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

October 22, 2010

Christopher D. Carroll\textsuperscript{1} Kiichi Tokuoka\textsuperscript{2}

\textsuperscript{1}JHU \textsuperscript{2}IMF

Abstract

Krusell and Smith (1998) showed that it is possible to construct rational expectations macroeconomic models with serious microfoundations. We argue that three modifications to their framework are required to fulfill its promise. First, we replace their assumption about household income dynamics with a process that matches microeconomic data. Second, our agents have finite lifetimes \textit{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 Microfoundations, Wealth Inequality, Marginal Propensity to Consume

JEL codes D12, D31, D91, E21

PDF: \url{http://econ.jhu.edu/people/ccarroll/papers/BSinKS.pdf}
Web: \url{http://econ.jhu.edu/people/ccarroll/papers/BSinKS/}
Archive: \url{http://econ.jhu.edu/people/ccarroll/papers/BSinKS.zip}

\textsuperscript{1}Carroll: Department of Economics, Johns Hopkins University, Baltimore, MD, \url{http://econ.jhu.edu/people/ccarroll/}, \texttt{ccarroll@jhu.edu}
\textsuperscript{2}Tokuoka: International Monetary Fund, \texttt{ktokuoka@imf.org}. 
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 between two microfoundational camps. “Bottom up” modelers (e.g. Modigliani and Brumberg (1954); Friedman (1957)) drew wisdom from microeconomic data and argued that macro models should be constructed by aggregation of models that matched robust micro facts. “Top down” modelers (e.g., Samuelson (1958); Diamond (1965); Hall (1978)) 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 interpretation of the term.

The tractability of representative agent models has made them appealing for business cycle analysis. 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. 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 is essential.

The bottom-up approach, however, has also had its problems. Even judged by a sympathetic standard that asks whether such models can match measured wealth heterogeneity, bottom-up models have not been as successful as might be hoped. 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 the upper parts of the wealth distribution (Carroll (1997); Cagetti (2003)). 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

\(^1\)See, e.g., Campbell and Mankiw (1991) and the vast related literature following Hall (1978). A newer literature attempts to fix the problems identified in that literature by introducing habit formation (see, e.g., Fuhrer (2000)); but this is at the cost of intensifying the conflict with microeconomic evidence (see the next footnote).

\(^2\)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).

\(^3\)Gorman (1953) shows that the essential requirement is that marginal propensities to consume be identical for all consumers. See Kirman (1992) and Solow (2003) for discussions of the deficiencies of representative agent models.
answer to the question of how to analyze systematic macroeconomic fluctuations (business cycles) in bottom-up models.

This paper aims to reconcile the warring camps. Specifically, we argue that a workhorse model that answers the main objections to both kinds of models 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). 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.’

The resulting framework differs sharply from the KS model in its implications for important microeconomic and macroeconomic questions. A timely example of such a macroeconomic question is how aggregate consumption will respond to an ‘economic stimulus payment,’ interpreted here as a one-time lump sum transfer to households. In response to a $1-per-capita payment, the baseline version of the KS model implies that the annual marginal propensity to consume (MPC) is about 0.05, almost irrespective of how the money 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 of the wealth-to-income distribution, the MPC will be about 0.22, which counts as at least a partial success since empirical evidence does seem to confirm the prediction from theory (Carroll and Kimball (1996)) that MPC’s are higher for lower-wealth households. Furthermore, since other kinds of macroeconomic shocks might also be distributed across the population in systematically different ways (for example, labor income shocks may affect a different set of households than capital income shocks), this improvement in realism may matter for important general macroeconomic questions as well as for narrower fiscal policy issues.

Section 2 of the paper begins building the model’s structure by adding microe-
conomic modeling elements to a benchmark representative agent model. Using this model (without macroeconomic dynamics), the section closes by estimating the degree of heterogeneity in impatience necessary to match the degree of inequality in the U.S. wealth distribution; we find that relatively small differences 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, and presents comparisons of our model with theirs. Section 4 improves the model by introducing an aggregate income process that is analytically simpler than the KS aggregate 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, we begin by briefly sketching a standard perfect foresight representative agent model.

The aggregate production function is

\[ Z_t K_t^\alpha (\bar{L} L_t)^{1-\alpha}, \]  

where \( Z_t \) is aggregate productivity in period \( t \), \( K_t \) is capital, \( \bar{L} \) is time worked per employee, and \( L_t \) is employment. The representative agent’s goal is to maximize discounted utility from consumption

\[ \max_{\{C_t\}} \sum_{n=0}^{\infty} \beta^n u(C_{t+n}) \]

for a CRRA utility function \( u(\bullet) = \bullet^{1-\rho}/(1 - \rho) \).\(^9\) The representative agent’s state at the time of the consumption decision is defined by two variables: \( M_t \) is market resources, and \( Z_t \) is aggregate productivity.

The transition process for \( M_t \) 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

\[ A_t = M_t - C_t, \]

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

\[ K_{t+1} = A_t. \]

\(^9\)Substitute \( u(\bullet) = \log \bullet \) for the case where \( \rho = 1 \).
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 period, when output has been produced and incorporated into resources but has not yet been consumed:

\[
M_{t+1} = \frac{\gamma K_{t+1} + Z_t K^\alpha_{t+1} (\bar{L}_{t+1})^{1-\alpha}}{\Gamma_{t+1} + K_t r_{t+1} + (\bar{L}_{t+1}) W_{t+1}},
\]

where \( r_{t+1} \) is the interest rate, \( W_{t+1} \) is the wage rate, and \( \gamma = (1 - \delta) \) is the depreciation factor for capital.

After normalizing by the productivity factor \( X_t = \frac{Z_t^{1/(1-\alpha)} (\bar{L}_t)}{1-\alpha} \), the representative agent’s problem is

\[
V(M_t, Z_t) = \max_{\{C_t\}} u(C_t) + \beta \mathbb{E}_t \left[ \Gamma_{t+1}^{1-\rho} V(M_{t+1}, Z_{t+1}) \right]
\]

s.t.

\[
A_t = M_t - C_t
\]

\[
K_{t+1} = \frac{A_t}{\Gamma_{t+1}}
\]

\[
M_{t+1} = \frac{\gamma K_{t+1} + K_t^\alpha_{t+1}}{\Gamma_{t+1} + K_t^\alpha_{t+1}},
\]

where the non-bold variables are the corresponding bold variables divided by \( X_t \) (e.g., \( A_t = A_t/X_t \), \( M_t = M_t/X_t \); \( \Gamma_{t+1} = X_{t+1}/X_t \); 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 Table 1). The model is calibrated at the quarterly frequency. When aggregate shocks are shut down (\( Z_t = 1 \) and \( \bar{L}_t = L_t \)), 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 Table 1.

Henceforth, we refer to the version of the model solved by the papers in the special JEDC volume as the “KS-JEDC” model, while we call the original KS model solved in Krusell and Smith (1998) “KS-Orig” model. (The only effective difference between the two models is the introduction of unemployment insur-

---

10 Equal to the marginal product of capital, \( \alpha Z_t K^{\alpha-1}_{t+1} (\bar{L}_{t+1})^{1-\alpha} \).
11 Equal to the marginal product of labor, \( (1-\alpha)Z_t K^{\alpha}_{t+1} (\bar{L}_{t+1})^{1-\alpha} \).
12 Details of this normalization are discussed in Carroll (2000a).
13 Examples of such authors include Young (2010) and Algan, Allais, and Haan (2008).
14 In the steady state, \( K_t/(\bar{L}_t) = k = (\alpha \beta/(1-\beta \alpha))^{1/(1-\alpha)} = 38.9 \), \( r = \alpha k^{\alpha-1} \), and \( W = (1-\alpha)k^\alpha \).
Table 1  Parameter Values and Steady State for the Representative Agent Model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time discount factor ( \beta )</td>
<td>0.99</td>
</tr>
<tr>
<td>Coefficient of relative risk aversion ( \rho )</td>
<td>1</td>
</tr>
<tr>
<td>Capital share ( \alpha )</td>
<td>0.36</td>
</tr>
<tr>
<td>Depreciation rate ( \delta )</td>
<td>0.025</td>
</tr>
<tr>
<td>Time worked per employee ( \bar{l} )</td>
<td>1/0.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Steady State</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital output ratio ( K/Y )</td>
<td>10.26</td>
</tr>
<tr>
<td>Interest rate ( r )</td>
<td>0.04</td>
</tr>
<tr>
<td>Wage rate ( W )</td>
<td>2.37</td>
</tr>
</tbody>
</table>

ance in the KS-JEDC version, which does not matter much for any substantive points.\(^{15,16}\)

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

\[
p_{t+1} = G_{t+1}p_t \psi_{t+1}
\]

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 permanent shock to income. Actual income is equal to a product of the permanent component of income, a mean-one transitory shock, and the wage rate:

\[
y_{t+1} = p_{t+1}\xi_{t+1}W_{t+1}.
\]

After taking logarithms, this income process is strikingly similar to Friedman (1957)’s characterization of income as having permanent and transitory components. Because this process has been used widely in the literature on buffer stock

\(^{15}\)To be very precise, another difference is the introduction of \( \bar{l} \) (time worked per employee) in the KS-JEDC model, but this does not have a real impact.

\(^{16}\)Details about the unemployment insurance scheme are described later in the paper.
saving, and though similar to Friedman’s formulation is not identical to it, we henceforth refer to it as the Friedman/Buffer Stock (or ‘FBS’) process.17

Table 2 summarizes the annual variances of log permanent shocks ($\sigma_\psi^2$) and log transitory shocks ($\sigma_\xi^2$) estimated by a selection of papers from the extensive literature.18 Some authors have used a process of this kind to describe the labor income process for an individual worker (top panel) while others have used it to describe the process for overall household income (bottom panel); it seems to work reasonably well in both cases.

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 the KS-JEDC model.19 This row of the table makes our point forcefully: The empirical procedures that have actually been applied to empirical micro data, if used to measure the income process households experience in a KS economy, would have produced estimates of $\sigma_\psi^2$ and $\sigma_\xi^2$ 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 KS-JEDC model’s well-known difficulty in matching the degree of wealth inequality is largely explained by its unrealistic assumption about the income process.20

2.3 Finite Lifetimes and the Finite Variance of Permanent Income in the Cross-Section

One might wish to use the FBS 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,

17Guvenen (2007) refers to a process like this one as a ‘restricted income process’ (RIP) as distinguished from a process that he proposes which is similar but which allows each agent to have a distinct idiosyncratic mean growth rate. While Guvenen’s process is plausible, Hryshko (2010) argues that there is strong evidence against the proposition that Guvenen’s preferred process describes the data as well as what Guvenen calls the ‘restricted’ process. Since incorporation of Guvenen’s income process would introduce serious modeling difficulties, it seems prudent to avoid using it unless the evidence becomes much more compelling.

18All 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.

19First, we generated income draws according to the income process in the KS-JEDC model. 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 $y_t = p_t \xi_t$ where $p_t = p_{t-1} \psi_t$. In doing so, as in Carroll and Samwick (1997), the draws of $y_t$ are excluded when $y_t$ is very low relative to its mean (see Carroll and Samwick (1997) for details about this restriction).

20The 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

<table>
<thead>
<tr>
<th></th>
<th>$\sigma^2_{\psi}$</th>
<th>$\sigma^2_{\xi}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Using individual data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MaCurdy (1982)$^\dagger$</td>
<td>0.013</td>
<td>0.045</td>
</tr>
<tr>
<td>Topel (1990)</td>
<td>0.013</td>
<td>0.017</td>
</tr>
<tr>
<td>Topel and Ward (1992)</td>
<td>0.017</td>
<td>0.013</td>
</tr>
<tr>
<td>Meghir and Pistaferri (2004)$^*$</td>
<td>0.031</td>
<td>0.032</td>
</tr>
<tr>
<td>Low, Meghir, and Pistaferri (2005)</td>
<td>0.011</td>
<td>—</td>
</tr>
<tr>
<td>Jensen and Shore (2008)$^*$</td>
<td>0.054</td>
<td>0.171</td>
</tr>
<tr>
<td>Hryshko (2009)$^*$</td>
<td>0.038</td>
<td>0.118</td>
</tr>
<tr>
<td>Guvenen (2009)</td>
<td>0.015</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Using household data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carroll (1992)</td>
<td>0.016</td>
<td>0.027</td>
</tr>
<tr>
<td>Carroll and Samwick (1997)</td>
<td>0.022</td>
<td>0.044</td>
</tr>
<tr>
<td>Storesletten, Telmer, and Yaron (2004a)</td>
<td>0.017</td>
<td>0.063</td>
</tr>
<tr>
<td>Storesletten, Telmer, and Yaron (2004b)</td>
<td>0.008 - 0.026</td>
<td>0.316</td>
</tr>
<tr>
<td>Blundell, Pistaferri, and Preston (2008)$^*$</td>
<td>0.010 - 0.030</td>
<td>0.029 - 0.055</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implied by KS-JEDC</td>
<td>0.000</td>
<td>0.039</td>
</tr>
<tr>
<td>Implied by Castaneda et al. (2003)</td>
<td>0.030</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Notes: $^\dagger$: MaCurdy (1982) did not explicitly separate $\psi_t$ and $\xi_t$, but we have extracted $\sigma^2_{\psi}$ and $\sigma^2_{\xi}$ as implications of statistics that his paper reports. First, we calculate $\text{var}(\log y_{t+d} - \log y_{t+d-1})$ and $\text{var}(\log y_{t+d-1} - \log y_{t+d-2})$ using his estimate (we set $d = 5$). Then, following Carroll and Samwick (1997) we obtain the values of $\sigma^2_{\psi}$ and $\sigma^2_{\xi}$ which can match these statistics, assuming that the income process is $y_t = p_t \xi_t$ and $p_t = p_{t-1} \psi_t$ (i.e., we solve $\text{var}(\log y_{t+d} - \log y_{t+d-1}) = d\sigma^2_{\psi} + 2\sigma^2_{\xi}$ and $\text{var}(\log y_{t+d-1} - \log y_{t+d-2}) = (d-1)\sigma^2_{\psi} + 2\sigma^2_{\xi}$). $^*$: 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. For example, Meghir and Pistaferri (2004) and Blundell, Pistaferri, and Preston (2008) assume an MA(1) process $\log \xi_t = v_t + \theta v_{t-1}$ and obtain estimates $(\sigma^2_{v_t}, \theta) = (0.0300, -0.2566)$ and $(0.0286 - 0.0544, 0.1132)$, respectively. $\sigma^2_{\xi}$ for these four articles reported in this table are calculated using their estimates.
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 \( D \) of dying before the beginning of period \( t + 1 \). (The probability of NOT dying is the cancellation of the probability of dying: \( D = 1 - \mathcal{D} \).) Households engage in a Blanchardian mutual insurance scheme: Survivors share the estates of those who die. Assuming a zero profit condition for the insurance industry, this means that the insurance scheme’s ultimate effect is simply to boost the rate of return (for survivors) by an amount exactly corresponding to the mortality rate.

In order to maintain a constant population (of mass one, uniformly distributed on the unit interval), we assume that dying households are replaced by an equal number of newborns; we write the population-mean operator as \( \mathbb{M} \left[ \bullet \right] \mid t = \int_0^1 \bullet \psi \, d\psi \). Newborns, we assume, begin life with a level of idiosyncratic permanent income equal to the mean level of idiosyncratic permanent income 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 \( \mathbb{M} [p] = 1 \) forever, while the mean of \( p^2 \) is given by\(^{21}\)

\[
\mathbb{M} [p^2] = \frac{D}{1 - \mathcal{D} \mathbb{E} [\psi^2]},
\]

The relation between \( p^2 \) and the variance of \( p \) is

\[
\sigma_p^2 = \mathbb{M} [(p - \mathbb{M} [p])^2] = \mathbb{M} [(p^2 - 2p\mathbb{M} [p] + (\mathbb{M} [p])^2)] = \mathbb{M} [p^2] - 1
\]

where the last line follows because \( \mathbb{M} [p] = 1 \).

Of course for all of this to be valid, it is necessary to impose the parametric restriction \( \mathcal{D} \mathbb{E} [\psi^2] < 1 \) (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

\(^{21}\)see Appendix A for the derivation.
Table 3 Variance of Permanent Income

<table>
<thead>
<tr>
<th>Year</th>
<th>var($p$)</th>
<th>$E[\psi^2]$</th>
<th>$\sigma^2_\psi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCF1992</td>
<td>2.5</td>
<td>1.015</td>
<td>0.015</td>
</tr>
<tr>
<td>SCF1995</td>
<td>7.5</td>
<td>1.018</td>
<td>0.018</td>
</tr>
<tr>
<td>SCF1998</td>
<td>3.1</td>
<td>1.015</td>
<td>0.015</td>
</tr>
<tr>
<td>SCF2001</td>
<td>3.6</td>
<td>1.016</td>
<td>0.016</td>
</tr>
<tr>
<td>SCF2004</td>
<td>5.2</td>
<td>1.017</td>
<td>0.017</td>
</tr>
<tr>
<td>KS-Orig or KS-JEDC</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

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. This corresponds well with our definition of permanent income (level) $p$ (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 $p^t \equiv p^t / M[p^t]$ among such households.

The results from this exercise are reported in Table 3 (with a final row that makes the point that both the KS-Orig and KS-JEDC models assumed that permanent shocks did not exist). Substituting these estimates for $\sigma^2_p$ into (7) and (8), we obtain estimates of the variance of $\psi$. Reassuringly, we can interpret the variances of $\psi$ thus obtained as being easily in the range of the estimated variances of $\log(\psi) = \sigma^2_\psi$ in Table 2. Such a correspondence, across two quite different methods of measurement, suggests there is considerable robustness to the measurement of the size of permanent shocks. (Below, we will choose a calibration for $\sigma^2_\psi$ that is in the middle range of estimates from either method.)

2.4 The Wealth Distribution with Transitory and Permanent Shocks

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

$$y_t = p_t \xi_t W_t$$  \hspace{1cm} (9)

$$p_t = p_{t-1} \psi_t$$  \hspace{1cm} (10)

22SCF1992 only asked whether the income level was about ‘normal’ or not.

23We restrict the sample to households between the ages of 25 and 60, because the interpretation of the question becomes problematic for retired households.

24So long as the variance of the permanent shocks is small, these two measures should be approximately the same.
\[ W_t = (1 - \alpha)Z_t(K_t/L_t)^\alpha, \] (11)

where \( y_t \) is noncapital income for the household in period \( t \), equal to the permanent component of noncapital income \( y_t \) multiplied by a mean-one iid shock factor \( \xi_t \) (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 \( \psi_t \), \( \mathbb{E}_t[\psi_{t+n}] = 1 \) for all \( n \geq 1 \). Lastly, \( K_t \) is per capita capital and \( L_t = 1 - u_t \) is the employment rate, where \( u_t \) is the unemployment rate. Since there is no aggregate shock, \( Z_t, K_t, L_t, \) and \( W_t \) are constant (\( Z_t = Z = 1, K_t = K, L_t = L, \) and \( W_t = W = (1 - \alpha)(K/L)^\alpha \)).

Following the assumptions in the JEDC volume, the distribution of \( \xi_t \) is as follows:

\[
\begin{align*}
\xi_t &= \mu \text{ with probability } u_t \\
&= (1 - \tau_t)\bar{\theta}_t \text{ with probability } 1 - u_t,
\end{align*}
\]

(12) (13)

where \( \mu > 0 \) is the unemployment insurance payment when unemployed and \( \tau_t = \mu u_t/\bar{L}_t \) is the rate of tax collected to pay the unemployment benefits.\(^{25}\) The probability of unemployment is constant \( (u_t = u) \); later we allow it to vary over time.

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

\[
v(m_t) = \max_{\{c_t\}} u(c_t) + \beta \mathcal{D} \mathbb{E}_t \left[ \psi_{t+1}^{1-\rho} v(m_{t+1}) \right]
\]

s.t.

\[
\begin{align*}
a_t &= m_t - c_t \\
a_t &\geq 0 \\
k_{t+1} &= a_t/(\mathcal{D}\psi_{t+1}) \\
m_{t+1} &= (1 + r)k_{t+1} + \xi_{t+1}
\end{align*}
\]

(14) (15) (16)

where the non-bold ratio variables are defined as the bold (level) variables divided by the level of permanent income \( p_t = p_tW \) (e.g., \( m_t = m_t/(p_tW) \)). The only state variable is (normalized) cash-on-hand \( m_t \). The household’s employment status is not a state variable, unlike in the KS-JEDC 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 \( \mathcal{D} \) between periods, the ef-

\(^{25}\) The KS-Orig model assumed no unemployment insurance (\( \mu = 0 \)).
Table 4 Parameter Values for Heterogenous Agents Model

<table>
<thead>
<tr>
<th>Description</th>
<th>Param</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemp Ins Payment</td>
<td>μ</td>
<td>0.15</td>
<td>Den Haan, Judd, and Juillard (2010)</td>
</tr>
<tr>
<td>Unemp Rate</td>
<td>u</td>
<td>0.07</td>
<td>Den Haan, Judd, and Juillard (2010)</td>
</tr>
<tr>
<td>Probability of Death</td>
<td>D</td>
<td>0.005</td>
<td>Yields 50 year working life</td>
</tr>
<tr>
<td>Variance of Log ( \theta_{t,i} )</td>
<td>( \sigma^2_\theta )</td>
<td>0.027 \cdot 4</td>
<td>Carroll (1992)</td>
</tr>
<tr>
<td>Variance of Log ( \psi_{t,i} )</td>
<td>( \sigma^2_\psi )</td>
<td>0.016/4</td>
<td>Carroll (1992); median in Table 3</td>
</tr>
</tbody>
</table>

Effective discount factor is \( \beta D \) (in (14)). Note that the effective interest rate is \((\gamma + r)/D\) (combining (15) and (16)).

Aside from heterogeneity in impatience (introduced below), three parameters characterize our modifications to the KS-JEDC model: \( D, \sigma^2_\theta, \) and \( \sigma^2_\psi \). \( D = 0.005 \) 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 \( \sigma^2_\theta \) is from Carroll (1992), and our calibration of \( \sigma^2_\psi = 0.016 \) is from the same source (but note that this value also matches the median value in Table 3). Other parameter values (\( \rho, \alpha, \delta, \) 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) and Bewley (1977)) has shown that models of this kind do not have a well-defined solution unless the condition

\[
\left( \frac{(\gamma + r) / \bar{D}}{\Gamma} \right) < 1
\]

is satisfied. Note that \( (\gamma + r) / \bar{D} \) is scaled by \( 1/\bar{D} \) due to the Blanchardian mutual insurance scheme as described in the previous subsection.

We admit that Carroll (1992)'s estimate is an approximation for \( \sigma^2_\theta \). While we use Carroll (1992)'s estimate as the variance of transitory income shocks conditional on being employed, Carroll (1992) did not estimate it by restricting the sample to employed households but by excluding households whose income fell once (or more) to a low level (below 10 percent of the average over the sample period).

This paper assumes that each period corresponds to a quarter, while \( \sigma^2_\psi = 0.027 \) (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 \( \sigma^2_\psi = 0.027 \) is more modest than other estimates such as in Carroll and Samwick (1997) (= 0.044). The reason why \( \sigma^2_\psi = 0.027 \) is used in this paper is that Carroll and Samwick (1997) themselves argue that their estimate of \( \sigma^2_\psi \) is almost certainly increased by measurement error.

Since \( \sigma^2_\psi \) 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).

Using quarterly income draws generated by this section’s income process with these parameter values, we have estimated the \textit{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.
holds where

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

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 income is growing (see also Szeidl (2006)). This condition is an 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.\(^{31}\)

We begin by searching for the time preference factor \( \hat{\beta} \) such that if all households had an identical \( \hat{\beta} = \hat{\beta} \) the steady-state value of the capital-to-output ratio \( (K_t/Y_t) \) would match the value that characterized the steady-state of the perfect foresight model.\(^{32}\) \( \hat{\beta} \) turns out to be 0.9887.

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 shut down a potentially very important important reason for variation in the level of wealth.

Compared to the wealth distribution implied by the KS-Orig model (or our solution of the KS-JEDC model\(^{33}\)), our model does indeed yield a substantial improvement (compare the first, third and fourth columns to the last column). However, even our model falls substantially short of matching the empirical degree of wealth inequality. 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 (see the first row of the table). This failure reflects the fact that, empirically, the distribution of wealth is considerably more unequal than the distribution of permanent income.

\(^{31}\)This near-equivalence explains why we do not bother to include a growth term in the process for noncapital income in (9)-(11) despite the presence of such a term in (6); inclusion of the income growth term should mostly just result in an offsetting effect on our estimated time preference rate, but would complicate our simulations unnecessarily.

\(^{32}\)Output is the sum of noncapital and capital income.

\(^{33}\)Our solution of the KS-JEDC model matches well the results of the KS-Orig model in terms of wealth distribution.
### Table 5 Proportion of Total Wealth Held by Percentile (in percent) and the Marginal Propensity to Consume

<table>
<thead>
<tr>
<th>Micro Income Process</th>
<th>Friedman/Buffer Stock</th>
<th>KS-Orig</th>
<th>KS-JEDC</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Point Discount Factor†</td>
<td>Uniformly Distributed Discount Factors*</td>
<td>Our solution</td>
<td>U.S. Data*</td>
</tr>
<tr>
<td>Top 1%</td>
<td>11.6</td>
<td>25.7</td>
<td>3.0</td>
<td>2.3</td>
</tr>
<tr>
<td>Top 10%</td>
<td>39.0</td>
<td>65.3</td>
<td>19.0</td>
<td>17.8</td>
</tr>
<tr>
<td>Top 20%</td>
<td>55.2</td>
<td>80.4</td>
<td>35.0</td>
<td>31.9</td>
</tr>
<tr>
<td>Top 40%</td>
<td>75.8</td>
<td>92.5</td>
<td>55.5</td>
<td>55.5</td>
</tr>
<tr>
<td>Top 60%</td>
<td>88.9</td>
<td>97.1</td>
<td>74.7</td>
<td>74.7</td>
</tr>
<tr>
<td>Bottom 20%</td>
<td>3.0</td>
<td>0.7</td>
<td>9.9</td>
<td>9.9</td>
</tr>
<tr>
<td>Average Annual MPC</td>
<td>0.10</td>
<td>0.19</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>$K_t/Y_t$</td>
<td>10.3</td>
<td>10.3</td>
<td>10.3</td>
<td></td>
</tr>
</tbody>
</table>

Notes: †: $\beta = 0.9887$. *: $(\beta, \nabla) = (0.9866, 0.056)$, which implies $(\beta - \nabla, \beta + \nabla) = (0.9810, 0.9922)$. *: U.S. data is the SCF reported in Castaneda, Diaz-Gimenez, and Rios-Rull (2003).

### 2.5 Heterogeneous Impatience

As the simplest method to address this defect, we introduce heterogeneity in time preference factors: Each household is assumed to have an idiosyncratic (but fixed) time preference factor. We do not think of this assumption as only capturing actual variation in pure rates of time preference across people (though such variation surely exists). Instead, we view 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 (17), 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.

---

34 This differs 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.

35 We could of course model age effects directly, but it is precisely the inclusion of such realism that has made OLG models unpopular; they are too unwieldy to use for many practical research purposes and (perhaps more important) it is too difficult distill their mechanics into readily communicable insights. And our view is that, for macroeconomic analysis purposes, all that is gained in exchange for such complexity is a widening of
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 wealthiest 20, 40, and 60 percent of the population. Because these statistics have been targeted by other papers (e.g., Castaneda, Diaz-Gimenez, and Rios-Rull (2003)), we adopt a goal of matching them.

We replace the assumption that all households have the same time preference factor with an assumption that, for some $\gamma$, time preference factors are distributed uniformly in the population between $\beta - \gamma$ and $\beta + \gamma$. Then, using simulations, we search for the values of $\beta$ and $\gamma$ that match the fraction of wealth held by the top 20, 40, and 60 percent of the population, while at the same time matching the total aggregate wealth/output ratio from the perfect foresight model. Specifically, we solve the following minimization problem:

$$\min_{\beta, \gamma} \sum_{i=20,40,60} \left( w_i - \omega_i \right)^2$$

subject to the constraint that the aggregate wealth/output ratio matches the steady-state value from the perfect foresight model, where $w_i$ and $\omega_i$ are wealth held by the top $i$ percent in our model and in the data, respectively. The solution of this problem is $(\beta, \gamma) = (0.9866, 0.0056)$.

The introduction of time preference 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 formal 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 wealth points with only one free parameter. More impressive is the model’s match to locations in the wealth distribution that were not targeted; for example, the wealth shares of the top 10 percent and the top 1 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 has similar, or different, implications from the KS-JEDC model for serious macroeconomic questions like the reaction of aggregate consumption to an economic ‘stimulus’ payment.

---

The distribution of wealth-to-income ratios. We achieve the same effect more parsimoniously by incorporating discount factor heterogeneity.

36 Castaneda, Diaz-Gimenez, and Rios-Rull (2003) matched various wealth and income distribution statistics, including wealth held by the top 1, 5, 10, 20, 40, 60, 80 percent, and the Gini coefficient.

37 In estimating these parameter values, we approximate the uniform distribution by seven points ($\beta - 3\gamma/3.5$, $\beta - 2\gamma/3.5$, $\beta - \gamma/3.5$, $\beta$, $\beta + \gamma/3.5$, $\beta + 2\gamma/3.5$, $\beta + 3\gamma/3.5$). Increasing the number of points further does not notably change the results below.

38 As is clear from the minimization problem above, we are estimating two parameters ($\beta$ and $\gamma$). However, that estimation is subject to a constraint (matching the targeted aggregate wealth/income ratio) that effectively pins down one of the parameters ($\beta$), so effectively only $\gamma$ works to match the three wealth target points.
Table 6  Parameter Values for KS Aggregate Shocks

<table>
<thead>
<tr>
<th>Param</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta Z$</td>
<td>0.010</td>
</tr>
<tr>
<td>$u^g$</td>
<td>0.04</td>
</tr>
<tr>
<td>$u^b$</td>
<td>0.10</td>
</tr>
<tr>
<td>Agg Transition Probability</td>
<td>0.125</td>
</tr>
</tbody>
</table>

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 plan: Effective immediately, every household in the economy will receive a ‘stimulus check’ worth some modest amount $x$ (financed by a tax on unborn future generations).

The table shows that the immediate net MPC out of the stimulus payments in the preferred version of the model (with heterogeneous time preference factors) is roughly twice as large as in the version of our model with an identical (point) time preference factor. The MPC in our model is also roughly four to five times as large as that produced by our solution of the KS-JEDC model (0.05) or the perfect foresight partial equilibrium model (0.04).

3 Aggregate Shocks

In this section, we examine a model with an FBS household income process that also incorporates KS aggregate shocks. Krusell and Smith (1998) assumed that the level of aggregate productivity alternates between $Z_t = 1 + \Delta Z$ if the aggregate state is good and $Z_t = 1 - \Delta Z$ if it is bad; similarly, $L_t = 1 - u_t$ where $u_t = u^g$ if the state is good and $u_t = u^b$ if bad. (For reference, we reproduce their assumed parameter values in Table 6.)

The decision problem for an individual household in period $t$ can be written using normalized variables and the employment status $\iota_t$:

$$v(m_t, \iota_t; K_t, Z_t) = \max_{\{c_t\}} u(c_t) + \beta \mathbb{E}_t \left[ (T_{t+1} \psi_{t+1})^{1-\rho} v(m_{t+1}, \iota_t; K_{t+1}, Z_{t+1}) \right]$$

s.t.

$$a_t = m_t - c_t$$

$39$ This financing scheme, along with the lack of a bequest motive, eliminates any Ricardian offset that might otherwise occur.

$40$ The MPC’s that we calculate in the table are annual MPC’s given by $1 - (1 - \text{quarterly MPC})^4$ (recall again that the models in this paper are calibrated quarterly). We make this choice because most of the empirical literature that has attempted to estimate MPC’s has used annual micro data, and so the casual usage of the term ‘the MPC’ refers to the amount of extra spending that would occur over the course of a year in response to a one unit increase in resources.
\[
\begin{align*}
\alpha_t & \geq 0 \\
k_{t+1} & = \frac{a_t}{(\Psi T_{t+1} \psi_{t+1})} \\
m_{t+1} & = (\Gamma + r_{t+1}) k_{t+1} + y_{t+1} \\
r_{t+1} & = \alpha Z_{t+1} \left( \frac{K_{t+1}}{\bar{L}_{t+1}} \right)^{\alpha-1},
\end{align*}
\]

where

- the non-bold individual variables (lower-case variables except for \( \psi_t \)) are the bold (level) variables divided by \( X_t p_t \) (e.g., \( a_t = a_t / X_t p_t \), \( m_t = m_t / X_t p_t \)),
- \( \Gamma_{t+1} = X_{t+1} / X_t \),
- \( L_t = 1 - u_t \), and
- the income process is the same as in (9)-(13) but the employment transition process follows KS-JEDC.

There are more state variables in this version of the model than in the model with no aggregate shock: The aggregate variables \( Z_t \) and \( K_t \), and the household’s employment status \( \iota_t \) 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 Krusell and Smith (1998). See Appendix B for details about our solution method.

We now report the results of simulations, both for the model in which all households have the same time preference factor (“\( \beta \)-Point” model) and for the version with a uniform distribution of time preference factors (“\( \beta \)-Dist” model, with the parameters estimated in Section 2). Both models are solved with the KS aggregate shocks described above. Results using our solution of the KS-JEDC model (where \( \theta_t = 1 \) and \( \psi_t = 1 \) for all \( t \) but no death (\( D = 0 \)) are also reported for comparison.

3.1 Some Macroeconomic Statistics

Table 7 shows that aggregate statistics which we think useful for macroeconomic analysis: The serial correlation of consumption growth, and correlation between consumption growth, income growth, and interest rates at several frequencies. For comparison to our solution of the KS-JEDC model, the result of Maliar, Maliar, and Valli (2008), who report consumption growth correlation, is included in the table.\footnote{The minor difference between the results in Maliar, Maliar, and Valli (2008) and ours reflects approximation error in solving the consumption function.} We also report results for the representative agent model with the
Table 7  Aggregate Statistics with KS Aggregate Shocks

<table>
<thead>
<tr>
<th></th>
<th>Buffer Stock/Friedman</th>
<th>KS-JEDC</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β-Point</td>
<td>β-Dist</td>
<td>Our solution</td>
</tr>
<tr>
<td>( corr(\Delta \log C_t, \Delta \log C_{t-1}) )</td>
<td>0.20</td>
<td>0.08</td>
<td>0.26</td>
</tr>
<tr>
<td>( corr(\Delta \log C_t, \Delta \log Y_t) )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( corr(\Delta \log C_t, \Delta \log Y_{t-1}) )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( corr(\Delta \log C_t, \Delta \log Y_{t-2}) )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( corr(\Delta \log C_t, r_t) )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( corr(\Delta \log C_t, r_{t-1}) )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( corr(\Delta_4 \log C_t, \Delta_3 \log Y_t) )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( corr(\Delta_8 \log C_t, \Delta_3 \log Y_t) )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( K_t/Y_t )</td>
<td>10.3</td>
<td>10.3</td>
<td>10.3</td>
</tr>
</tbody>
</table>

Notes: \( \Delta_4 \) and \( \Delta_8 \) are one-year and two year growth rates, respectively.

KS aggregate income shock parameters (last column of Table 7), the results of which are similar to those of our solution of the KS-JEDC model.

The classic reference point for consumption growth measurement is the random walk model of Hall (1978), and the large literature that rejects the random walk proposition in favor of models that either contain ‘rule-of-thumb’ some consumers who set spending equal to income in every period (Campbell and Mankiw (1989)) or, more popular recently, models with habit formation or sticky expectations that imply serial correlation in consumption growth (see Carroll, Sommer, and Slacalek (2010) for evidence).

The KS-JEDC model produces a relatively high correlation coefficient \( corr(\Delta \log C_t, \Delta \log C_{t-1}) \), 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).\(^4\) Our β-Point and β-Dist models also imply positive \( corr(\Delta \log C_t, \Delta \log C_{t-1}) \), although not as high as that predicted by the KS-JEDC 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

\(^4\) 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 (2010)). This would imply that the models above do not reproduce stickiness in aggregate consumption growth well.
Table 8  Proportion of Total Wealth Held by Percentile (in percent)

<table>
<thead>
<tr>
<th>Micro Income Process</th>
<th>Friedman/Buffer Stock</th>
<th>KS-JEDC</th>
<th>KS-Orig</th>
<th>Hetero</th>
<th>US data</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-Point</td>
<td>β-Dist</td>
<td>Our solution</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top 1%</td>
<td>10.3</td>
<td>24.2</td>
<td>2.3</td>
<td>24.0</td>
<td>29.6</td>
</tr>
<tr>
<td>Top 10%</td>
<td>38.2</td>
<td>65.0</td>
<td>17.8</td>
<td>73.0</td>
<td>66.1</td>
</tr>
<tr>
<td>Top 20%</td>
<td>54.5</td>
<td>80.3</td>
<td>31.9</td>
<td>88.0</td>
<td>79.5</td>
</tr>
<tr>
<td>Top 40%</td>
<td>75.5</td>
<td>92.5</td>
<td>55.5</td>
<td>-</td>
<td>92.9</td>
</tr>
<tr>
<td>Top 60%</td>
<td>88.8</td>
<td>97.1</td>
<td>74.7</td>
<td>-</td>
<td>98.7</td>
</tr>
</tbody>
</table>

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

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-JEDC model, but after some investigation we determined that the KS-JEDC 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).

3.2 Wealth Distribution Statistics

Krusell and Smith themselves pointed out that an important failure of the baseline version of their original model was its inability to match the wealth distribution. The column labeled ‘KS-JEDC’ in Table 8 shows, unsurprisingly, that this failure is not remedied by the minor changes made to their model for the JEDC volume: The proportion of wealth held by the richest 1 percent of households in the KS-JEDC model is only 2.3 percent, compared to 29.6 percent in the U.S. data.

The model’s problems in matching the wealth distribution are not confined to the top. In fact, perhaps a bigger problem is that the 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 model. We have calculated, for example, that in steady state about 80 percent of all households in the KS-JEDC model have wealth between 0.5 times mean wealth and 1.5 times mean wealth; the corresponding fraction ranges from only 20 to 25 percent in the SCF data from 1992-2004.

To address this shortcoming of the baseline version of their model, Krusell and Smith examined a variant of their model that incorporated a form of discount rate
heterogeneity. 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-Orig Hetero” column in the table).

Even without discount factor heterogeneity, our \( \beta \)-Point model also increases inequality in the wealth distribution relative to the KS-JEDC model; the proportions held by the top 1 percent and the top 10 percent of wealth-holders are about 10 percent and nearly 40 percent, respectively (first column). However, these are still much lower than the ratios in the U.S. data.

But our \( \beta \)-Dist model does a much better job in reproducing the U.S. wealth distribution (second column in the table and Figure 1). Although even our \( \beta \)-Dist model still underpredicts wealth holdings at the top 1 percent a bit, it closely matches the U.S. distribution in the middle part, substantially outperforming even the version of the original KS model that incorporated heterogeneity.

3.3 The Aggregate Marginal Propensity to Consume

The distribution of wealth has implications for the aggregate MPC out of transitory shocks to income. To see why, Figure 2 plots our \( \beta \)-Point 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.

---

43Specifically, 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.
The figure shows that MPC is higher when the level of normalized cash on hand is lower and vice versa, implying that the average MPC is higher when a larger fraction of households has less (normalized) cash on hand.

There are many more households with little wealth in our $\beta$-Point model than in the KS-JEDC model, as illustrated by comparison of the short-dashing and the long-dashing lines in Figure 1. Our greater concentration of wealth at the bottom should produce a higher average MPC in our $\beta$-Point model, given the concave consumption function (Figure 2).

Indeed, the average MPC out of transitory income in our $\beta$-Point model is 0.10 in annual terms, about double the value in the KS-JEDC model (0.05) (Table 9). 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 $\beta$-Dist model (second column of the table) produces a higher average MPC (0.2) which is at the lower bound of this range, since in the $\beta$-Dist model there are even more households with less wealth (Figure 1). The results of the $\beta$-Dist model do not change much with an alternative set of income variance parameters. For example, even if we pick much lower variances of income shocks (annual $\sigma^2_\psi = \sigma^2_\theta = 0.01$), which are at the lower bound of empirical estimates (Table 2), the average MPC is only slightly lower (0.18).

---

44 Consumption functions of the KS-JEDC and $\beta$-Dist models have a similar form.
45 The average MPC calculated here can be interpreted as how much is spent on average when one dollar is disbursed to all households.
### Table 9: Average Marginal Propensity to Consume in Annual Terms

<table>
<thead>
<tr>
<th>Micro Income Process</th>
<th>Friedman/Buffer Stock</th>
<th>KS-JEDC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β-Point</td>
<td>β-Dist</td>
</tr>
<tr>
<td>Overall average</td>
<td>0.10</td>
<td>0.20</td>
</tr>
<tr>
<td>By wealth/permanent income ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top 1%</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Top 10%</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Top 20%</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Top 40%</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Top 60%</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>Bottom 1/2</td>
<td>0.13</td>
<td>0.31</td>
</tr>
<tr>
<td>By income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top 1%</td>
<td>0.08</td>
<td>0.13</td>
</tr>
<tr>
<td>Top 10%</td>
<td>0.08</td>
<td>0.14</td>
</tr>
<tr>
<td>Top 20%</td>
<td>0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>Top 40%</td>
<td>0.11</td>
<td>0.16</td>
</tr>
<tr>
<td>Top 60%</td>
<td>0.11</td>
<td>0.17</td>
</tr>
<tr>
<td>Bottom 1/2</td>
<td>0.08</td>
<td>0.22</td>
</tr>
<tr>
<td>By employment status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>0.09</td>
<td>0.18</td>
</tr>
<tr>
<td>Unemployed</td>
<td>0.17</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Notes: Annual MPC is calculated by \(1 - (1 - \text{quarterly MPC})^4\).

MPC’s are generally higher among low wealth/income households and the unemployed in both our β-Point and β-Dist models (rest of the rows in Table 9). 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 stimulate consumption, it makes much more sense to target them to low-wealth households than to distribute them uniformly to the population as a whole.
Table 10 Parameter Values for More Plausible Aggregate Process

<table>
<thead>
<tr>
<th>Description</th>
<th>Param</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance of Log $\Psi_t$</td>
<td>$\sigma^2_{\Psi}$</td>
<td>0.00004</td>
</tr>
<tr>
<td>Variance of Log $\Xi_t$</td>
<td>$\sigma^2_{\Xi}$</td>
<td>0.00001</td>
</tr>
</tbody>
</table>

4 A More Plausible and More Tractable Aggregate Process

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

The aggregate production function is the same as equation (1), but following Carroll and Slacalek (2008) the aggregate state (good or bad) does not exist in this model ($Z_t = 1$). Aggregate productivity is instead captured by $L_t$. Specifically, $L_t = P_t \Xi_t$; $P_t$ is aggregate permanent productivity, where $P_{t+1} = P_t \Psi_{t+1}$; $\Psi_{t+1}$ is the aggregate permanent shock; and $\Xi_t$ is the aggregate transitory shock (note that $\Psi$ is the capitalized version of the Greek letter $\psi$ used for the idiosyncratic permanent shock; similarly (though less obviously), $\Xi$ is the capitalized $\xi$). Both $\Psi_t$ and $\Xi_t$ 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 10).

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 elimination of the 'good' and 'bad' aggregate states reduces the number of state variables to two ($m_t$ and $K_t$) 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 $K_t$, which can be obtained by following essentially the same method as described in the Appendix B.

The performance of the $\beta$-Dist model with this (Carroll/Slacalek) aggregate process is similar to the $\beta$-Dist model under the KS aggregate process, using the same set of parameter values.\footnote{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 confirms that this version of the $\beta$-Dist
model can replicate closely the U.S. wealth distribution. The MPCs are also close those reported earlier (Table 11).

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.
Table 11  Average Marginal Propensity to Consume in Annual Terms in
$\beta$-Dist Model under ‘Plausible’ Aggregate Process

<table>
<thead>
<tr>
<th></th>
<th>Carroll/Slacalek Aggregate Process</th>
<th>KS Aggregate Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall average</td>
<td>0.17</td>
<td>0.20</td>
</tr>
<tr>
<td>By wealth/permanent income ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top 1%</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Top 10%</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Top 20%</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Top 40%</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Top 60%</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>Bottom 1/2</td>
<td>0.27</td>
<td>0.31</td>
</tr>
<tr>
<td>By income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top 1%</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>Top 10%</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td>Top 20%</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Top 40%</td>
<td>0.17</td>
<td>0.16</td>
</tr>
<tr>
<td>Top 60%</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td>Bottom 1/2</td>
<td>0.17</td>
<td>0.22</td>
</tr>
<tr>
<td>By employment status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>0.16</td>
<td>0.18</td>
</tr>
<tr>
<td>Unemployed</td>
<td>0.33</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Notes: The Friedman/Buffer Stock income process is assumed. Annual MPC is calculated by $1 – (1 – \text{quarterly MPC})^4$. 

25
Appendix

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

The evolution of the square of $p$ is given by

\[
\begin{align*}
    p_{t+1,i} &= p_{t,i} \psi_{t+1,i}(1 - d_{t+1,i}) + d_{t+1,i} \\
    p_{t+1,i}^2 &= (p_{t,i} \psi_{t+1,i}(1 - d_{t+1,i}))^2 + 2p_{t,i} \psi_{t+1,i} d_{t+1,i}(1 - d_{t+1,i}) + d_{t+1,i}^2,
\end{align*}
\]

where $d_{t+1,i} = 1$ if household $i$ dies.

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

\[
\mathbb{E}_t[p_{t+1,i}^2] = \mathbb{E}_t[(p_{t,i} \psi_{t+1,i}(1 - d_{t+1,i}))^2] + D
\]

\[
= p_{t,i}^2 \mathbb{D}_t \mathbb{E}[^2] + D,
\]

so we have

\[
\mathbb{M}[p_{t+1}^2] = \mathbb{M}[p_{t}^2] \mathbb{D}_t \mathbb{E}[\psi^2] + D,
\]

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

\[
\begin{align*}
    \mathbb{M}[p^2] &= D + \mathbb{D}_t \mathbb{E}[\psi^2] \mathbb{M}[p^2] \\
    \mathbb{M}[p^2] &= \left( \frac{D}{1 - \mathbb{D}_t \mathbb{E}[\psi^2]} \right).
\end{align*}
\]

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 Section 3:

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-JEDC model according to this categorization. All articles in the table except Kim, Kim, and Kollmann (2010) solve the exact KS-JEDC model using various methods.\footnote{Kim, Kim, and Kollmann (2010) modified the form of the utility function.}

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
Table 12  Method of Solving KS-JEDC Model

<table>
<thead>
<tr>
<th>Method of Solving</th>
<th>Authors and Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relying on simulation</td>
<td>Young (2010)</td>
</tr>
<tr>
<td></td>
<td>Den Haan (2010b)</td>
</tr>
<tr>
<td>(In principle) not relying on simulation</td>
<td>Algan, Allais, and Haan (2008)</td>
</tr>
<tr>
<td></td>
<td>Reiter (2010)</td>
</tr>
</tbody>
</table>

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 (2010) and Den Haan (2010b)'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,\(^{48}\) 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 (2010) use a perturbation method that linearizes the system. The problem with this method is that they are unable to solve the exact same KS-JEDC 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.

B.2 Our Algorithm

Our algorithm to solve the problem in Section 3 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 \(K_t\) is sufficient to predict per

\(^{48}\)Simulation plays a part in Algan, Allais, and Haan (2008)'s method (they use simulation to find the function).
capita capital tomorrow ($K_{t+1}$). 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 $b = (b_0, b_1, b'_0, b'_1)$:

$$\log K_{t+1} = b_0 + b_1 \log K_t$$

if the aggregate state in period $t$ is good ($Z_t = 1 + \Delta Z$), and

$$\log K_{t+1} = b'_0 + b'_1 \log K_t$$

if the aggregate state is bad ($Z_t = 1 - \Delta Z$)

2. Simulate the economy populated by 7,000 households (which experiments determined is enough to suppress idiosyncratic noise) for 1,100 periods (following Maliar, Maliar, and Valli (2010)). 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 $Z_t = 1 + \Delta Z$ (the aggregate state is good). (The steady state level of $k_{t,i}$ is calculated by $ar{k} = (\alpha \beta / (1 - \beta \gamma))^{1/(1-\alpha)}$. With $k_{t,i} = 38.0$ for all $i$, $K_t = 41.2$.) 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 $\beta$-Dist model. (Our experiments found that we can reach the solution faster with $\eta = 1/2$.)

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

From the second iteration and thereafter, we use the terminal distribution of wealth (and permanent component of income ($p$)) in the previous iteration as the initial one. For the case of the $\beta$-Dist model, the number of households is multiplied by 10 in the final two (or three) iterations to reduce cross-sectional simulation error.

---

49 In our analysis below, the process is iterated until the difference between each estimate ($b_0, b_1, b'_0, b'_1$) and its previous value is smaller than 1 percent.

50 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.
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.\(^{51}\)

### B.4 Estimated Law of Motion

When we simulate the \( \beta \)-Point model in Section 3, the estimated law of motion is

\[
\log K_{t+1} = 0.141 + 0.962 \log K_t
\]

if the aggregate state is good, and

\[
\log K_{t+1} = 0.123 + 0.966 \log K_t
\]

\(^{51}\)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.
if the aggregate state is bad.\textsuperscript{52}

In the case of the $\beta$-Dist model, the estimated law of motion is

$$\log K_{t+1} = 0.154 + 0.959 \log K_t$$

if the aggregate state is good, and

$$\log K_{t+1} = 0.141 + 0.961 \log K_t$$

if the aggregate state is bad.\textsuperscript{53}

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

$$\log K_{t+1} = 0.138 + 0.963 \log K_t$$ \hspace{1cm} (19)

if the aggregate state is good, and

$$\log K_{t+1} = 0.122 + 0.966 \log K_t$$ \hspace{1cm} (20)

if the aggregate state is bad.\textsuperscript{54} The coefficients in (19) and (20) are very close to those estimated in other articles that examine the KS-JEDC model (e.g., Maliar, Maliar, and Valli (2010)).

\* \* Experiment to Understand Sticky Consumption Growth

Although $\text{corr}(\Delta \log C_t, \Delta \log C_{t-1})$ 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 $\text{corr}(\Delta \log C_t, \Delta \log C_{t-1})$.

Previous studies that have examined KS type models have not investigated this issue. Using the KS-JEDC 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.\textsuperscript{55}

Figure 4 plots $\Delta \log C_t$ 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 $\Delta \log C_t$ is very persistent (it is negative in the bad state and positive in the good state), resulting in a relatively high $\text{corr}(\Delta \log C_t, \Delta \log C_{t-1})$.

\textsuperscript{52}$R^2$ 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 (2010a) discusses details).

\textsuperscript{53}$R^2$ is greater than 0.9999 for both states.

\textsuperscript{54}$R^2$ is greater than 0.99999 for both states.

\textsuperscript{55}Because one state switches to another with a probability of 0.125, the average length of each state is eight quarters in typical simulation.
It is easy to understand that $\Delta \log C_t$ is higher when the state is good (and vice versa) given the following facts:

- A first order approximation of the Euler equation yields:
  \[
  \Delta \log C_t \approx \rho^{-1}(r_t - (1 - \beta + \delta)) + \varepsilon_t, \tag{21}
  \]
  where $\rho$ is the coefficient of relative risk aversion, $r_t$ is the interest rate, $\beta$ is the time preference factor, $\delta$ is the depreciation rate, and $E_{t-1}[\varepsilon_t] = 0$. Indeed, when we conduct an IV regression of equation (21) using the data that produced Table 7,\textsuperscript{56} the estimate of $\rho^{-1}$ is 0.95 (with a standard deviation of 0.08) and close to the actual value of $\rho^{-1} (= 1)$, which suggests that the linear approximation (21) is largely valid.

- When the state is good, $r_t = \alpha Z_t (K_t/L_t)^{\alpha-1}$ (from (18)) is higher because $Z_t$ (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 $corr(\Delta \log C_t, \Delta \log C_{t-1})$) for the same mechanisms as in the experiment above.

In sum, a relatively high $corr(\Delta \log C_t, \Delta \log C_{t-1})$ in the KS-JEDC 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 $\varepsilon_t$ is the error term in (21)), which measures the autocorrelation after the effects of the interest rate are removed, is much lower than $corr(\Delta \log C_t, \Delta \log C_{t-1})$ and only 0.01.\textsuperscript{57}

\section*{D Transition Matrix Method}

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

\subsection*{D.1 Method}

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

\textsuperscript{56}We use $r_{t-1}$ as an instrument of $r_t$.

\textsuperscript{57}When the AR1 coefficient of $\varepsilon_t$ 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 4  Dynamics of $\Delta \log C_t$ in KS-JEDC Model

Figure 5  Dynamics of $r_t$ in KS-JEDC Model
\( d = Td \). 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.

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 two parameter values (\( \beta \) and \( \nabla \)) 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.\(^{58}\)

The estimates of \( \bar{\beta} - \nabla \) and \( \bar{\beta} + \nabla \) are 0.9837 and 0.9900, respectively. The interval \( (\bar{\beta} + \nabla) - (\bar{\beta} - \nabla) = 2\nabla \) is narrower than the estimate in Section 2 (Table 5). Table 13 compares the wealth/(quarterly) permanent income ratios produced by the \( \beta \)-Dist 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 \( \beta \)-Dist model’s ability to match the middle part of the distribution.

\(^{58}\)Because the SCF reports annual “normal” income, it is divided by 4 when calculating the wealth/(quarterly)permanent income ratios.
Table 13  Wealth/(Quarterly) Permanent Income Ratio

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 10%</td>
<td>35.7</td>
<td>22.9</td>
<td>22.9</td>
<td>29.3</td>
<td>30.9</td>
<td>29.6</td>
</tr>
<tr>
<td>Top 20%</td>
<td>17.1</td>
<td>14.1</td>
<td>14.1</td>
<td>17.3</td>
<td>18.5</td>
<td>18.5</td>
</tr>
<tr>
<td>Top 40%</td>
<td>6.9</td>
<td>7.0</td>
<td>6.8</td>
<td>8.0</td>
<td>8.7</td>
<td>8.8</td>
</tr>
<tr>
<td>Top 60%</td>
<td>3.8</td>
<td>3.1</td>
<td>3.2</td>
<td>3.4</td>
<td>3.8</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Figure 6  Distribution of Wealth/Permanent Income Ratio
References


