\mathsf{W}_t = \mathsf{W} = (1-\alpha)(\boldsymbol{K}/\bar{l}\boldsymbol{L})^{\alpha}$).

The distribution of ξ_t is as follows:

$$\begin{aligned} \xi_t &= \mu \text{ with probability } u_t && (10) \\ &= (1-\tau_t)\bar{l}\theta_t \text{ with probability } 1 - u_t, && (11) \end{aligned}$$

where μ is the unemployment insurance payment when unemployed and $\tau_t = \mu u_t / \bar{l}\boldsymbol{L}_t$ is the rate of tax collected to pay the unemployment benefits. The probability of unemployment is constant ($u_t = u$); later we allow it to vary over time. (Again for comparability, these assumptions about the nature of unemployment follow the structure and calibrations in the JEDC volume mentioned above).

The decision problem for the household in period t can be written using normalized variables:

$$\begin{aligned} \mathrm{v}(m_t) &= \max_{\{c_t\}} \ \mathrm{u}(c_t) + \beta\not{D}\mathbb{E}_t\left[\psi_{t+1}^{1-\rho}\mathrm{v}(m_{t+1})\right] && (12) \\ &\text{s.t.} \\ a_t &= m_t - c_t \\ a_t &\geq 0 \\ k_{t+1} &= a_t/(\not{D}\psi_{t+1}) && (13) \\ m_{t+1} &= (\daleth + \mathsf{r})k_{t+1} + \xi_{t+1} && (14) \end{aligned}$$

where the non-bold ratio variables are defined as the bold (level) variables divided by the level of permanent income $\boldsymbol{p}_t = p_t\mathsf{W}$ (e.g., $m_t = \boldsymbol{m}_t/(p_t\mathsf{W})$). The only state variable is (normalized) cash on hand m_t. Note that the household's employment status is not a state variable, unlike in the KS model, where tomorrow's employment status depends on today's status. This constitutes a substantial improvement in simplicity (which is useful for computational and analytical purposes), arguably without too much sacrifice of realism (except possibly for detailed studies of the behavior of households during extended unemployment spells).

Since households die with a constant probability D between periods, the effective discount factor is $\beta\not{D}$ (in (12)). Note that the effective interest rate is $(\daleth+\mathsf{r})/\not{D}$

(in (13) and (14)).[22]

We choose standard parameter values (Table 4). $\mu = 0.15$ is from Den Haan, Judd, and Julliard (2007).[23] $\mathsf{D} = 0.005$ implies the average length of working life is $1/0.005 = 200$ quarters $= 50$ years (dating from entry into the labor force at, say, age 25). The variance of log transitory income shocks σ_θ^2 is from Carroll (1992),[24] and our calibration of $\sigma_\psi^2 = 0.016$ is from the same source (which conveniently also matches the median value in Table 3).[25,26] Other parameter values (ρ, α, δ, and $\bar{l}$) are from Table 1.

The one remaining unspecified parameter is the time preference factor. As a preliminary theoretical consideration, note that Carroll (2009) (generalizing Deaton (1991)) has shown that models of this kind do not have a well-defined solution unless the condition

$$\left(\frac{(\mathsf{R}\beta)^{1/\rho}}{\acute{\Gamma}}\right) < 1 \tag{15}$$

holds where

$$\acute{\Gamma} = \left(\mathbb{E}[\psi^{-\rho}]\right)^{-1/\rho}\Gamma.$$

Carroll (2009) dubs this the 'Growth Impatience Condition' because it is the condition required to guarantee that consumers are sufficiently impatient to prevent the indefinite increase in the *ratio* of net worth to permanent income when

[22] $(\text{⅂}+\mathsf{r})$ is scaled by $1/\not{\mathsf{D}}$ due to the Blanchardian mutual insurance scheme as described in the previous subsection.

[23] $\mu = 0.15$ is different from the choice in Krusell and Smith (1998), where $\mu = 0$.

[24] This paper assumes that each period corresponds to a *quarter*, while $\sigma_\theta^2 = 0.027$ (from Carroll (1992)) is estimated using annual data. Therefore, following Carroll and Slacalek (2008), 0.027 needs to be multiplied by 4 since the variance of log transitory income shocks of *quarterly* data is *four* times as large as that of annual data. Note further that $\sigma_\theta^2 = 0.027$ is more modest than other estimates such as in Carroll and Samwick (1997) ($= 0.044$). The reason why $\sigma_\theta^2 = 0.027$ is used in this paper is that Carroll and Samwick (1997) themselves argue that their estimate of σ_θ^2 is almost certainly increased by measurement error.

[25] Since σ_ψ^2 in Table 3 (0.016) is estimated using annual data, it needs to be *divided* by 4, following Carroll and Slacalek (2008) (recall that our model is calibrated quarterly).

[26] Using quarterly income draws generated by this section's income process with these parameter values, we have estimated the *annual* ARMA process for $\log(\xi_t)$ assumed in Moffitt and Gottschalk (1995): $\log(\xi_t) = a_1 \log(\xi_{t-1}) + v_t + m_1 v_{t-1}$. The estimates of a_1 and m_1 are positive and negative, respectively, in line with the coefficients estimated by Moffitt and Gottschalk (1995) using the U.S. data (Panel Study of Income Dynamics). This suggests that Moffitt and Gottschalk's findings are consistent with the other papers in this literature, and with our own calibration of the income process.

Table 4: Parameter Values for Heterogenous Agents Model

Description	Param	Value	Source
Unemp Insurance Payment	μ	0.15	Den Haan, Judd, and Julliard (2007)
Unemp Rate	u	0.07	Mean in Krusell and Smith (1998)
Probability of Death	D	0.005	Yields 50 year working life
Variance of Log $\theta_{t,i}$	σ_θ^2	$0.027 \cdot 4$	Carroll (1992)
Variance of Log $\psi_{t,i}$	σ_ψ^2	0.016/4	Carroll (1992); median in Table 3

income is growing (see also Szeidl (2006)). This condition is a complex amalgam of the pure time preference factor, expected growth, the relative risk aversion coefficient, and the real interest factor so that, for example, a consumer can be 'impatient' in the required sense even if $\beta = 1$, so long as expected income growth is positive.

We begin by searching for the time preference factor $\grave{\beta}$ such that if all households had $\beta = \grave{\beta}$ the steady-state value of the capital-to-output ratio ($\boldsymbol{K}_t/\boldsymbol{Y}_t$) would match the value that characterized the steady-state of the perfect foresight model.[27] $\grave{\beta}$ turns out to be 0.9887 (first column of Table 5).

We now ask whether the model with realistically calibrated income and finite lifetimes can reproduce the degree of wealth inequality evident in the micro data. An improvement in the model's ability to match the data is to be expected, since in buffer stock models agents strive to achieve a target *ratio* of wealth to permanent income. By assuming that there is no dispersion in permanent income across households, KS's income process assumption shut down a potentially very important important reason for variation in the level of wealth.

Our model's implied distribution of wealth does indeed improve substantially over the distribution implied by the baseline KS model (third column). However, even our model falls substantially short of generating the empirical degree of wealth dispersion. In particular, the proportion of wealth held by households in the top 1 percent of the distribution is far less in the model than in the data. This failure reflects the fact that, empirically, the distribution of wealth is considerably more unequal than the distribution of permanent income.

As the simplest method to address this defect, we introduce heterogeneity in time preference factors in the population: Each household has an idiosyncratic (but fixed) rate of discounting.[28] However, we do not think of this assumption

[27] Output is the sum of noncapital and capital income.

[28] This differs subtly from KS's experiment with heterogeneity, in which a household's discount factor could change suddenly; they interpreted such a change as reflecting a dynastic transition.

Table 5: Wealth Distributions and Marginal Propensity to Consume

	$\grave{\beta} =$ 0.9887	$\{\beta_{low}, \beta_{high}\} =$ $\{0.9810, 0.9922\}$	Krusell and Smith (1998)	Perfect Foresight Partial Equilibrium	U.S. Data
Top 1%	11.6	25.7	3.0		29.6
Top 10%	39.0	65.3	19.0		66.1
Top 20%	55.2	80.4	35.0		79.5
Top 40%	75.8	92.5			92.9
Top 60%	88.9	97.1			98.7
Bottom 20%	3.0	0.7			−0.4
MPC	0.10	0.19		0.04	
$\boldsymbol{K}_t/\boldsymbol{Y}_t$	10.3	10.3			

Notes: U.S. data is the SCF reported in Castaneda, Diaz-Gimenez, and Rios-Rull (2003).

as only capturing actual variation in pure rates of time preference across people (though there is surely substantial variation of that kind). Instead, we view the incorporation of discount-factor heterogeneity as a shortcut that captures the essential consequences of many other kinds of heterogeneity as well. For example, a robust pattern in most countries is that income grows much faster for young people than for older people. According to (15), young people should therefore tend to act, financially, in a more 'impatient' fashion than older people. In particular, we should expect them to have lower target wealth-to-income ratios. Thus, what we are capturing by allowing heterogeneity in time preference factors is probably also some portion of the difference in behavior that reflects differences in age.[29]

One way of gauging a model's predictions for wealth inequality is to ask how well it is able to match the proportion of total wealth held by the top 20 percent, the top 40 percent, and the top 60 percent of the population. Because these statistics have been targeted by other papers (notably Castaneda, Diaz-Gimenez, and Rios-Rull (2003)), we adopt a goal of matching them.

[29]We could of course model age effects directly, but it is precisely the inclusion of such elements of realism that has made OLG models unpopular; they are simply too unweildy to use for many practical research purposes. And our view is that, for macroeconomic analysis purposes, all that is gained in exchange for this complexity is a widening of the distribtution of wealth-to-income ratios. We achieve the same effect much more parsimoniously by incorporating discount factor heterogeneity.

Instead of all households having the same time preference factor, we assume that, for some Δ, time preference factors are distributed uniformly in the population between $\grave{\beta} - \Delta$ and $\grave{\beta} + \Delta$. Then, using simulations, we search for the value of Δ such that our model's mean squared error in matching the three moments (the proportion of wealth held by the top 20 percent, the top 40 percent and the top 60 percent) is minimized (subject to the constraint that the total aggregate wealth/output ratio in the economy continues to match the steady-state value from the perfect foresight model).[30]

The introduction of heterogeneity sharply improves model's fit to the targeted proportions of wealth holdings (second column of the table). The ability of the model to match the targeted moments does not, of course, constitute a test, except in the loose sense that a model with such strong structure might have been unable to get nearly so close to three target points using only one parameter (Δ). Somewhat more impressive is the model's match to locations in the wealth distribution that were not explicitly targeted; for example, the wealth shares of the top 10 percent and the bottom 20 percent are also included in the table, and the model performs reasonably well in matching them.

But perhaps the question of greatest interest is whether a model that manages to match the distribution of wealth in the population has similar, or different, implications from the baseline KS model for serious macroeconomic questions like the reaction of aggregate consumption to a temporary tax cut.

We pose the question as follows. The economy has been in its steady-state equilibrium leading up to date t. Before the consumption decision is made in that period, the government announces the following 'stimulus' plan: Effective immediately, every household in the economy will receive a 'stimulus check' worth some modest amount $\$x$.

The table shows that the immediate net MPC out of the stimulus checks in the preferred version of the model (with heterogeneous time preference factors) is roughly twice as large as that in the version of our model with a unique time preference factor. The MPC in our model is also much larger than that produced by the perfect foresight partial equilibrium model (0.04) (or our solution of the baseline KS model (0.05) – see below).[31]

[30] In estimating these parameter values, we approximate the uniform distribution by seven points ($\bar{\beta}-3\triangle$, $\bar{\beta}-2\triangle$, $\bar{\beta}-\triangle$, $\bar{\beta}$, $\bar{\beta}+\triangle$, $\bar{\beta}+2\triangle$, $\bar{\beta}+3\triangle$, where $\bar{\beta} = (\beta_{low}+\beta_{high})/2$ and $\triangle = (\beta_{high} - \beta_{low})/7$). Increasing the number of points further does not change the results below much.

[31] The MPC's that we calculate in the table are annual MPC's given by $1 - (1 -$ quarterly MPC$)^4$ (recall again that the models in this paper are calibrated quarterly). Most of the empirical literature that has attempted to estimate MPC's has used annual micro data, and so the usual usage of the term 'the MPC' refers to the amount of extra

Table 6: Parameter Values for Aggregate Shocks

Param	Value
$\triangle^{\Psi}$	0.01
u^g	0.04
u^b	0.10
Agg Transition Probability	1/8

3 Aggregate Shocks

KS assumed that the level of aggregate productivity alternates between $\Psi_t = 1 + \triangle^{\Psi}$ if the aggregate state is good and $\Psi_t = 1 - \triangle^{\Psi}$ if it is bad; similarly, $\boldsymbol{L}_t = 1 - u_t$ where $u_t = u^g$ if the state is good and $u_t = u^b$ if bad. (For convenience, we reproduce their assumed parameter values in Table 6.)

Table 7 reports some characteristics of a representative agent model like the one described in section 2.1 above, modified to incorporate an aggregate shock process of the Krusell-Smith type. Using this representative agent model, the standard deviation of consumption ($\sigma_{\log \boldsymbol{C}}$) is lower than that of output, which is consistent with a stylized fact that consumption is less volatile than output in aggregate data.[32,33]

We next examine a model with our preferred household income process that also incorporates KS aggregate shocks. The decision problem for an individual household in period t can be written using normalized variables and the employment status ι_t:

$$
\begin{aligned}
\mathrm{v}(m_t, \iota_t; \boldsymbol{K}_t, \Psi_t) &= \max_{\{c_t\}} \ \mathrm{u}(c_t) + \beta \not{D} \mathbb{E}_t \left[(\Gamma_{t+1}\psi_{t+1})^{1-\rho} \mathrm{v}(m_{t+1}, \iota_t; \boldsymbol{K}_{t+1}, \Psi_{t+1}) \right] \\
&\text{s.t.} \\
a_t &= m_t - c_t \\
a_t &\geq 0 \\
k_{t+1} &= a_t / (\not{D} \Gamma_{t+1} \psi_{t+1}) \\
m_{t+1} &= (\daleth + \mathsf{r}_{t+1}) k_{t+1} + y_{t+1} \\
\mathsf{r}_{t+1} &= \alpha \Psi_{t+1} (\boldsymbol{K}_{t+1} / \bar{l} \boldsymbol{L}_{t+1})^{\alpha - 1},
\end{aligned}
\qquad (16)
$$

where

spending that would occur over the course of a year in response to a one unit increase in resources.

[32] Output $\boldsymbol{Y}_t$ is the sum of both noncapital and capital income.

[33] Note that these statistics are calculated using raw values (not HP or otherwise filtered).

Table 7: Some Statistics for the Representative Agent Model

	Rep Agent Model
Std. Dev. of Output ($\sigma_{\log \boldsymbol{Y}}$)	0.034
Std. Dev. of Consumption ($\sigma_{\log \boldsymbol{C}}$)	0.017
$\sigma_{\log \boldsymbol{Y}}/\sigma_{\log \boldsymbol{C}}$	0.49
Std. Dev. of Investment ($\sigma_{\log \boldsymbol{I}}$)	0.107
$\sigma_{\log \boldsymbol{I}}/\sigma_{\log \boldsymbol{Y}}$	3.14
$corr(\boldsymbol{Y}_t, \boldsymbol{Y}_{t-1})$	0.80
$corr(\boldsymbol{C}_t, \boldsymbol{Y}_t)$	0.68
$corr(\boldsymbol{C}_t, \boldsymbol{Y}_{t-1})$	0.70
$corr(\boldsymbol{C}_t, \boldsymbol{C}_{t-1})$	0.99
$corr(\boldsymbol{I}_t, \boldsymbol{Y}_t)$	0.94
$corr(\boldsymbol{I}_t, \boldsymbol{Y}_{t-1})$	0.68
$corr(\boldsymbol{I}_t, \boldsymbol{I}_{t-1})$	0.73
$corr(\triangle \log \boldsymbol{C}_t, \triangle \log \boldsymbol{C}_{t-1})$	0.25
$\boldsymbol{K}_t/\boldsymbol{Y}_t$	10.3

- the non-bold *individual* variables (lower-case variables except for ι_t and ψ_t) are defined as the bold (level) variables divided by $\mathrm{Z}_t p_t$ (e.g., $a_t = \boldsymbol{a}_t/\mathrm{Z}_t p_t$, $m_t = \boldsymbol{m}_t/\mathrm{Z}_t p_t$),
- $\Gamma_{t+1} = \mathrm{Z}_{t+1}/\mathrm{Z}_t$,
- $\boldsymbol{L}_t = 1 - u_t$, and
- the income process is the same as in (7)-(11) but the process of employment transition follows KS.[34]

There are more state variables in this version of the model than in the model with no aggregate shock: The aggregate variables Ψ_t and $\boldsymbol{K}_t$, and the household's employment status ι_t whose transition process depends on the aggregate state. Solving the full version of the model above with both idiosyncratic and aggregate shocks is not straightforward; indeed, the basic idea for the solution method is the key insight of KS. See Appendix B for details about our solution method.

[34]The transition probabilities of employment status are (originally) from Krusell and Smith (1998), but following Den Haan, Judd, and Julliard (2007) $\mu = 0.15$ (this is 0 in Krusell and Smith (1998)) as in model with no aggregate shock.

Table 8: Statistics for Heterogeneous Agents Model

	Buffer Stock		Krusell-Smith	
	Baseline	Hetero	Our solution	Maliar et al. (2008)
$\sigma_{\log \boldsymbol{Y}}$	0.034	0.033	0.034	
$\sigma_{\log \boldsymbol{C}}$	0.016	0.017	0.016	
$\sigma_{\log \boldsymbol{C}}/\sigma_{\log \boldsymbol{Y}}$	0.48	0.52	0.47	
$\sigma_{\log \boldsymbol{I}}$	0.104	0.095	0.105	
$\sigma_{\log \boldsymbol{I}}/\sigma_{\log \boldsymbol{Y}}$	3.09	2.85	3.12	
$corr(\boldsymbol{Y}_t, \boldsymbol{Y}_{t-1})$	0.81	0.80	0.81	0.81
$corr(\boldsymbol{C}_t, \boldsymbol{Y}_t)$	0.70	0.80	0.68	0.68
$corr(\boldsymbol{C}_t, \boldsymbol{Y}_{t-1})$	0.72	0.79	0.70	0.71
$corr(\boldsymbol{C}_t, \boldsymbol{C}_{t-1})$	0.98	0.96	0.99	0.99
$corr(\boldsymbol{I}_t, \boldsymbol{Y}_t)$	0.95	0.95	0.95	0.94
$corr(\boldsymbol{I}_t, \boldsymbol{Y}_{t-1})$	0.69	0.68	0.70	0.68
$corr(\boldsymbol{I}_t, \boldsymbol{I}_{t-1})$	0.75	0.74	0.75	0.74
$corr(\triangle \log \boldsymbol{C}_t, \triangle \log \boldsymbol{C}_{t-1})$	0.20	0.08	0.26	0.28
$\boldsymbol{K}_t/\boldsymbol{Y}_t$	10.3	10.3	10.3	

3.1 Comparisons of the Models

We now report the results of simulations, using both the model with a unique time preference factor and with heterogeneous preference factors (estimated in Section 2). Henceforth, we call the former the "baseline" model and the latter the "hetero" model. Both models are solved with aggregate shocks introduced above. Results using our solution of the original KS model (where $\theta_t = 1$ and $\psi_t = 1$ for all t but no death ($\mathsf{D} = 0$)) are also reported for comparison.

3.1.1 Business Cycle Statistics

Aggregate business cycle statistics of the three models (baseline model, hetero model, and our solution of the baseline KS model) are generally similar, and replicate the stylized fact that aggregate consumption is less volatile than aggregate income (Table 8). For comparison, we also report the results in Maliar, Maliar, and Valli (2008) for the KS model. Their results are, as expected, very close to those produced by our solution of the KS model reported in the third column.[35]

With respect to aggregate consumption dynamics, our solution of the KS model

[35]The minor differences between the results in Maliar, Maliar, and Valli (2008) and ours reflect approximation error in solving the consumption function.

gives a relatively high correlation coefficient $corr(\triangle \log \boldsymbol{C}_t, \triangle \log \boldsymbol{C}_{t-1})$, which is closer to the U.S. data (where the statistic is about one-third) than that produced by standard consumption models stemming from Hall (1978).[36] Our baseline and hetero models also imply positive $corr(\triangle \log \boldsymbol{C}_t, \triangle \log \boldsymbol{C}_{t-1})$, although not as high as that predicted by the KS model.

As argued in Hall (1978), a standard consumption model implies that consumption growth can be approximated by a random walk, and economists often introduce habits into the utility function or 'sticky expectations' (Carroll and Slacalek (2008)) to generate sticky consumption growth to match the empirical data. At first blush, it seems puzzling that this model (which includes neither sticky expectations nor habits) generates a substantial violation of the random walk proposition. This puzzle does not seem to have been noticed in the previous literature on the KS model, but after some investigation we determined that the KS model's sticky consumption growth is produced by the high degree of serial correlation in interest rates in the model, which results from the assumption about the process of aggregate productivity shocks (see Appendix C for details).

3.1.2 Distribution Statistics

Our simulations confirm the well-known fact (originally noted by Krusell and Smith (1998) themselves) that the (baseline) KS model performs badly in matching the wealth distribution, in particular the ratios of wealth held by the top percentiles (third column of Table 9). For example, the model predicts that the top 1 percent own only 2 percent of aggregate wealth, while the U.S. data reports that the top 1 percent hold as much as about 30 percent. Further, the KS model generates a distribution of wealth in which most households' wealth levels are not very far from the wealth target of a representative agent in the perfect foresight version of the model. Our solution of the KS model implies that as much as 80 percent of all households have wealth between 0.5 times mean wealth and 1.5 times mean wealth, while the corresponding fraction is only 20-25 percent in the SCF1992-2004.

Krusell and Smith (1998) well understood the shortcoming of the baseline version of their model, and therefore examined a variant of the model that incorporated a form of discount rate heterogeneity as an experiment to improve

[36] However, it should be noted that the serial correlation coefficient for consumption growth calculated using the U.S. data may be significantly underestimated because of measurement error and some other factors (Carroll, Sommer, and Slacalek (2008)). This would imply that the models above do not reproduce stickiness in aggregate consumption growth well.

Table 9: Distributions of Wealth (in percent)

	Buffer Stock		Krusell-Smith	KS hetero	U.S. data
	Baseline	Hetero	Our solution		
Top 1%	10.3	24.2	2.3	24.0	29.6
Top 10%	38.2	65.0	17.8	73.0	66.1
Top 20%	54.5	80.3	31.9	88.0	79.5
Top 40%	75.5	92.5	55.5	–	92.9
Top 60%	88.8	97.1	74.7	–	98.7

Notes: U.S. data is the SCF from Castaneda, Diaz-Gimenez, and Rios-Rull (2003).

the model's coherence with the wealth distribution data.[37] As they showed, this simple form of heterogeneity did improve the model's ability to match the wealth holdings of the top percentiles (see "KS hetero" column in the table).

Our baseline model increases inequality in the wealth distribution relative to the (baseline) KS model, and the ratios of the top 1 percent and the 10 percent are as high as about 10 percent and nearly 40 percent, respectively (first column). However, these are still much lower than the ratios in the U.S. data.

Finally, our hetero model does a much better job in reproducing the U.S. wealth distribution (second column in the table and Figure 1). Although the model still underpredicts wealth holdings at the top 1 percent a bit, it closely matches the U.S. distribution in the middle part, sustantially outperforming the KS hetero model.

3.1.3 The Aggregate Marginal Propensity to Consume

Distributions of wealth have implications for the aggregate MPC out of transitory income. Figure 2 plots our baseline model's individual consumption function in the good (aggregate) state, with the horizontal axis being cash on hand normalized by the level of (quarterly) permanent income. The figure shows that MPC is higher when the level of normalized cash on hand is lower and vice versa,[38] implying that the average MPC is higher when a larger fraction of households has less (normalized) cash on hand.

There are more households with little wealth in our baseline model than in the (baseline) KS model ("KS" in the figure), as illustrated in Figure 1. This should produce a higher average MPC in our baseline model given the concave

[37] Specifically, they assume that the discount factor takes one of the three values 0.9858, 0.9894 and 0.9930, and that the transition follows a Markov process.

[38] Consumption functions of the KS and hetero models have a similar form.

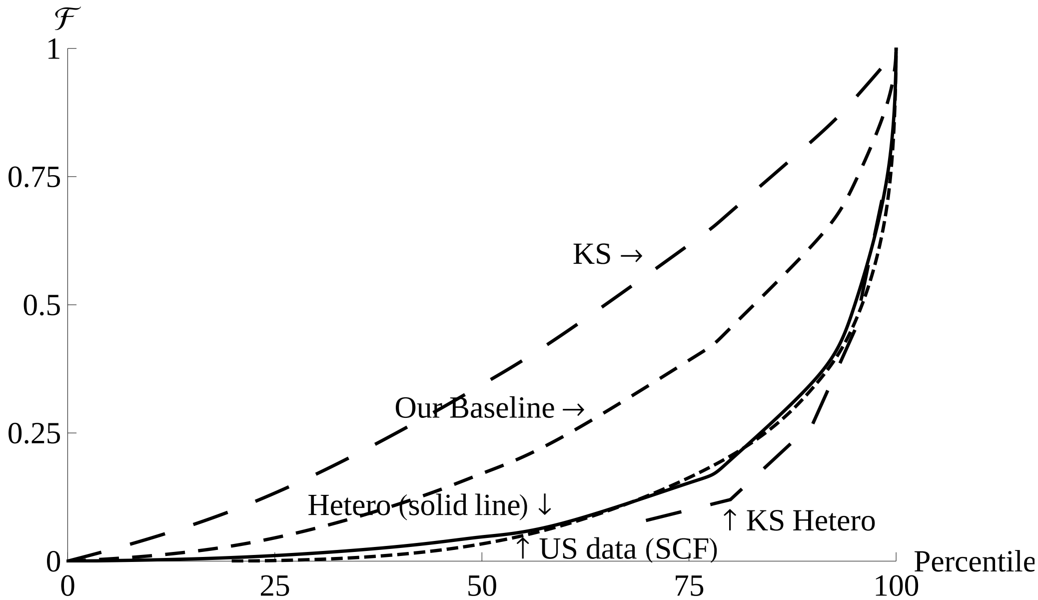

Figure 1: Cumulative Distribution of Wealth

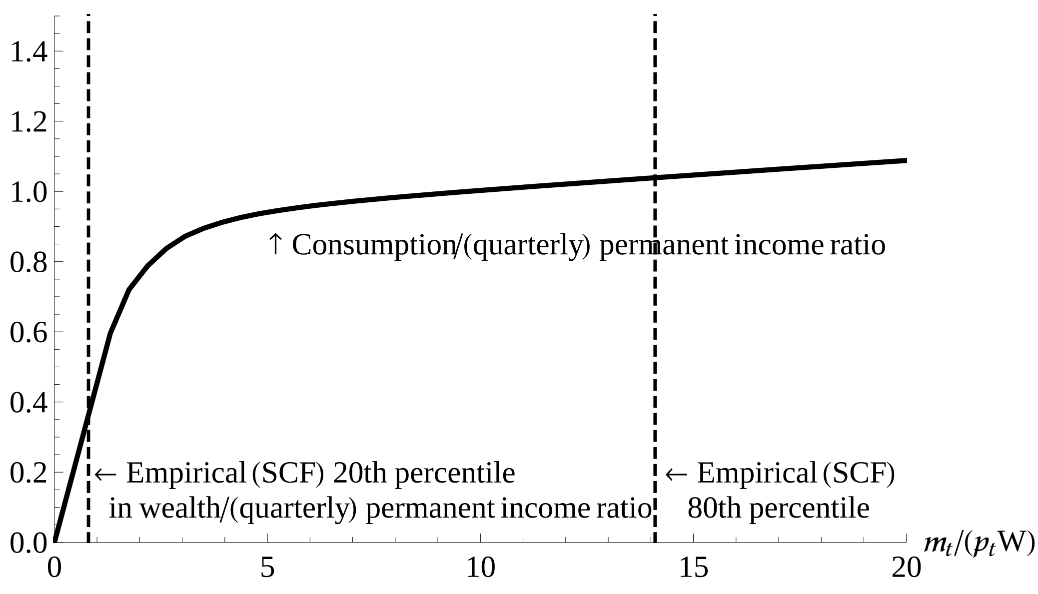

Figure 2: Consumption Function

consumption function (Figure 2).

Indeed, the average MPC out of transitory income in our baseline model is 0.10 in annual terms, about double the value in the KS model (0.05) (Table 10).[39] However, this is still much lower than typical empirical estimates which are between 0.2 and 0.5 (e.g., Parker (1999), Souleles (1999), Lusardi (1996)). Our 'hetero' model produces a higher average MPC (0.2) which is at the lower bound of this range, since in the hetero model there are even more households with less wealth (Figure 1).

MPC's are generally higher among low wealth/income households and the unemployed in both our baseline and hetero models (rest of the rows in Table 10). These results provide the basis for a common piece of conventional wisdom about the effects of economic stimulus mentioned in our introduction: If the purpose of the stimulus payments is to stimulte consumption, it makes much more sense to target them to low-wealth households than to distribute them uniformly to the population as a whole.

4 A More Plausible and More Tractable Aggregate Process

The KS process for aggregate productivity shocks has little empirical foundation; indeed, it seems to have been intended by the authors more as an illustration of how one might incorporate business cycles than as a serious candidate for a realistic description of aggregate dynamics. This section introduces an aggregate income process that is considerably more tractable, and that is also a much closer match to the aggregate data, than the KS process. We regard the version of our model with this new income process as the 'preferred' version of our model to be used as a starting point for future research.

Aggregate Process Following Carroll and Slacalek (2008), consider the aggregate production function:

$$\boldsymbol{K}_t^{\alpha}(\boldsymbol{L}_t)^{1-\alpha},$$

where $\boldsymbol{K}_t$ is per capita capital and $\boldsymbol{L}_t$ is aggregate labor supply. Most importantly, the aggregate state (good or bad) does not exist in this model. With aggregate productivity Ψ_t removed from the production function, productivity is now captured by $\boldsymbol{L}_t$. $\boldsymbol{L}_t = P_t \Xi_t$; P_t is aggregate permanent productivity, where $P_{t+1} = P_t \Psi_{t+1}$; Ψ_{t+1} is the aggregate permanent shock; and Ξ_t is the aggregate transitory shock.

[39] The average MPC calculated here can be interpreted as how much is spent on average when one dollar is disbursed to all households.

Table 10: Marginal Propensity to Consume in Annual Terms

	Buffer Stock		Krusell-Smith
	Baseline	Hetero	Our solution
Average	0.10	0.20	0.05
By wealth/permanent income ratio			
Top 1%	0.06	0.06	0.04
Top 10%	0.06	0.06	0.04
Top 20%	0.06	0.06	0.04
Top 40%	0.06	0.06	0.04
Top 60%	0.07	0.08	0.04
Bottom 1/2	0.13	0.31	0.05
By income			
Top 1%	0.08	0.13	0.05
Top 10%	0.08	0.14	0.05
Top 20%	0.09	0.15	0.05
Top 40%	0.11	0.16	0.05
Top 60%	0.11	0.17	0.05
Bottom 1/2	0.08	0.22	0.05
By employment status			
Employed	0.09	0.18	0.05
Unemployed	0.17	0.36	0.06

Notes: Annual MPC is calculated by $1 - (1-$ quarterly MPC$)^4$.

Table 11: Parameter Values for Aggregate Shocks

Description	Param	Value
Variance of Log Ψ_t	σ^2_Ψ	0.00004
Variance of Log Ξ_t	σ^2_Ξ	0.00001

Both Ψ_t and Ξ_t are assumed to be log normally distributed with mean one, and their log variances are from Carroll and Slacalek (2008), who have estimated them using U.S. data (Table 11).

Solution By keeping the model structure the same as in the previous section (other than the aggregate process above), the new model is easier to solve. In particular, the nonexistence of the aggregate state (and thus the state-dependent aggregate transition process) allows us to reduce the number of state variables to two (m_t and $\boldsymbol{K}_t$) after normalizing the model by $p_t P_t$ (as elaborated in Carroll and Slacalek (2008)). As before, the household needs to know the law of motion of $\boldsymbol{K}_t$, which can be obtained by following essentially the same method as described in the Appendix B.

Model Performance The performance of the models is similar to that under the KS aggregate process, using the same set of parameter values.[40] Figure 3 confirms that the hetero model can replicate closely the U.S. wealth distribution. The MPCs (not reported here) are also very similar to those in Table 10 produced by the models with the KS aggregate process.

5 Conclusion

This paper found that the performance in replicating wealth distributions of a KS type model can be improved significantly by introducing i) a microfounded income process, ii) finite lifetimes, and iii) heterogeneity in time preference factors. Moreover, such a model can produce plausible macroeconomic implications such as those about the MPC.

[40] Given that there is no aggregate state in the economy, we assume that the unemployment rate u_t is fixed at 0.07 (same as in Section 2).

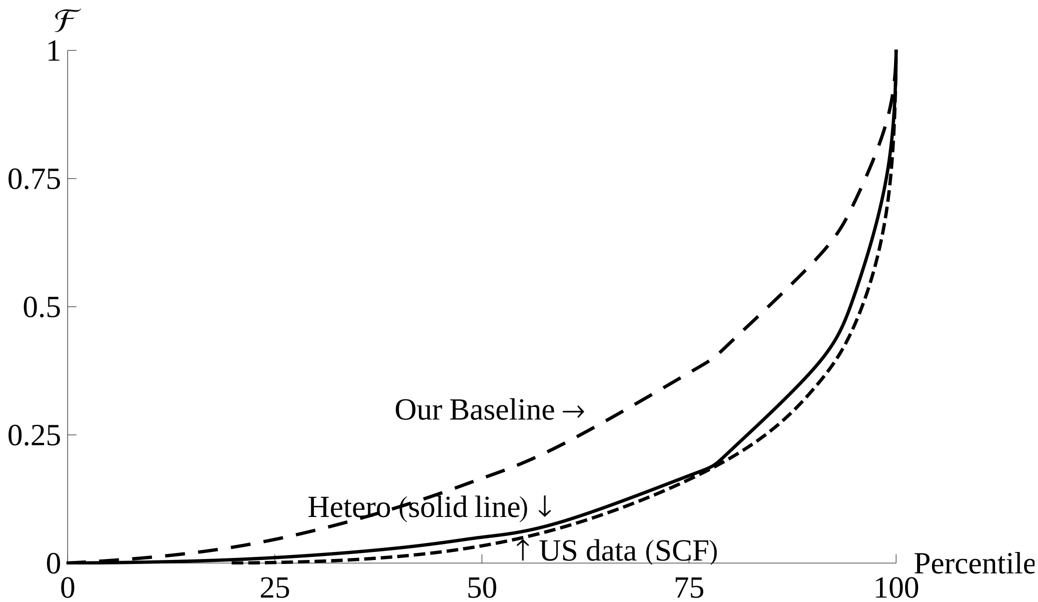

Figure 3: Cumulative Distribution of Wealth

Appendix

A Derivation of $\mathbb{M}[p^2]$

The evolution of the square of p is given by

$$\begin{aligned} p_{t+1,i} &= p_{t,i}\psi_{t+1,i}(1-\mathsf{d}_{t+1,i}) + \mathsf{d}_{t+1,i} \\ p^2_{t+1,i} &= (p_{t,i}\psi_{t+1,i}(1-\mathsf{d}_{t+1,i}))^2 + 2p_{t,i}\psi_{t+1,i}\underbrace{\mathsf{d}_{t+1,i}(1-\mathsf{d}_{t+1,i})}_{=0} + \mathsf{d}^2_{t+1,i}, \end{aligned}$$

where $\mathsf{d}_{t+1,i} = 1$ if household i dies.

Because $\mathbb{E}_t[(1-\mathsf{d}_{t+1,i})^2] = 1-\mathsf{D}$ and $\mathbb{E}_t[\mathsf{d}^2_{t+1,i}] = \mathsf{D}$, we have

$$\begin{aligned} \mathbb{E}_t[p^2_{t+1,i}] &= \mathbb{E}_t[(p_{t,i}\psi_{t+1,i}(1-\mathsf{d}_{t+1,i}))^2] + \mathsf{D} \\ &= p^2_{t,i}\not{\mathsf{D}}\mathbb{E}[\psi^2] + \mathsf{D}, \end{aligned}$$

so we have

$$\mathbb{M}\left[p^2_{t+1}\right] = \mathbb{M}[p^2_t]\not{\mathsf{D}}\mathbb{E}[\psi^2] + \mathsf{D},$$

and the steady state expected level of $\mathbb{M}[p^2] \equiv \lim_{t\to\infty}\mathbb{M}[p^2_t]$ can be found from

$$\begin{aligned} \mathbb{M}[p^2] &= \mathsf{D} + \not{\mathsf{D}}\mathbb{E}[\psi^2]\mathbb{M}[p^2] \\ \mathbb{M}[p^2] &= \left(\frac{\mathsf{D}}{1-\not{\mathsf{D}}\mathbb{E}[\psi^2]}\right). \end{aligned}$$

B Solution Method to Obtain Law of Motion

B.1 Solution Methods

Broadly speaking, the literature takes one of the following two approaches in solving the problem in Subsection **??**:

1. Relying on simulation to obtain the law of motion of per capita capital
2. (In principle) not relying on simulation to obtain the law of motion of per capita capital

Table 12 lists some existing articles that solve the KS model according to this categorization. All articles in the table except Kim, Kim, and Kollmann (2007) solve the exact KS model using various methods.[41]

[41] Kim, Kim, and Kollmann (2007) modified the form of the utility function.

Table 12: List of Existing Articles

Relying on simulation	
Young (2007)	Grid-based method
Den Haan (2010)	Grid-based method
(In principle) not relying on simulation	
Algan, Allais, and Haan (2008)	Parameterization method
Reiter (2010)	Parameterization method
Kim, Kim, and Kollmann (2007)	Perturbation method
Den Haan and Rendahl (2008)	Explicit aggregation method

The advantage of the first approach is that simulation performed to obtain the law of motion generates micro data, which can be used directly to investigate issues such as wealth distribution. The disadvantage is that this approach is generally subject to cross-sectional sampling variation, because this approach typically performs simulation using a finite number of households. Young (2007) and Den Haan (2010)'s approaches can also be categorized in the first approach but avoid cross-sectional sampling variation by running *nonstochastic* simulation that approximates the density of wealth with a histogram.

The advantages of the second approach are: i) there is no cross-sectional sampling variation; ii) it is generally faster than the first approach. There are some studies that have used the second approach. Algan, Allais, and Haan (2008) and Reiter (2010) find a wealth distribution function of various moments,[42] while Reiter (2010) calculates a matrix for the transition probabilities of individual wealth (see Appendix D for details about his technique). Kim, Kim, and Kollmann (2007) use a perturbation method that linearizes the system. The problem with this method is that they are unable to solve the exact same KS model and thus modify the form of the utility function, although they can solve a related problem very quickly.

While we could use the second approach, we adopt the first approach. The first approach directly generates various micro data (e.g., individual wealth and MPC), which can be used to examine key issues in this chapter, such as wealth distribution and the aggregate MPC. Details about our algorithm are in the next subsection.

[42] Simulation plays a part in Algan, Allais, and Haan (2008)'s method (they use simulation to find the function).

B.2 Our Algorithm

Our algorithm to solve the problem in Subsection **??** follows closely that in Krusell and Smith (1998), which relies on stochastic simulation. Their contribution is that they find that only per capita capital today ($\boldsymbol{K}_t$) is sufficient to predict per capita capital tomorrow ($\boldsymbol{K}_{t+1}$). The specific procedure we take based on their finding is as follows:

1. Solve for the optimal individual decision rules given some "beliefs" b that determine the (expected) law of motion of per capita capital. The law of motion is assumed to take the following log-linear form determined by $b = (b_0, b_1, b_0', b_1')$:
$$\log \boldsymbol{K}_{t+1} = b_0 + b_1 \log \boldsymbol{K}_t$$
if the aggregate state in period t is good ($\Psi_t = 1 + \triangle^{\Psi}$), and
$$\log \boldsymbol{K}_{t+1} = b_0' + b_1' \log \boldsymbol{K}_t$$
if the aggregate state is bad ($\Psi_t = 1 - \triangle^{\Psi}$).

2. Simulate the economy populated by $7,000$ households (for the case of the hetero model) assuming that they follow these decision rules for $1,100$ periods.[43] When starting a simulation, $p_{t,i} = 1$ for all i, the distribution of $m_{t,i}$ is generated assuming $k_{t,i}$ is equal to its steady state level (38.0) for all i, and $\Psi_t = 1 + \triangle^{\Psi}$ (the aggregate state is good).[44] If households are dead and replaced by unrelated newborns, they start a life with $p_{t,i} = 1$ and $k_{t,i} = 0$.

3. Estimate $\tilde{b}$, which determines the law of motion of per capita capital, using the last $1,000$ periods of data generated by the simulation (we discard the first 100 periods).

4. Compute an improved vector for the next iteration by $\hat{b} = (1 - \eta)\tilde{b} + \eta b$. ($\eta = 1/2$ is used for the hetero model.[45])

We repeat this process until $\hat{b} = b$ with a given degree of precision.[46]

From the second iteration and thereafter, we use the terminal distribution of wealth (and permanent component of income (p)) in the previous iteration as

[43] The length of the simulation ($1,100$ periods) is from Maliar, Maliar, and Valli (2010).

[44] The steady state level of $k_{t,i}$ is calculated by $\bar{k} = (\alpha\beta/(1 - \beta\daleth))^{1/(1-\alpha)}$. With $k_{t,i} = 38.0$ for all i, $\boldsymbol{k}_{t,i} = \boldsymbol{K}_t = 41.2$.

[45] Our experiments found that we can reach the solution faster with $\eta = 1/2$.

[46] In our analysis below, the process is iterated until the difference between each estimate (b_0, b_1, b_0', or b_1') and its previous value is smaller than 1 percent.

the initial one. For the case of the hetero model, the number of households is multiplied by 10 in the final two (or three) iterations to reduce cross-sectional simulation error.[47]

While we can eventually obtain some solution whatever the initial b is, we use b obtained using the representative agent model as the starting point. This can significantly reduce the time needed to obtain the solution.

Parameter values to solve the model are from Table 1, Table 4 (except for the unemployment rate u_t), and Table 6. The time preference factors are the estimates in Section 2.

B.3 Tricks to Reduce Simulation Errors

In obtaining the aggregate law, some tricks are introduced in the simulation to reduce simulation errors (or to speed up the solution given a degree of estimate precision). These tricks are applied to elements including:

- Death. When death is concentrated among households at the very top of the wealth distribution, then per capita capital would be at a lower than normal level. To alleviate simulation errors from this source, each period we: i) sort households by wealth level, ii) construct groups, the size of which is the inverse of the death probability (under our parameter choice, the size of each group is 200 and the first group contains households from the wealthiest to the 200th), and iii) pick one household that dies within each group.

- Permanent income shocks. In our methodology, permanent shocks to income are approximated by n discrete points. Similarly to the death element, after sorting we set up groups each of size n. We randomize shocks within each group subject to the constraint that each shock point is experienced by one of the group members every period, making the group mean of the shocks equal to the theoretical mean.[48]

B.4 Estimated Law of Motion

When we simulate the baseline model in Subsection **??**, the estimated law of motion is

---

[47] This is enough to ensure that the maximum deviation of each estimate of b_0, b_1, b'_0 and b'_1 from its previous value is less than 1 percent.

[48] This idea is motivated by Braun, Li, and Stachurski (2009), who proposed the estimation of densities with smaller simulation errors by calculating conditional densities given simulated data.

$$\log \boldsymbol{K}_{t+1} = 0.141 + 0.962 \log \boldsymbol{K}_t$$

if the aggregate state is good, and

$$\log \boldsymbol{K}_{t+1} = 0.123 + 0.966 \log \boldsymbol{K}_t$$

if the aggregate state is bad.[49]

In the case of the hetero model, the estimated law of motion is

$$\log \boldsymbol{K}_{t+1} = 0.154 + 0.959 \log \boldsymbol{K}_t$$

if the aggregate state is good, and

$$\log \boldsymbol{K}_{t+1} = 0.141 + 0.961 \log \boldsymbol{K}_t$$

if the aggregate state is bad.[50]

Finally, for our solution of the KS model, we estimate the law of motion as follows:

$$\log \boldsymbol{K}_{t+1} = 0.138 + 0.963 \log \boldsymbol{K}_t \tag{17}$$

if the aggregate state is good, and

$$\log \boldsymbol{K}_{t+1} = 0.122 + 0.966 \log \boldsymbol{K}_t \tag{18}$$

if the aggregate state is bad.[51] The coefficients in (17) and (18) are very close to those estimated in other articles that examine the KS model (e.g., Maliar, Maliar, and Valli (2010)).[52]

---

[49] R^2 is high and over 0.9999 for both states. We should interpret R^2 with caution, because a high R^2 does not necessarily mean a high accuracy of the solution; R^2 only measures in-sample fit (Den Haan (2007) discusses details).

[50] R^2 is greater than 0.9999 for both states.

[51] R^2 is greater than 0.99999 for both states.

[52] Using parameter values proposed in Den Haan, Judd, and Julliard (2007), Maliar, Maliar, and Valli (2010) estimate the law of motion in the KS model as follows:

$$\log \boldsymbol{K}_{t+1} = 0.138 + 0.963 \log \boldsymbol{K}_t \quad R^2 > 0.9999$$

if the aggregate state is good, and

$$\log \boldsymbol{K}_{t+1} = 0.124 + 0.966 \log \boldsymbol{K}_t \quad R^2 > 0.9999$$

if the aggregate state is bad.

C Experiment to Understand Sticky Consumption Growth

Although $corr(\triangle \log \boldsymbol{C}_t, \triangle \log \boldsymbol{C}_{t-1})$ produced in Subsection 3.1 may not be high enough relative to that observed in the U.S. data, it is still not clear why they produce a relatively high $corr(\triangle \log \boldsymbol{C}_t, \triangle \log \boldsymbol{C}_{t-1})$.

Previous studies that have examined KS type models have not investigated this issue. Using the KS model, we performed an experiment to understand the phenomenon better. In this experiment we assume that the aggregate state switches from good to bad (or from bad to good) every eight quarters.[53]

Figure 4 plots $\triangle \log \boldsymbol{C}_t$ for 24 quarters in the experiment (the state is bad for the first eight quarters, good for the next eight quarters, and bad for the final eight quarters). The figure shows that $\triangle \log \boldsymbol{C}_t$ is very persistent (it is negative in the bad state and positive in the good state), resulting in a relatively high $corr(\triangle \log \boldsymbol{C}_t, \triangle \log \boldsymbol{C}_{t-1})$.

It is easy to understand that $\triangle \log \boldsymbol{C}_t$ is higher when the state is good (and vice versa) given the following facts:

- A first order approximation of the Euler equation yields:

$$\triangle \log \boldsymbol{C}_t \approx \rho^{-1}(\mathsf{r}_t - (1 - \beta + \delta)) + \varepsilon_t, \tag{19}$$

  where ρ is the coefficient of relative risk aversion, r_t is the interest rate, β is the time preference factor, δ is the depreciation rate, and $\mathbb{E}_{t-1}[\varepsilon_t] = 0$. Indeed, when we conduct an IV regression of equation (19) using the data that produced Table 8,[54] the estimate of ρ^{-1} is 0.95 (with a standard deviation of 0.08) and close to the actual value of ρ^{-1} $(= 1)$, which suggests that the linear approximation (19) is largely valid.

- When the state is good, $\mathsf{r}_t = \alpha \Psi_t (\boldsymbol{K}_t / \bar{l} \boldsymbol{L}_t)^{\alpha - 1}$ (from (16)) is higher because Ψ_t (aggregate productivity) is higher, as can be seen in Figure 5, which plots the dynamics of r_t for the 24 quarters.

Unlike in this experiment, one state generally does not last for exactly eight quarters in typical simulation. However, one state shifts to another with only a low probability $(= 0.125)$, producing sticky aggregate consumption growth (and a relatively high $corr(\triangle \log \boldsymbol{C}_t, \triangle \log \boldsymbol{C}_{t-1})$) for the same mechanisms as in the experiment above.

[53] Because one state switches to another with a probability of 0.125, the average length of each state is eight quarters in typical simulation.

[54] We use r_{t-1} as an instrument of r_t.

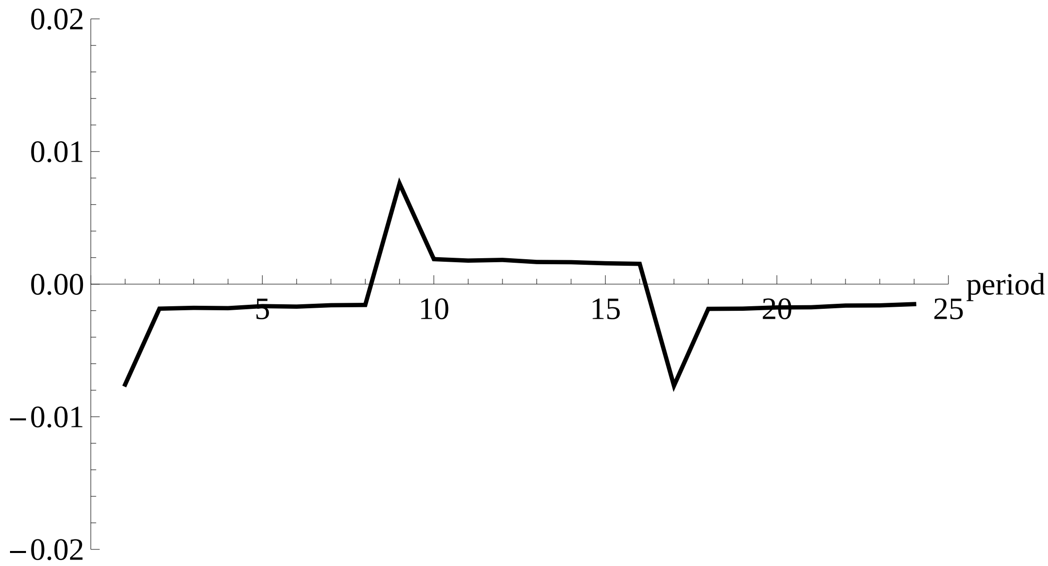

Figure 4: Dynamics of $\triangle \log \boldsymbol{C}_t$

In sum, a relatively high $corr(\triangle \log \boldsymbol{C}_t, \triangle \log \boldsymbol{C}_{t-1})$ in the KS model can be interpreted as a consequence of the persistent behavior of the interest rate r_t. Indeed, $corr(\varepsilon_t, \varepsilon_{t-1})$ (where ε_t is the error term in (19)), which measures the autocorrelation after the effects of the interest rate are removed, is much lower than $corr(\triangle \log \boldsymbol{C}_t, \triangle \log \boldsymbol{C}_{t-1})$ and only 0.01.[55]

D Transition Matrix Method

This appendix summarizes Reiter (2010)'s transition matrix method and its application to the models in this chapter.

D.1 Method

Based on a heterogeneous agents model, Reiter (2010) proposes to calculate a matrix T (n by n matrix) for transition probabilities of individual wealth (level) between periods and compute a vector d for the steady state distribution of wealth

[55]When the AR1 coefficient of ε_t is estimated (the equation $\varepsilon_t = \phi \varepsilon_{t-1}$ is estimated), the estimate is 0.01 and is not statistically significant (the standard deviation is 0.03).

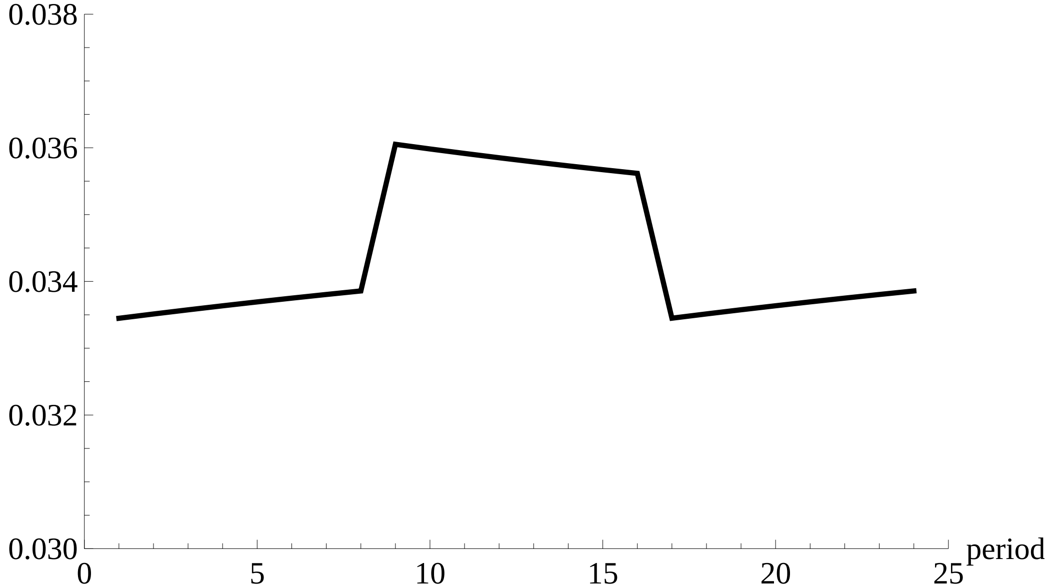

Figure 5: Dynamics of r_t

using this matrix (the steady state distribution d is calculated by solving $d = Td$). There are a couple of advantages in using such a nonsimulation method; it is generally fast and does not produce simulation errors. The downside with this technique is that if there are two or more state variables as in the models in this chapter, this method is computationally harder to use. The reason is that the steady state needs to be described by a matrix with two dimensions or higher (no longer a vector) and solving the transition matrix and the steady state is very costly.

However, this does not mean that Reiter (2010)'s transition matrix method is not useful for our models where there are two or more state variables. In fact, this chapter partially applies Reiter's method to the models in Section 2, significantly speeding up the search. In the models without aggregate shocks, it is relatively easy to compute a transition matrix for wealth/permanent income *ratios* (k) and its steady state given parameter values (remember that the models are solved with one state variable after normalization by permanent income). Therefore, it is straightforward to estimate parameter values that match the ratio of aggregate capital (level) to output with its target, assuming a certain distribution of permanent income. These estimates can be used as a good initial guess for the formal search (using simulation) of the parameter values.

Table 13: Wealth/(Quarterly) Permanent Income Ratio

	Hetero	SCF1992	SCF1995	SCF1998	SCF2001	SCF2004
Top 10%	35.7	22.9	22.9	29.3	30.9	29.6
Top 20%	17.1	14.1	14.1	17.3	18.5	18.5
Top 40%	6.9	7.0	6.8	8.0	8.7	8.8
Top 60%	3.8	3.1	3.2	3.4	3.8	3.8

D.2 Application to a Model with Heterogeneity

Furthermore, if we only match wealth/permanent income ratios (k) (not level variables), we can fully apply Reiter (2010)'s method to estimate the parameter values. Taking this approach, below we search for the parameter values (β_{low} and β_{high}) of the model with heterogeneous time preference factors in Section 2 that match the wealth/(quarterly) permanent income ratios at the top 20 percent, 40 percent, and 60 percent in the distribution (subject to the constraint that the average wealth/(quarterly) permanent income ratio matches its steady state level). The empirical counterparts are calculated using the SCF, interpreting "normal" income reported in the SCF as permanent income.[56]

The estimates of β_{low} and β_{high} are 0.9837 and 0.9900, respectively. The interval $\beta_{high} - \beta_{low}$ is narrower than the estimate in Section 2 (Table 5). Table 13 compares the wealth/(quarterly) permanent income ratios produced by the hetero model with the U.S. data, while Figure 6 plots the model's prediction and the U.S. data (SCF1998). These results again confirm our hetero model's ability to match the middle part of the distribution.

[56]Because the SCF reports *annual* "normal" income, it is divided by 4 when calculating the wealth/(quarterly)permanent income ratios.

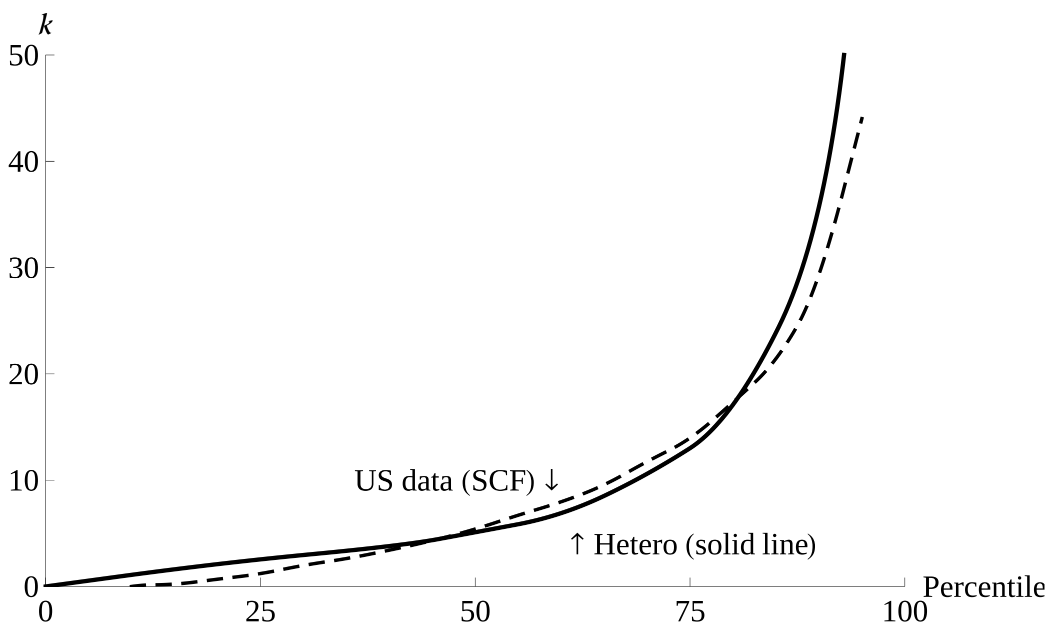

Figure 6: Distribution of Wealth/Permanent Income Ratio

References

Algan, Yann, Oliver Allais, and Wouter J. Den Haan (2008): "Solving Heterogeneous-agent Models with Parameterized Cross-sectional Distributions," Journal of Economic Dynamics and Control, 32(3), 875–908.

Blanchard, Olivier J. (1985): "Debt, Deficits, and Finite Horizons," Journal of Political Economy, 93(2), 223–247.

Blundell, Richard, Luigi Pistaferri, and Ian Preston (2008): "Consumption Inequality and Partial Insurance," Manuscript.

Braun, Richard Anton, Huiyu Li, and John Stachurski (2009): "Computing Densities and Expectations in Stochastic Recursive Economies: Generalized Look-Ahead Techniques," Manuscript.

Campbell, John Y., and N. Gregory Mankiw (1991): "The Response of Consumption to Income: A Cross-Country Investigation," European Economic Review, 35, 723–67.

Carroll, Christopher D. (1992): "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," Brookings Papers on Economic Activity, 1992(2), 61–156, Available at `http://econ.jhu.edu/people/ccarroll/BufferStockBPEA.pdf`.

——— (2000a): "Requiem for the Representative Consumer? Aggregate Implications of Microeconomic Consumption Behavior," American Economic Review, 90(2), 110–115.

——— (2000b): "Why Do the Rich Save So Much?," in Does Atlas Shrug? The Economic Consequences of Taxing the Rich, ed. by Joel B. Slemrod. Harvard University Press, Available at `http://econ.jhu.edu/people/ccarroll/Why.pdf`.

——— (2009): "Theoretical Foundations of Buffer Stock Saving," Revise and Resubmit, *Review of Economic Studies*, Latest version available at http://econ.jhu.edu/people/ccarroll/BufferStockTheory.pdf.

Carroll, Christopher D., and Miles S. Kimball (1996): "On the Concavity of the Consumption Function," Econometrica, 64(4), 981–992, Available at `http://econ.jhu.edu/people/ccarroll/concavity.pdf`.

Carroll, Christopher D., and Andrew A. Samwick (1997): "The Nature of Precautionary Wealth," Journal of Monetary Economics, 40(1), 41–71, Available at http://econ.jhu.edu/people/ccarroll/nature.pdf.

Carroll, Christopher D., and Jiri Slacalek (2008): "Sticky Expectations and Consumption Dynamics," Manuscript.

Carroll, Christopher D., Martin Sommer, and Jiri Slacalek (2008): "International Evidence on Sticky Consumption Growth," Johns Hopkins University Working Paper Number 542, Available at http://econ.jhu.edu/people/ccarroll/papers/cssIntlStickyC http://econ.jhu.edu/people/ccarroll/papers/cssIntlStickyC.pdf http://econ.jhu.edu/people/ccarroll/papers/cssIntlStickyC.zip.

Castaneda, Ana, Javier Diaz-Gimenez, and Jose-Victor Rios-Rull (2003): "Accounting for the U.S. Earnings and Wealth Inequality," Journal of Political Economy, 111(4), 818–857.

Deaton, Angus S. (1991): "Saving and Liquidity Constraints," Econometrica, 59, 1221–1248, Available at http://ideas.repec.org/a/ecm/emetrp/v59y1991i5p1221-48.html.

Den Haan, Wouter J. (2007): "Assessing the Accuracy of the Aggregate Law of Motion in Models with Heterogeneous Agents," Manuscript.

——— (2010): "Comparison of Solutions to the Incomplete Markets Model with Aggregate Uncertainty," Journal of Economic Dynamics and Control, 34(1), 4–27.

Den Haan, Wouter J., Ken Judd, and Michel Julliard (2007): "Description of Model B and Exercises," Manuscript.

Den Haan, Wouter J., and Pontus Rendahl (2008): "Solving the Incomplete Markets Models with Aggregate Uncertainty Using Explicit Aggregation," Manuscript.

Diamond, Peter A. (1965): "National Debt in a Neoclassical Growth Model," American Economic Review, 55, 1126–1150, Available at http://www.ssc.wisc.edu/~manuelli/712/NationalDebtDiamond.pdf.

Friedman, Milton A. (1957): A Theory of the Consumption Function. Princeton University Press.

Guvenen, Fatih (2009): "An Empirical Investigation of Labor Income Processes," Review of Economic Dynamics, 12.

Hall, Robert E. (1978): "Stochastic Implications of the Life-Cycle/Permanent Income Hypothesis: Theory and Evidence," Journal of Political Economy, 96, 971–87, Available at http://www.stanford.edu/~rehall/Stochastic-JPE-Dec-1978.pdf.

Hryshko, Dmytro (2009): "RIP to HIP: The Data Reject Heterogeneous Labor Income Profiles," Manuscript, University of Alberta.

——— (2010): "RIP to HIP: The Data Reject Heterogeneous Labor Income Profiles," Manuscript, University of Alberta.

Hubbard, R. Glenn, Jonathan S. Skinner, and Stephen P. Zeldes (1994): "The Importance of Precautionary Motives for Explaining Individual and Aggregate Saving," in The Carnegie-Rochester Conference Series on Public Policy, ed. by Allan H. Meltzer, and Charles I. Plosser, vol. 40, pp. 59–126.

Jensen, Shane T., and Stephen H. Shore (2008): "Changes in the Distribution of Income Volatility," Manuscript.

Kennickell, Arthur B. (1995): "Saving and Permanent Income: Evidence from the 1992 SCF," Board of Governors of the Federal Reserve System.

Kim, Jinill, Sunghyun Henry Kim, and Robert Kollmann (2007): "Perturbation Analysis of a Heterogenous-Agent Model," Manuscript.

Krusell, Per, and Anthony A. Smith (1998): "Income and Wealth Heterogeneity in the Macroeconomy," Journal of Political Economy, 106(5), 867–896.

Low, Hamish, Costas Meghir, and Luigi Pistaferri (2005): "Wage Risk and Employment Over the Life Cycle," Manuscript, Stanford University.

Lucas, Robert E. (1976): "Econometric Policy Evaluation: A Critique," in The Phillips Curve and Labour Markets, ed. by Karl Brunner, and Allan H. Meltzer, vol. 1, pp. 19–46. Journal of Monetary Economics (Supplement).

Lusardi, Annamaria (1996): "Permanent Income, Current Income, and Consumption: Evidence from Two Panel Data Sets," Journal of Business and Economic Statistics, 14(1), 81–90.

MacCurdy, Thomas (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," Journal of Econometrics, 18(1), 83–114.

Maliar, Lilia, Serguei Maliar, and Fernando Valli (2008): "Comparing Numerical Solutions of Models with Heterogeneous Agents (Models B): a Grid-based Euler Equation Algorithm," Manuscript.

——— (2010): "Solving the Incomplete Markets Model with Aggregate Uncertainty Using the Krusell Smith Algorithm," Journal of Economic Dynamics and Control, 34(1), 42–49.

Meghir, Costas, and Luigi Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," Journal of Business and Economic Statistics, 72(1), 1–32.

Modigliani, Franco, and Richard Brumberg (1954): "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data," in Post-Keynesian Economics, ed. by Kenneth K. Kurihara, pp. 388–436. Rutgers University Press, New Brunswick, N.J.

Moffitt, Robert A., and Peter Gottschalk (1995): "Trends in the Covariance Structure of Earnings in the U.S.: 1969-1987," Manuscript.

Nielsen, Helena Skyt, and Annette Vissing-Jorgensen (2006): "The Impact of Labor Income Risk on Educational Choices: Estimates and Implied Risk Aversion," Manuscript.

Parker, Jonathan A. (1999): "The Reaction of Household Consumption to Predictable Changes in Social Security Taxes," American Economic Review, 89(4), 959–973.

Reiter, Michael (2010): "Solving the Incomplete Markets Model with Aggregate Uncertainty by Backward Induction," Journal of Economic Dynamics and Control, 34(1), 28–35.

Samuelson, Paul A. (1958): "An exact consumption loan model of interest with or without the social contrivance of money," Journal of Political Economy, 66, 467–482.

Solow, Robert M. (2003): "Dumb and Dumber in Macroeconomics," Manuscript, MIT.

Souleles, Nicholas S. (1999): "The Response of Household Consumption to Income Tax Refunds," American Economic Review, 89(4), 947–958.

Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron (2004a): "Consumption and Risk Sharing Over the Life Cycle," Journal of Monetary Economics, 51(3), 609–633.

——— (2004b): "Cyclical Dynamics in Idiosyncratic Labor-Market Risk," Journal of Political Economy, 112(3), 695–717.

Szeidl, Adam (2006): "Invariant Distribution in Buffer-Stock Saving and Stochastic Growth Models," Manuscript, University of California at Berkeley.

Topel, Robert H. (1990): "Specific Capital, Mobility and Wages: Wages Rise with Job Seniority," NBER Working Paper Number W3294.

Topel, Robert H., and Michael P. Ward (1992): "Job Mobility and the Careers of Young Men," Quarterly Journal of Economics, 107(2), 439–479.

Young, Eric R. (2007): "Finite Forecasting Function Solutions to the Stochastic Growth Model," Manuscript.