Macroeconomics and Household Heterogeneity*

Dirk Krueger
University of Pennsylvania, CEPR, CFS, NBER and Netspar

Kurt Mitman
IIES, Stockholm University and CEPR

Fabrizio Perri
Federal Reserve Bank of Minneapolis, CEPR and NBER

February 15, 2016

Abstract

Using the great recession as a case study, this paper will survey models of the macro economy with household heterogeneity and aggregate fluctuations. Our main focus is methodological: we wish to determine under what conditions cross-sectional household heterogeneity is important for the macroeconomic response to a business cycle shock, and what mechanisms are suitable for generating an empirically plausible model income and wealth distribution in the first place. We also investigate the role social insurance policies (such as unemployment insurance) play for shaping the cross-sectional income and wealth distribution, and through it, the dynamics of business cycles. Our main conclusion is that the wealth distribution can matter for the macroeconomic response to business cycle shocks if (and only if) it features sufficiently many households with very little net worth, as is empirically observed in the wealth distribution for the U.S.

Keywords: Recessions, Wealth Inequality, Social Insurance

JEL Classifications: E21, E32, J65

*We wish to thank the editors John Taylor and Harald Uhlig, discussants of the paper Claudio Michelacci, Tony Smith and Chris Tonetti as well as seminar participants at the Stanford Handbook for Macroeconomics Conference, Sveriges Riksbank-EABCN Conference on Inequality & Macroeconomics, the Workshop on Macroeconomic Dynamics (EIEF), the Econometric Society Wolrd Congress in Montreal as well as the Bank of Italy, the Bank of Norway, the ECB, NYU Abu Dhabi and Warwick for many useful comments and Simone Civale for outstanding research assistance. Krueger gratefully acknowledges financial support from the NSF under grants SES 1123547 and SES 1326781.

1 Introduction

How important is household heterogeneity for understanding the macro economy? The objective of this paper is to give a quantitative answer to a narrow version of this broad question.[1] Specifically, we narrow the focus of this question along two dimensions. First, in its empirical context in that we mainly focus on a specific macroeconomic event, namely the Great recession of 2007-2009. Second, we focus on a specific dimension of household heterogeneity, namely that in wealth, and its associated correlations with and consequences for the cross-sectional inequality in earnings, disposable income and consumption expenditures.

The great recession was the largest negative macroeconomic downturn since World War II. The initial decline in economic activity was deep, occurred world-wide, impacted all macroeconomic aggregates (and notably, private aggregate consumption and employment), and the recovery has been slow. Is the cross-sectional distribution of wealth an important determinant of the dynamics of the initial downturn and the ensuing recovery? That is, does household heterogeneity matter in terms of aggregate economic activity (as measured by output and labor input), its composition between consumption and investment, and, eventually, the cross-sectional distribution of consumption and welfare?

To address these questions *empirically*, we make use of recent waves of the Panel Study of Income Dynamics (PSID) which provides household level panel data on earnings, income, consumption expenditures and wealth for the U.S. To answer these questions *theoretically and quantitatively* we study various versions of the canonical real business cycle model with aggregate technology shocks and ex-post household heterogeneity induced by the realization of uninsurable idiosyncratic labor earnings shocks, as in Krusell and Smith (1998). In the model, a recession is associated with lower aggregate wages and higher unemployment (i.e a larger share of households with low labor income). The main empirical and model-based focus of the paper is on the dynamics of macroeconomic variables (and specifically, aggregate consumption, investment and output) in response to such a business cycle shock, and specifically on the conditions under which the degree of wealth inequality plays a quantitatively important role for shaping this response. We also study how stylized social insurance programs, and unemployment insurance more specifically, shape the cross-sectional distribution of wealth and welfare, and how it impacts the recovery of the aggregate economy after a great-recession like event.

We proceed in four steps. First, we make use of the PSID earnings, income, consumption and wealth data to document three sets of cross-sectional (or inequality) facts.

[1] In this paper we focus on household heterogeneity. A sizeable literature has investigated similar questions in models with firm heterogeneity. Representative contributions from this literature include Khan and Thomas (2008) and Bachmann, Caballero and Engel (2013). We abstract from firm heterogeneity in this chapter, but note that the methodological

First we summarize the key features of the joint income, wealth and consumption distribution prior to the great recession (i.e. for the year 2006). Second we show how this joint distribution has changed during the great recession (in the 2006-2010 interval) and third, we use the panel dimension of the data to investigate how individual households fared and adjusted their consumption-savings behavior during this time interval. The purpose of this empirical analysis is too-fold. First, we believe the facts are interesting in their won right as they characterize the distributional consequences of the great recession. Second, the facts serve as important moments for the evaluation of the different versions of the quantitative heterogeneous household model we study next.

Therefore, in a second step, we construct, calibrate and compute various versions of the canonical Krusell-Smith (1998) model and study its cross-sectional and dynamic properties. We first revisit the well-known finding that idiosyncratic income risk and incomplete financial markets alone are insufficient to generate a sufficiently dispersed model-based cross-sectional wealth distribution. The problem is two-fold: the very wealthy are not nearly wealthy enough and households at the bottom of the wealth distribution have too much net worth in the model, relative to the data. We then argue that it is this last finding that implies an aggregate consumption response to a negative aggregate technology shock essentially identical to the one in the representative agent model. We then study extensions of the model in which preference heterogeneity and a (somewhat rudimentary) life cycle structure interacts with the presence of unemployment insurance and social security to deliver a wealth distribution in which the bottom 40% of households hold very little net worth and the top 20% own 80% of wealth, as in the data. We then show that in these economies the decline in aggregate consumption is substantially larger than in the representative agent economy. This is primarily due to the fact that now these economies are populated by more wealth-poor households whose consumption responds strongly to the aggregate shock, both for those households that experience a transition from employment to unemployment, but also households that have not yet lost their job, but understand they are facing a potentially long lasting recession with elevated unemployment *risk*. The more severe consumption recessions in economies with larger wealth inequality go hand in hand with a lesser investment collapse and thus a faster recovery from the great recession, although this effect is quantitatively small.

In light of the previous finding that larger wealth inequality, and more specifically, a large fraction of wealth poor households, is an important contributor to an aggregate consumption collapse in the great recession, the next question this chapter addresses is whether social insurance stabilizing income of the unemployed avoids

Third, in the context of the high wealth inequality economies we determine whether the presence (and size) of a public unemployment insurance is important for the dynamics of the economy in response to an aggregate shock. The answer to this last question is twofold: a) for a *given wealth distribution* (consistent with that in the data) the absence of a sizable unemployment insurance system implies a significantly stronger negative consumption response (and thus a weaker investment response and speedier recovery); b) forward looking households respond to lower public insurance by increasing their precautionary saving, the resulting wealth distribution has fewer people with zero assets, which in turn softens the aggregate consumption decline (and slows the recovery).

In the previous model, the wealth distribution has a potentially large effect on the *distribution* of aggregate output between consumption and investment, but not on output itself. In a final step we therefor study an economy with a New Keynesian flavor where an aggregate demand externality generates an endogenous feedback effect from private consumption to total factor productivity and thus aggregate output. In this model social insurance policies might not only be beneficial in providing public insurance, but also can also serve a potentially positive role for stabilizing aggregate output.

The paper is organized as follows. Section 2 documents key dimensions of heterogeneity among U.S. households, prior to and during the Great Recession. Sections 3 and 4 presents our benchmark real business cycle model with household heterogeneity and discuss how we calibrate it. Section 5 studies to what extent the benchmark model is consistent with the facts presented in section 2, and section 6 documents how the response of the aggregate economy to a great recession shock depends on the cross-sectional wealth distribution. It also measures the welfare cost of the great recession and shows it to be substantially heterogeneous across the population. In section 7 we augment the model with endogenous labor supply choices and demand externalities in order to investigate the importance of cross-sectional wealth heterogeneity for the dynamics of aggregate output (and not just its distribution between consumption and investment, the main focus of section 6). Section 8 discusses the vast literature on this topic from the perspective of our own findings and section 9 concludes. The appendix contains details about the construction of the empirical facts as well as the computational algorithm used in the paper.

2 The Great Recession: a Heterogeneous Household Perspective

In this section we present the basic facts about the cross-sectional distribution of earnings, income, consumption and wealth before and during the great recession. The main data set we employ is the Panel Survey of Income Dynamics (PSID) for the years 2004, 2006, 2008 and 2010. This dataset has two key advantages for the purpose of this study. First, it contains information about household earnings, income, a broad and comprehensive measure of consumption expenditures and wealth for a sample of households representative of the US population. Second, it has a panel dimension so we can, in the same data set, measure both the key dimensions of cross-sectional household heterogeneity as well as investigate how different groups in the income and wealth distribution have fared during the Great Recession[2].

The purpose of this section is two-fold: first, it presents some basic stylized cross-sectional facts that motivative the construction of the type of models we will study in the remainder of this chapter. But second, it provides simple and direct evidence for the importance of household heterogeneity for macroeconomic questions. If, as we will document, there are significant differences in behavior (for example, along the consumption and savings margin) across different groups of the earnings and wealth distribution during the Great Recession, then keeping track of the cross-sectional earnings and wealth distribution and understanding their dynamics is important for analyzing the unfolding of the Great Recession from a macroeconomic and distributional perspective.

2.1 Aggregates

We start our analysis by comparing the evolution of basic U.S. macroeconomic aggregates from the National Income and Product Accounts (NIPA) with the aggregates for the same variables obtained from the PSID. In Figure 1 below we compare trends in aggregate Per Capita Disposable Income (panel A) and Per Capita Consumption Expenditures (panel B) from the Bureau of Economic Analysis (BEA) with the corresponding series obtained aggregating household level in PSID, for the years 2004 through 2010, the last available data point for PSID.[3]

The main conclusion we draw from figure 1 is that both NIPA and the PSID paint the same qualitative picture of the period prior to, and during the great recession. Both disposable income and consumption expenditures experience a slowdown, which is somewhat more pronounced in the PSID. Furthermore, PSID consumption expenditure data also display a much weaker aggregate recovery than what is observed in the NIPA data.[4]

2.2 Inequality before the Great Recession

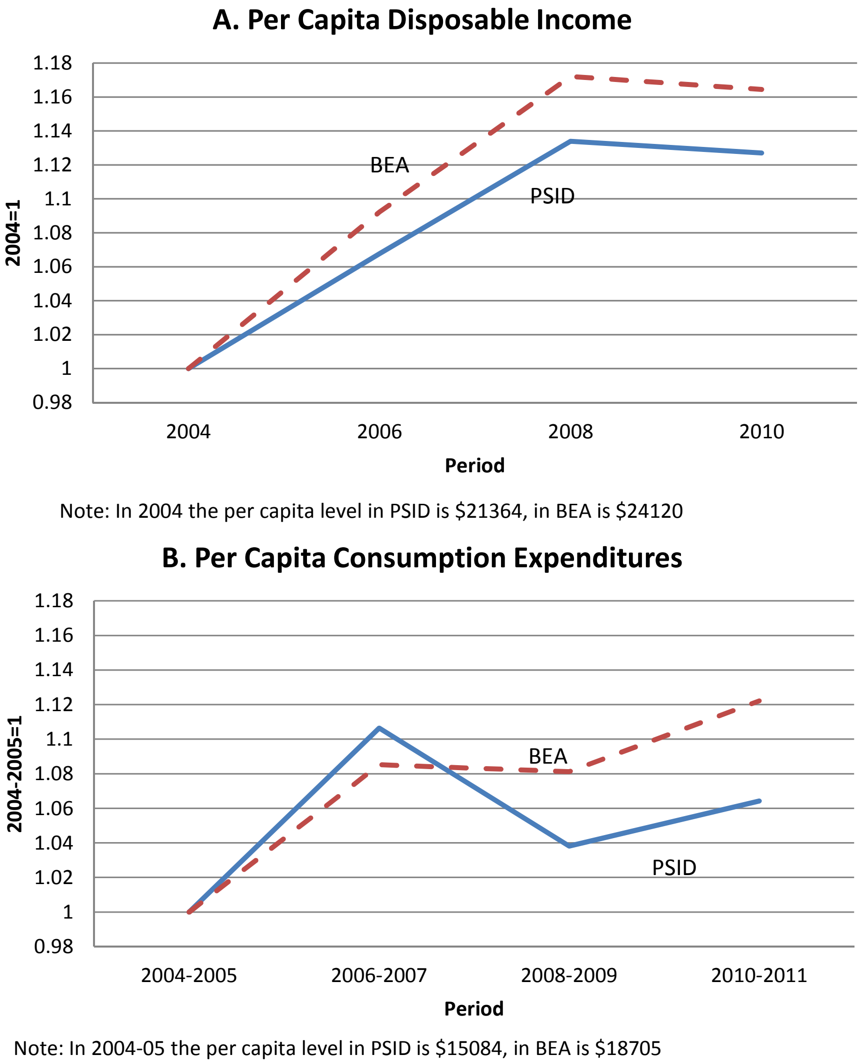
In this section we document basic inequality facts in the United States for the year 2006, just before the Great Recession hit the economy. Since the Great Recession greatly impacted households in the labor market and our models below focus on labor market earnings risk, we focus on households which have at least one member of age between

[2] See Smith and Tonetti (2014) for a novel method of constructing an income-consumption panel using both the PSID and Consumer Expenditure Survey (CEX).

[3] In appendix A we describe in detail how the two series are constructed from the PSID.

[4] As Heathcote, Perri and Violante (2010) document, this discrepancy between macro data and aggregated micro data is also observed in previous recoveries from U.S. recessions.

Figure 1: The Great Recession in the NIPA and in the PSID Data



22 and 60. Table 1 reports statistics that characterize, for this group of households, the distributions of earnings (which we define to include all sources of labor income plus transfers minus tax liabilities), disposable income (which is earnings plus unemployment benefits, plus income from capital, including rental equivalent income of the main residence of the household), consumption expenditures (which includes all expenditure categories reported by PSID plus the rental equivalent of the main residence).

Table 1 reports, for each variable (earnings, disposable income, consumption expenditures and net worth), the cross sectional average (in 2006 dollars), as well as the share of the total value held by each of the five quintiles of the corresponding distribution, and finally the share of each variable held by the households between the 90th and 95th percentile, between the 95th and 99th percentile and by those in the top 1% of the respective distribution. The last column of the table reports the same statistics for net worth computed from a different data set, the 2007 Survey of Consumer Finances

Table 1: Means and Marginal Distributions in 2006

	Variable				
	Earn.	Disp Y	Cons. Exp.	NetW	NetW (SCF 07)
Mean (2006$)	52,783	62,549	43,980	291,616	497,747
% Share held by:					
*Q*1	3.4	4.3	5.7	-1.2	-0.3
*Q*2	9.7	9.7	10.7	0.7	0.9
*Q*3	15.2	15.1	15.6	4.1	4.2
*Q*4	22.8	22.9	22.5	13.3	11.8
*Q*5	48.7	48.0	45.5	83.1	83.4
90 – 95	11.1	10.8	10.4	14.0	11.1
95 – 99	13.3	13.1	11.4	23.2	25.6
Top 1%	7.8	7.8	8.0	30.2	34.1
Sample Size		6442			14725

(SCF), which is the most commonly used dataset for studying the U.S. wealth distribution.

The table reveals features that are typical of distributions of resources across households in developed economies. Earnings and disposable income are both quite concentrated, with the bottom quintiles of the respective distributions holding shares smaller than 5% (3.4% and 4.3% to be exact) and the top quantiles holding almost 50% (48.7% and 48% to be precise). The distributions of earnings and disposable income look quite similar, since for the households in our sample (aged 22 to 60) capital income is a fairly small share of total disposable income (constituting only roughly 1/6 of disposable income).[5] Consumption expenditures are less unequally distributed, with the bottom quintile accounting for a bigger fraction (5.7%) of total expenditures.

Net worth is by far the most concentrated variable, especially at the top of the distribution. The bottom 40% of households hold essentially no wealth at all, whereas the top quintile owns 83% of all wealth, and the top 10% holds around 70% of total wealth. Comparing the last two columns demonstrates that, although the average level of wealth in the PSID is substantially lower than in the SCF, the distribution of wealth across the five quintiles lines up quite closely between the two data sets, suggesting that the potential underreporting or mis-measurement of wealth in the PSID might affect the overall amount of wealth measured in this data set, but not the cross-sectional distribution which is remarkably comparable to that in the SCF.

[5] Recall that earnings are net of taxes and include government transfers already as well. **Need to reconcile income data at top with Piketty and Saez who report a larger share of the top 1%. Their income measure is different though and they look at entire population.**

Although the marginal distributions of earnings, income and wealth are interesting in their own right, the more relevant object for our purposes is the joint distribution of wealth, earnings, disposable income and consumption expenditures.[6] To document the salient features of this joint distribution we divide the households in our 2006 PSID sample into net worth quintiles, and then for each *net worth quintile* we report, in Table 2, the share of the relevant variable held by that quintile.

Table 2: Earnings, Disposable Income and Expenditures by Net Worth in 2006

	% Share of:			% Expend. Rate	
Quintile(NetW)	Earn.	Disp Y	Expend.	Earn.	Disp Y
$Q1$	9.6	8.6	11.3	98.1	92.2
$Q2$	12.3	10.7	12.4	84.1	81.3
$Q3$	18.0	16.6	16.8	77.9	70.9
$Q4$	22.8	22.6	22.4	81.6	69.6
$Q5$	37.2	41.4	37.2	83.0	63.1
	Correlation with net worth				
	0.26	0.39	0.16		

The table shows two important features of the data. The first is that, perhaps not surprisingly, households with higher net worth tend to have higher earnings and higher disposable incomes. The last row of the table shows the extend to which these two variables are positively correlated with net worth. The second observation is that consumption expenditures are also positively correlated with net worth, but less so than the two income variables. The reason is that, as can be seen in the last two columns of the table, the lower is net worth, the higher the consumption rate.[7] The differences in the consumption rates across wealth quintiles are economically significant: for example, between the bottom and the top wealth quintile the differences in the consumption rates range between 20% and 30%.

Another way to look at the same issue is to notice from tables 1 and 2 that the households in the bottom two wealth quintiles, although they basically hold no wealth (see table 1 above), are responsible for $11.3\% + 12.4\% = 23.7\%$ of total consumption expenditures, making this group quantitatively consequential for aggregate consumption dynamics. The differences across groups delineated by wealth constitute *prima-facie* evidence that the shape of the wealth distribution *could* matter for the aggregate consumption response to macroeconomic shocks such as the ones responsible for the

6 The class of models we will construct below will have wealth -in addition to current earnings- as the crucial state variable, and thus we stress the correlation of net worth with earnings, income and especially consumption here.

7 We measure the consumption rate by computing total consumption expenditures in for a specific wealth quintile, and dividing it by total earnings (or disposable income) in that wealth quintile.

Great Recession. In the next subsection we will go beyond household heterogeneity at a given point in time and empirically evaluate how, during the Great Recession, consumption and saving changed differentially for households across the wealth distribution.

2.3 The Great Recession across the Income and Wealth Distributions

In tables 3 and 4 we report for households in each of the five wealth quintiles of the net worth distribution, the changes in net worth, earnings, disposable income, consumption expenditures and consumption expenditure rates.[8] Table 3 reports the changes for the 2004-2006 period, to establish a benchmark from a pre-recession period, whereas table 4 reports the changes for the 2006-2010 period, which covers the whole recession.[9]

Table 3: Changes (2004-2006) in Selected Variables across PSID Net Worth

		%			% Expend. Rate (pp)	
NW Q	Net Worth[a]	Earn.	Disp Y	Cons. Exp.	Earn.	Disp Y
$Q1$	27.0k (+∞)	11.6	14.3	12.7	1.0	-1.4
$Q2$	40.0k (140%)	12.0	13.8	11.8	-0.0	-1.5
$Q3$	40.8k (50%)	9.1	9.4	18.4	7.3	6.3
$Q4$	60.7k (28.2%)	9.9	10.8	8.8	-1.0	-1.3
$Q5$	266.1k (21.5%)	2.9	3.4	5.9	2.2	1.4

[a] *In 000's of dollars. Percentage change in parenthesis*

Both tables reveal a number of interesting facts that we want to highlight. From the first column of table 3 notice that all groups of households experienced solid growth in net worth between 2004 and 2006, mainly due to the rapid growth in asset prices (stock prices and especially real estate prices) during this period, with low wealth households experiencing the strongest growth in wealth (but of course from very low levels, see again table 1). Turning to earnings and disposable income (second and third column of table 3), we observe that households originally at the bottom of the wealth distribution experience faster income growth than those in higher wealth quintiles. This is most likely due to mean reversion in income: low wealth households are also low income households, and on average low income households experience faster income growth.

8 We construct these changes as follows: we keep the identity of the households fixed; for example, to compute the 2004-2006 change in net worth for $Q1$ of the net worth distribution we select all households in the bottom quintile of the wealth distribution in 2004, compute their average net worth (or earnings, income or consumption) in 2004 and 2006, and then calculate the percent difference between the two averages. For the consumption expenditure rates we report percentage point differences.

9 In tables A2 and A3 in the data appendix we report the changes for the 2006-2008 time period and for the 2008-2010 separately.

Finally, expenditure growth roughly tracked the growth of income variables between 2004 and 2006, and as a result the consumption rates of each group remained roughly constant, perhaps with the exception of households initially in the middle quantile who saw strong consumption expenditure growth, and thus their consumption rate displays a marked rise.

Table 4: Changes (2006-2010) in Selected Variables across PSID Net Worth

		%			% Expend. Rate (pp)	
NW Q	Net Worth[a]	Earn.	Disp Y	Cons. Exp.	Earn.	Disp Y
*Q*1	12.6k(+∞)	12.6	12.3	2.0	-9.7	-8.8
*Q*2	7.7k(35%)	8.5	9.1	3.8	-3.9	-4.2
*Q*3	7.3k(9%)	3.8	3.9	1.7	-1.6	-1.5
*Q*4	8.3k(4%)	4.3	3.3	-3.5	-5.6	-4.1
*Q*5	-118.8k (-11%)	-0.9	-2.3	-8.3	-5.2	-3.2

[a] *All changes in this table are rescaled to be comparable to the two-year changes in previous tables*
[b] *In 000's of dollars. Percentage change in parenthesis*

Now we turn to the dynamics in income, consumption and wealth during the Great Recession.[10] Table 4 displays the very significant changes throughout the wealth distribution, relative to the previous time period. Growth in net worth slowed down substantially for all quintiles, most significantly so at the top of the wealth distribution. In fact, on average net worth fell for households initially (that is, in 2006) in the top wealth quintile. Income growth also slowed down, although not uniformly across the wealth distribution. Comparing tables 3 and 4 we see that the slowdown in income growth is very modest at the bottom of the wealth distribution, whereas the middle and top quintiles experience a more substantial slowdown. For example, the 4th wealth quintile went from bi-annual disposable income growth 10.8% between 2004 and 2006 to a biannual growth rate of 3.3% between 2006 and 2010.

Most important for our purposes is the change in consumption expenditures at different points in the wealth distribution, especially in relation to the size of the associated earnings and disposable income change (as evident in the movement of the consumption rates over time). To highlight the starkest differences across the wealth distribution, focus on the difference between the top and the bottom wealth quintile. In 2004-2006 both the households in the bottom and in the top wealth quintile display small (less than 1.5%) changes in the consumption rate (out of disposable income). By contrast, in 2006 to 2010 all groups reduce their consumption rates, but most pronouncedly at the bottom end of the 2006 wealth distribution. For this group the consumption rate fell by 8.8%, whereas the top quintile's consumption rate declined only by 3.2%.

These differences suggest that household heterogeneity in net worth is relevant for

[10] Now the wealth ranking of households is based on the 2006 data.

determining the behavioral adjustments triggered by a massive aggregate economic downturn. Therefore the shape of the wealth distribution is potentially an important determinant of aggregate outcomes. In the next section we will present a standard (since Krusell and Smith, 1998) macroeconomic model with household heterogeneity in which the cross-sectional wealth distribution is the key aggregate state variable and study whether the exact form of this wealth distribution is an important determinant of macroeconomic consumption and investment dynamics in the presence of severe business cycle shocks. After parameterizing and solving the model, in section 5 we will assess the ability of the model to replicate the key empirical features of the cross-sectional distribution of resources documented in this section, prior to studying its business cycle properties and the importance of the wealth distribution in shaping it.

2.4 Consumption in the Great Recession: Changes in Income or Changes in Behavior?

Here we could add the results from Fabrizio's decomposition: informative about what a good consumption model with household heterogeneity should deliver.

3 A Canonical Business Cycle Model with Household Heterogeneity

In this section we lay out the benchmark model on which this chapter is built. The model is a slightly modified version of the original Krusell and Smith (1998) real business cycle model with household wealth and preference heterogeneity[11], and shares many features of the model recently studied by Carroll, Slacalek, Tokuoka and White (2015).

3.1 Technology

In the spirit of real business cycle theory aggregate shocks take the form of productivity shocks to the aggregate production function

$$Y = Z^{*}F(K, N) \tag{1}$$

Total factor productivity Z^{*} in turn is given by

$$Z^{*} = ZC^{\omega} \tag{2}$$

11 Krusell and Smith (1998) in turn build on stationary versions of the model with household wealth heterogeneity, and thus on Bewley (1986), Imrohoroglu (1989), Huggett (1993, 1997) and Aiyagari (1994).

where the level of technology Z follows a first order Markov process with transition matrix $\pi(Z'|Z)$. Here C is aggregate consumption and the parameter $\omega \geq 0$ measures the importance of an aggregate demand externality. In the benchmark model we consider the case of $\omega = 0$ in which case total factor productivity is exogenous and determined by the stochastic process for Z (and in which case we do not distinguish between Z and Z^*). In section 7 we consider a situation with $\omega > 0$. In that case current TFP and thus output is partially demand-determined.

In either case, in order to aid the interpretation of the results we will mainly focus on a situation in which the exogenous technology Z can take two values, $Z \in Z_l, Z_h$. We then interpret Z_l as a recession and Z_h as an expansion.

Finally, we assume that capital depreciates at a constant rate $\delta \in [0, 1]$.

3.2 Household Demographics, Endowments and Preferences

3.2.1 Demographics and the Life Cycle

In each period a measure 1 of potentially infinitely lived households populates the economy. Households are either young, working households (denoted by W) and participate in the labor market or are old and retired (and denoted by R). We denote a household's age by $j \in \{W, R\}$. Young households have a constant probability of retiring $1-\theta \in [0, 1]$ and old households have a constant probability of dying $1-\nu \in [0, 1]$. Deceased household are replaced by new young households. Given these assumption the distribution of the population is given by

$$\Pi_W = \frac{1-\theta}{(1-\theta)+(1-\nu)}$$
$$\Pi_R = \frac{1-\nu}{(1-\theta)+(1-\nu)}$$

3.2.2 Preferences

I In the benchmark model households do not value leisure, but have preference defined over stochastic consumption streams, determined by a period utility function $u(c)$ with the standard concavity and differentiability properties, as well as a time discount factor β that may be heterogeneous across households (but is fixed over time for a given household). Denote by B the finite set of possible time discount factors.

3.2.3 Endowments

Since households do not value leisure in the utility function young households supply their entire time endowment (which is normalized to 1) to the market. However, they face idiosyncratic labor productivity and thus earnings risk. This earnings risk comes from two sources. First, households are subject to unemployment risk. We denote by $s \in S = \{u, e\}$ the current employment status of a household, with $s = u$ indicating unemployment. Employment follows a first order Markov chain with transitions $\pi(s'|s, Z', Z)$ that depend on the aggregate state of the world. This permits the dependence of unemployment-employment transitions on the state of the business cycle.

In addition, conditional on being employed a household's labor productivity $y \in Y$ is stochastic and follows a first order Markov chain; denote by $\pi(y'|y) > 0$ denote the conditional probability of transiting from state y today to y' tomorrow, and by $\Pi(y)$ the associated (unique) invariant distribution. In the benchmark model we assume that, conditional on being employed, transitions of labor productivity are independent of the aggregate state of the world.[12]

For both idiosyncratic shocks (s, y) we assume a law of large numbers, so that idiosyncratic risk averages out, and only aggregate risk determines the number of agents in a specific idiosyncratic state $(s, y) \in S \times Y$. Furthermore, we assume that the share of households in a given idiosyncratic employment state s only depends on the current aggregate state[13] Z, and thus denote by $\Pi_Z(s)$ the deterministic fraction of households with idiosyncratic unemployment state s if the aggregate state of the economy is given by Z. We denote the cross-sectional distribution over labor productivity by $\Pi(y)$; by assumption this distribution does not depend on the aggregate state Z.

Households can save (but not borrow)[14] by accumulating (risky) physical capital and have access to perfect annuity markets.[15] We denote by $a \in A$ the asset holdings of an individual household and by A the set of all possible asset holdings. Households are born with zero initial wealth, draw their unemployment status according to $\Pi_Z(s)$ and their initial labor productivity from $\Pi(y)$. The cross-sectional population distribution of employment status s, labor productivity y, asset holdings a and discount factors β is denoted as Φ and summarizes, together with the aggregate shock Z, the aggregate state of the economy at any given point in time.

---

12 Even for the unemployed, the potential labor productivity evolves in the background and determines the productivity upon finding a job, as well as unemployment benefits while being unemployed, as described below.

13 This assumption imposes consistency restrictions on the transition matrix $\pi(s'|s, Z', Z)$. By assumption the cross-sectional distribution over y is independent of Z to start with.

14 We therefore abstract from uncollateralized household debt, as modeled in Chatterjee et al. (2007) and Livshits, MacGee and Tertilt (2007). Herkenhoff (2015) provides an investigation of the impact of increased access to consumer credit on the U.S. business cycle.

15 Thus the capital of the deceased is used to pay an extra return on capital $\frac{1}{\theta}$ of the survivors

3.3 Government Policy

3.3.1 Unemployment Insurance

The government implements a balanced budget unemployment insurance system whose size is parameterized by a replacement rate $\rho = \frac{b(y,Z,\Phi)}{w(Z,\Phi)y}$ that gives benefits b as a fraction of potential earnings wy of a household[16], with $\rho = 0$ signifying the absence of public social insurance against unemployment risk. These benefits are paid to households in the unemployment state $s = u$ and financed by proportional taxes on labor earnings with tax rate $\tau(Z, \Phi)$. Taxes are levied on both labor earnings and unemployment benefits.[17]

Recall that by assumption the number of unemployed $\Pi_Z(u)$ only depends on the current aggregate state. The budget constraint of the unemployment insurance system then reads as

$$\Pi_Z(u) \sum_y \Pi(y) b(y, Z, \Phi) = \tau(Z, \Phi) \left[\sum_y \Pi(y) \left[\Pi_Z(u) b(y, Z, \Phi) + (1 - \Pi_Z(u)) \, w(Z, \Phi) y \right] \right]$$

Exploiting the fact that $b(y, Z, \Phi) = \rho w(Z, \Phi) y$ and that the cross-sectional distribution over y is identical among the employed and unemployed we can simply:

$$\Pi_Z(u)\rho = \tau(Z, \Phi) \left[\Pi_Z(u)\rho + (1 - \Pi_Z(u)) \right]$$

and conclude that the tax rate satisfies:

$$\tau(Z, \Phi; \rho) = \left(\frac{\Pi_Z(u)\rho}{1 - \Pi_Z(u) + \Pi_Z(u)\rho} \right) = \left(\frac{1}{1 + \frac{1 - \Pi_Z(u)}{\Pi_Z(u)\rho}} \right) = \tau(Z; \rho) \in (0, 1) \quad (3)$$

That is, the tax rate $\tau(Z; \rho)$ only depends (positively) on the exogenous policy parameter ρ measuring the size of the unemployment system as well as (negatively) on the exogenous ratio of employed to unemployed $\frac{1 - \Pi_Z(u)}{\Pi_Z(u)}$ which varies over the business cycle.

3.3.2 Social Security

The government runs a balanced budget PAYGO system whose size is determined by a constant payroll tax rate τ_{SS} (that applies to labor earnings and unemployment benefits). Socially security benefits $b_{SS}(Z, \Phi)$ of retirees are assumed to be independent

16 Recall that even unemployed households carry with them the idiosyncratic state y even though it does not affect their current labor earnings since they are unemployed.

17 Since labor earnings are exogenous in the benchmark version of the model the tax is a lump sum tax.

of past contributions, but because of fluctuations in the aggregate tax base will vary the the aggregate state of the economy Z. The budget constraint then determines the relationship between benefits and the tax rate according to:

$$b_{SS}(Z,\Phi)\Pi_o = \tau_{SS}\Pi_y\left[\sum_y \Pi(y)\left[\Pi_Z(u)\rho w(Z,\Phi)y + (1-\Pi_Z(u))\,w(Z,\Phi)y\right]\right]$$

and thus the soical security replacement rate is a function of the tax rate τ_{SS}, the old age dependency ratio $\frac{\Pi_y}{\Pi_o}$ and average labor productivity in the economy:

$$\frac{b_{SS}(Z,\Phi)}{w(Z,\Phi)} = \tau_{SS}\frac{\Pi_y}{\Pi_o}\left[\Pi_Z(u)\rho + (1-\Pi_Z(u))\right]$$

Note that with a UI replacement rate of $\rho = 1$ (and with average labor productivity productivity of working people equal to 1) we have

$$\tau_{SS} = \frac{b_{SS}(Z,\Phi)}{w(Z,\Phi)}\frac{\Pi_o}{\Pi_y}$$

In this case the social security tax rate is simply equal to the average replacement rate $\frac{b_{SS}(Z,\Phi)}{w(Z,\Phi)}$ times the old age dependency ratio $\frac{\Pi_y}{\Pi_o}$.

3.4 Recursive Competitive Equilibrium

As is well-known the state space in this economy includes the entire cross-sectional distribution Φ of individual characteristics,[18] (j,s,y,a,β). Since the dynamic programming problems of young, working age households and retired households differ significantly from each other (both in terms of individual state variables as well the budget constraint) it makes notation easier to separate age $j \in \{W,R\}$ from the other state variables. The dynamic programming problem of retired households then reads as

$$v_R(a,\beta;Z,\Phi) = \max_{c,a'\geq 0}\left\{u(c) + \nu\beta\sum_{Z'\in Z}\pi(Z'|Z)v_R(a',\beta;Z',\Phi')\right\}$$

subject to

$$c + a' \quad = \quad b_{SS}(Z,\Phi) + (1 + r(Z,\Phi) - \delta)a/\nu$$

[18] In order to make the computation of a recursive competitive equilibrium feasible we follow Krusell and Smith (1998), and many others since, and define and characterize a recursive competitive equilibrium with boundedly rational households who only use a small number of moments (and concretely here, just the mean) of the wealth distribution to forecast future prices. For a discussion of the various alternatives in computing equilibria in this class of models, see the special issue of the Journal of Economic Dynamics and Control.

$$\Phi' = H(Z, \Phi', Z')$$

For young, working household households, the decision problem is given by

$$\begin{aligned} v_W(s, y, a, \beta; Z, \Phi) = \{\max_{c, a' \geq 0} u(c) \quad &+ \quad \beta \sum_{(Z', s', y') \in (Z, S, Y)} \pi(Z'|Z)\pi(s'|s, Z', Z)\pi(y'|y) \\ &* \quad [\theta v_W(s', y', a', \beta; Z', \Phi') + (1-\theta) v_R(a', \beta; Z', \Phi')]\} \end{aligned}$$

subject to

$$\begin{aligned} c + a' &= (1 - \tau(Z; \rho) - \tau_{SS}) w(Z, \Phi) y \left[1 - (1-\rho) 1_{s=u}\right] + (1 + r(Z, \Phi) - \delta) a \\ \Phi' &= H(Z, \Phi', Z') \end{aligned}$$

where $1_{s=u}$ is the indicator function that takes the value 1 if the household is unemployed and thus labor earnings equal unemployment benefits $b(y, Z, \Phi) = \rho w(Z, \Phi) y$.

Definition 1 *A recursive competitive equilibrium is given by value and policy functions of working and retired households, v_j, c_j, a'_j, pricing functions r, w and an aggregate law of motion H such that*

1. *Given the pricing functions r, w, the tax rate given in equation (3) and the aggregate law of motion H, the value function v solves the household Bellman equation above and c, a' are the associated policy functions.*
2. *Factor prices are given by*

$$\begin{aligned} w(Z, \Phi) &= Z F_N(K(Z, \Phi), N(Z, \Phi)) \\ r(Z, \Phi) &= Z F_K(K(Z, \Phi), N(Z, \Phi)) \end{aligned}$$

3. *Budget balance in the unemployment system: equation (3) is satisfied*
4. *Market clearing*

$$\begin{aligned} N(Z, \Phi) &= (1 - \Pi_Z(u)) \sum_{y \in Y} y \Pi(y) \\ K(Z, \Phi) &= \int a d\Phi \end{aligned}$$

5. *The aggregate law of motion H is induced by the exogenous stochastic processes for idiosyncratic and aggregate risk as well as the optimal policy function a' for assets.*[19]

[19] We give the explicit statement of the law of motion in appendix B

3.5 A Taxonomy of Different Versions of the Model

The following table 5 summarizes the different versions of the model we will study in this paper, including the section of the paper in which it will appear. We start with a version of the model in which total factor productivity is exogenous. The only source of propagation of the aggregate shocks is the capital stock, which is predetermined in the short run (and thus output is exogenous), but responds in the medium run to technology shocks and/or reforms of the social insurance system. We study two versions of the model, the original Krusell-Smith (1998) economy without preference heterogeneity (which we will alternatively refer to as the KS-economy, the low-wealth inequality economy, or the homogeneous discount factor economy), and a model with permanent discount factor heterogeneity (which we refer to as high wealth inequality economy or heterogeneous discount factor economy). The latter economy also features an unemployment insurance system whose size is calibrated to U.S. data. In section 5.1 we discuss the extent to which both versions of this model match the empirically observed U.S. cross-sectional wealth distribution, and in section 6.1 we trace out the model-implied aggregate consumption-, investment- and output dynamics in response to a great-recession type shock.

Table 5: Taxonomy of Different Versions of the Model Used in the Paper

Name	Discounting	Techn.	Soc. Ins.	Section
KS	$\beta = \bar{\beta}$	$\omega = 0$	$\rho = 1\%$	Sec. 6.1
Het. β	$\beta \in [\bar{\beta} - \epsilon\, \bar{\beta} + \epsilon]$	$\omega = 0$	$\rho = 50\%$	Sec. 6.1
Het. β	$\beta \in [\bar{\beta} - \epsilon\, \bar{\beta} + \epsilon]$	$\omega = 0$	$\rho = 10\%$	Sec. 6.3
Dem. Ext.	$\beta \in [\bar{\beta} - \epsilon\, \bar{\beta} + \epsilon]$	$\omega > 0$	$\rho = 50\%$	Sec. 7

In order to assess the interaction of wealth inequality and social insurance policies for aggregate macro dynamics, in section 6.3 we then study a version of the heterogeneous discount factor economy with (close to) no unemployment insurance. In section 7 the assumption of exogenous TFP is relaxed, and we present a version of the model in which TFP and thus output is partially demand-determined. In this version of the model household heterogeneity not only has a potential impact on the size of the consumption recession, but on the magnitude of the output decline as well, and by stabilizing individual consumption demand unemployment insurance may act as a quantitatively important source of macroeconomic stabilization.

4 Calibration of the Benchmark Economy

In this section we describe how we map our economy to the data. Since we want to address business cycles and transitions into and out of unemployment we calibrate the model to *quarterly* data.

4.1 Technology and Aggregate Productivity Risk

Following Krusell and Smith (1998) we assume that output is produced according to a Cobb-Douglas production function

$$Y = ZK^{\alpha}N^{1-\alpha} \tag{4}$$

We set the capital share to $\alpha = 36\%$ and assume a depreciation rate of $\delta = 2.5\%$ per quarter. For the aggregate technology process we assume that aggregate productivity Z can take two values $Z \in \{Z_l, Z_h\}$, where we interpret Z_l as a potentially severe recession. The aggregate technology process is assumed to follow a first order Markov chain with transitions

$$\pi = \begin{pmatrix} \rho_l & 1-\rho_l \\ 1-\rho_h & \rho_h \end{pmatrix}.$$

The stationary distribution associated with this Markov chain satisfies

$$\begin{aligned} \Pi_l &= \frac{1-\rho_h}{2-\rho_l-\rho_h} \\ \Pi_h &= \frac{1-\rho_l}{2-\rho_l-\rho_h} \end{aligned}$$

With the normalization that $E(Z) = 1$ the aggregate productivity process is fully determined by the two persistence parameters ρ_l, ρ_h and the dispersion of aggregate productivity, as measured by Z_l/Z_h.

We consider two calibrations of the model, one to assess the standard business cycle properties of the model, and one to study the great recession from the perspective of our model. For the former, we adopt the calibration of the aggregate technology process originally proposed by Krusell and Smith (1998). For the latter we think of a $Z = Z_l$ realization as a severe recession such as the great recession that began in 2008 or the double-dip recession of the early 1980's (and a realization of $Z = Z_h$ as normal times). In this interpretation of the model by choice of the parameters $\rho_l, \rho_h, Z_l/Z_h$ we want the model to be consistent with the fraction of time periods spent in severe recessions, their expected length (conditional on slipping into one) and the decline in GDP per capita associated with severe recessions.

For this we note that with the productivity process set out above, the fraction of time spent in severe recessions is Π_l whereas, conditional on falling into one, the expected length is given by:

$$EL_l = 1 \times 1 - \rho_l + 2 \times \rho_l (1 - \rho_l) + ... = \frac{1}{1 - \rho_l} \tag{5}$$

This suggests the following calibration strategy:

1. Choose ρ_l to match the average length of a severe recession EL_l. This is a measure of the persistence of recessions.

2. Given ρ_l choose ρ_h to match the fraction of time the economy is in a severe recession, Π_l.

3. Choose $\frac{Z_l}{Z_h}$ to match the decline in GDP per capita in severe recessions relative to normal times

In order to measure the empirical counterparts of these entities in the data we need an operational definition of a severe recession. This definition could be based on GDP per capita, total factor productivity or on unemployment rates, given the model assumption that the aggregate unemployment rate $\Pi_Z(y_u)$ is only a function of the aggregate state of the economy. We chose the latter and define a severe recession to be one where the unemployment rate rises above 9% at least for one quarter and determine the length of the recession to be the period for which the unemployment rate is above 7%. Using this definition during the period from 1948 to 2014.III we identify two severe recession period, from 1980.II-1986.II and 2009.I-2013.III. This delivers a frequency of severe recessions of $\Pi_l = 16.48\%$ with expected length of 22 quarters. The average unemployment rate in these severe recession periods rate is $u(Z_l) = 8.39\%$ and the average unemployment rate in the non-severe recession periods is $u(Z_h) = 5.33\%$. The implied Markov transition matrix that delivers this frequency and length of severe recessions has $\rho_l = 0.9545$ and $\rho_h = 0.9910$ and thus is given by:

$$\pi = \begin{pmatrix} 0.9545 & 0.0455 \\ 0.0090 & 0.9910 \end{pmatrix}.$$

For the ratio $\frac{Z_l}{Z_h}$ we target a value of $\frac{Y_l}{Y_h} = 0.9298$, that is, a drop of GDP per capita of 7% relative to normal times.[20] With average labor productivity if employed equal to 1 and if unemployed equal to zero and unemployment rates in normal and recession states equal to $u(Z_l) = 8.39\%$ and $u(Z_h) = 5.33\%$ and a capital share $\alpha = 0.36$ this requires $\frac{Z_l}{Z_h} = 0.9614$, which, together with the normalization

$$Z_l \Pi_l + Z_h \Pi_h = 1$$

determines the levels of Z as $Z_l = 0.9676, Z_h = 1.0064$. Note that because of endogenous dynamics of the capital stock which falls significantly during a great recession,

[20] This is the decline in GDP per capita during the two recession periods we identified, after GDP per capita is linearly de-trended, between 1964 to 2014.

the dispersion in total factor productivity is smaller than what would be needed to engineer a drop of output by 7% only through TFP and increased unemployment (which is the drop in output on impact, given that the capital stock is predetermined).[21]

As a matter of comparison, the aggregate productivity process used by Krusell and Smith (1998) that we also explore has $Z_l = 0.99, Z_h = 1.01$ (and associated unemployment rates of 10% and 4%, respectively), and with a transition matrix given by

$$\pi = \begin{pmatrix} 0.8750 & 0.1250 \\ 0.1250 & 0.8750 \end{pmatrix}.$$

4.2 Idiosyncratic Earnings Risk

Recall that households face two types of idiosyncratic risks, countercyclical unemployment risk described by the transition matrices $\pi(s'|s, Z', Z)$ and, conditional on being employed, acyclical earnings risk determined by $\pi(y'|y)$. We describe both components in turn.

4.2.1 Unemployment Risk

Idiosyncratic unemployment risk is completely determined by the four 2 by 2 transition matrices $\pi(s'|s, Z', Z)$ summarizing the probabilities of transiting in and out of unemployment. Thus $\pi(s'|s, Z', Z)$ has the form

$$\begin{bmatrix} \pi_{u,u}^{Z,Z'} & \pi_{u,e}^{Z,Z'} \\ \pi_{e,u}^{Z,Z'} & \pi_{e,e}^{Z,Z'} \end{bmatrix} \tag{6}$$

where, for example, $\pi_{e,u}^{Z,Z'}$ is the probability that an unemployed individual finds a job between one period and the next, when aggregate productivity transits from Z to Z'. Evidently each row of this matrix has to sum to 1. Note that, in addition, the restriction that the aggregate unemployment rate only depends on the aggregate state of the economy imposes one additional restriction on each of these two by two matrices, of the form

[21] In the short run,

$$\frac{Y_l}{Y_h} = \frac{Z_l}{Z_h}\left(\frac{1-u(Z_l)}{1-u(Z_h)}\right)^{0.64}$$

so that in order to generate a drop of output of 7% in the short run would require:

$$\frac{Z_l}{Z_h} = \frac{0.9298}{\left(\frac{0.9161}{0.9467}\right)^{0.64}} = 0.9496$$

.

$$\Pi_{Z'}(u) = \pi_{u,u}^{Z,Z'} * \Pi_Z(u) + \pi_{e,u}^{Z,Z'} * (1 - \Pi_Z(u)) \quad (7)$$

Thus, conditional on targeted unemployment rates in recessions and expansions, (Π_l, Π_h) this equation imposes a joint restriction on $(\pi_{u,u}^{Z,Z'}, \pi_{e,u}^{Z,Z'})$, for each (Z, Z') pair. With these restrictions, the idiosyncratic transition matrices are uniquely pinned down by the job finding rates[22] $\pi_{u,e}^{Z,Z'}$.

We compute the job finding rate for a quarter as follows. We consider an individual that starts the quarter as unemployed and compute the probability that at the end of the quarter that individual is still unemployed. The possible ways that this can happen are (denoting as f_1, f_2, f_3 the job finding rates in months 1,2 and 3 of the quarter):

1. Doesn't find a job in month 1, 2 or 3, with prob $(1 - f_1) \times (1 - f_2) \times (1 - f_3)$
2. Finds a job in month 1, loses it in month 2, doesn't find in month 3, with prob $f_1 \times s_2 \times (1 - f_3)$
3. Finds a job in month 1, keeps it in month 2, loses in month 3, with prob $f_1 \times (1 - s_2) \times s_3$
4. Finds a job in month 2, loses in month 3, with prob $(1 - f_1) \times f_2 \times s_3$

Thus the probability that someone that was unemployed at the beginning of the quarter is not unemployed at the end of the quarter is:

$$f = 1 - ((1 - f_1)(1 - f_2)(1 - f_3) + f_1 s_2 (1 - f_3) + f_1 (1 - s_2) s_3 + (1 - f_1) f_2 s_3) \quad (8)$$

We follow Shimer (2005) to measure the job-finding and separation rates from CPS data[23] as averages for periods corresponding to specific Z, Z' transitions and equating it with $\pi_{u,e}^{Z,Z'}$ delivers the following employment-unemployment transition matrices:

- Aggregate economy is and remains in a recession: $Z = Z_l . Z' = Z_l$

$$\begin{pmatrix} 0.3378 & 0.6622 \\ 0.0606 & 0.9394 \end{pmatrix} \quad (9)$$

[22] One could alternatively use job separation rares $\pi_{e,u}^{Z,Z'}$.

[23] Let u_t = unemployment rate and u_t^S = short-term unemployment rate (people who are unemployed this month, but were not unemployed last month). The we can define the monthly job-finding rate as $1 - (u_{t+1} - u_{t+1}^S)/u_t$ and the separation rate as $u_{t+1}^S/(1 - u_t)$. The series we use from the CPS are the unemployment level (UNEMPLOY), the short-term unemployment level (UNEMPLT5) and civilian employment (CE16OV). There was a change in CPS coding starting in February 1994 (inclusive), so UNEMPLT5 in every month starting with February 1994 is replaced by $UEMPL5 \times 1.1549$.

- Aggregate economy is and remains in normal times: $Z = Z_h . Z' = Z_h$

$$\begin{pmatrix} 0.1890 & 0.8110 \\ 0.0457 & 0.9543 \end{pmatrix} \tag{10}$$

- Aggregate economy slips into recession: $Z = Z_h . Z' = Z_l$

$$\begin{pmatrix} 0.3382 & 0.6618 \\ 0.0696 & 0.9304 \end{pmatrix} \tag{11}$$

- Aggregate economy emerges from recession: $Z = Z_l . Z' = Z_h$

$$\begin{pmatrix} 0.2220 & 0.7780 \\ 0.0378 & 0.9622 \end{pmatrix} \tag{12}$$

We observe that the resulting matrices make intuitive sense. One possible (but quantitatively minor) exception is that the job-finding rate is higher if the economy remains in normal times than if it emerges from a recession (note that the job separation rates all make perfect sense). On the other hand, the lower job-finding rate is consistent with the experience during the great recession per our definition, as job-finding rates did not recover until well into 2014, whereas by our calibration the recession ended in 2013.

4.2.2 Earnings Risk Conditional on Employment

In addition to unemployment risk we add to the model earnings risk, conditional on being employed. This allows us to obtain a more empirically plausible earnings distribution and makes earnings risk a more potent determinant of wealth dispersion (and thus reduces the importance of preference heterogeneity for this purpose). We assume that, conditional on being employed, log-labor earnings of households follow a process with transitory and with persistent shocks:[24]

$$\log(y') = p + \epsilon \tag{13}$$
$$p' = \phi p + \eta \tag{14}$$

with persistence ϕ and innovations of the persistent and transitory shocks (η, ϵ), respectively. The associated variances of the shocks are denoted by $(\sigma^2_\eta, \sigma^2_\epsilon)$, and therefore the entire process is characterized by the parameters $(\phi, \sigma^2_\eta, \sigma^2_\epsilon)$.

[24] We assume that the variance and persistence of this process is independent of the state of the business cycle. Earnings risk in our benchmark economy *is* countercyclical, as stressed by Storesletten, Telmer and Yaron (2004, 2007) and Guvenen, Song and Ozkan (2014), but in our model only because of countercyclical unemployment risk.

We estimate this process for household labor earnings after taxes (after first removing age, education and time effects) from *annual* PSID data[25] and find estimates of $(\hat{\phi}, \hat{\sigma}_\eta^2, \hat{\sigma}_\epsilon^2) = (0.9695, 0.0384, 0.0522)$. Next we translate these estimates into a quarterly persistence and variance.[26] We then use the Rouwenhorst procedure to discretize the persistent part of the process into a seven state Markov chain.[27] The iid shock only enters the computation of the expectation on the right hand side of the Euler equation.[28] We approximate the integral calculating the expectation using a Gauss-Hermite quadrature scheme with 3 nodes. Thus, we effectively approximate the continuous state space process by a discrete Markov chain with $7 \times 3 = 21$ states.[29]

4.3 Preferences

In the benchmark economy with exogenous labor supply choice we assume that the period utility function over current consumption is given by a constant relative risk aversion utility function with parameter $\sigma = 1$. As described above, we study two versions of the model, the original Krusell-Smith (1998) economy in which households have identical time discount factors, and a model in which households, as in Carroll et al. (2015) have permanently different time discount factors (and die with positive probability, in order to insure a bounded wealth distribution).

For the model with preference heterogeneity we adopt the specification proposed by Carroll et al. (2015). We found that this specification permits us to jointly match the high wealth concentration at the top of the distribution as well as the very significant share of the population with virtually zero net worth. Specifically, we assume that households at the beginning of their life draw their permanent β from a uniform distri-

[25] For the exact definition of the labor earnings after taxes, sample selection criteria and estimation method, please see Appendix A.

[26] In order to insure that quarterly log-earnings has the same persistence as annual log-earnings we choose the persistence of the quarterly AR(1) to be $\phi = \hat{\phi}^{\frac{1}{4}}$. For the variances, we note that the main purpose of the earnings shocks is to help deliver a plausible cross-sectional distribution of labor income. Therefore we aim to maintain the same cross-sectional distribution of earnings at the quarterly frequency as we estimate at the annual frequency. Choosing a quarterly transitory variance equal to its annual counterpart and

$$\frac{\sigma_\eta^2}{1-\phi^2} = \frac{\hat{\sigma}_\eta^2}{1-\hat{\phi}^2}$$

achieves this goal.

[27] See Kopecky and Suen (2010) for a detailed description and evaluation of the Rouwenhorst method.

[28] Since we use cash at hand and the persistent income state as state variables in the individual household dynamic programming problem.

[29] For the computation of the distributional statistics we simulate a panel of households. In this simulation the persistent shock remains on the grid, but the transitory shock is drawn from a normal distribution and thus is not restricted to fall on one of the quadrature points.

bution[30] with support $[\bar{\beta}-\epsilon, \bar{\beta}+\epsilon]$ and choose $(\bar{\beta}, \epsilon)$ so that the model wealth distribution (with an unemployment insurance replacement rate of 50%) has a Gini coefficient of 81.6% as in the data and a quarterly wealth-to-output ratio of 10.26 (as in Carroll et al., 2015) This requires $(\bar{\beta} = 0.9883, \epsilon = 0.01004)$. Finally, we set the death probability to $1/160$ for an expected working lifetime of 40 years, and thus $\theta = 0.9938$. The fact that households have finite (in expectation) lifetimes and that newborns start with zero wealth and draw their β afresh prevents the highest β households from asymptotically holding all the wealth in the economy.

For the original Krusell-Smith economy we choose the common quarterly discount factor $\beta = 0.989975$ to insure that the capital-output ratio in this economy (again at quarterly frequency) equals that in the heterogeneous β economy.

4.4 Government Unemployment Insurance Policy

The size of the social insurance (or unemployment insurance, more concretely) system is determined by the replacement rate ρ that given unemployment benefits as a fraction of average wages in the economy. For the benchmark economy that we calibrate to U.S. data we assume $\rho = 50\%$.

5 Evaluating the Benchmark Economy

5.1 The Joint Distribution of Earnings, Income, Wealth and Consumption in the Benchmark Economy

In this section we evaluate the extent to which our benchmark model is consistent with the main empirical facts characterizing the joint distribution of wealth, income and consumption expenditures, as well as the changes in this distribution when the economy is subjected to a large negative aggregate shock.

---

30 In practice we discretize this distribution and assume that each household draws one of five possible β's with equal probability; thus $B = \{\beta_1, ...\beta_5\}$ and $\Pi(\beta) = 1/5$. Since, conditional on survival, β is constant over a household's life, $\pi(\beta'|\beta) = I$. We also experimented with stochastic β's as in Krusell and Smith (1998) but found that the formulation we adopt enhances the model's ability to generate sufficiently many wealth-poor households. The results for the stochastic β economy generally lie in between those obtained in the original Krusell and Smith (1998) economy documented in detail in this paper, and the results obtained in the model with permanent β heterogeneity, also documented in great detail below.

5.1.1 Wealth Inequality in the Benchmark Economy

We have argued in the introduction that a model-implied cross-sectional wealth distribution that is consistent with the empirically observed concentration, and especially, with a share of wealth of the bottom 40% of close to zero, is crucial when using the model as a laboratory for studying aggregate fluctuations. We now document that our benchmark economy has this property whereas an economy akin to the one studied in Krusell and Smith's (1998) original work in which wealth inequality is entirely driven by idiosyncratic unemployment shocks and incomplete financial markets does not.[31]

Table 6: Net Worth Distributions: Data v/s Models

	Data		Models	
% Share held by:	PSID, 06	SCF, 07	Bench	KS
$Q1$	-1.2	-0.3	0.3	6.9
$Q2$	0.7	0.9	1.2	11.7
$Q3$	4.1	4.2	4.7	16.0
$Q4$	13.3	11.8	16.0	22,3
$Q5$	83.1	83.4	77.8	43.0
$90-95$	14.0	11.1	17.9	10.5
$95-99$	23.2	25.6	26.0	11.8
$T1\%$	30.2	34.1	14.2	5.0
Gini	0.77	0.79	0.77	0.35

Table 6 reports selected statics for the wealth distribution, both the one computed from the data (PSID and SCF) as well as from two model economies, the original Krusell-Smith (1998) economy and our benchmark model with idiosyncratic income risk, incomplete markets, a rudimentary life cycle structure, unemployment insurance and heterogeneous discount factors.[32] As indicated in the calibration section, through appropriate choice of the time discount factor(s) both economies have the same average (over the business cycle) capital-output ratio, and the benchmark economy displays a wealth Gini coefficient in line with the micro data. All other moments of the wealth distribution were not targeted in the calibration.

From the table we note that, overall, the benchmark model fits the empirical wealth

[31] We retain *our* calibration of unemployment risk, and thus our cross-sectional wealth distribution differs from their original numbers, but not in a substantial magnitude.

[32] Recall that in the data we restrict attention to households with at least one member of working age. Consequently, when we report cross-sectional statistics from the benchmark model (which includes a retirement phase) we restrict attention to households in the working stages of their life.

distribution in the data quite well, both at the bottom and at the top of the distribution. Specifically, it captures the fact that households constituting the bottom two quintiles of the wealth distribution hardly have any wealth, and that the top wealth quintile holds more than 80% of all net worth in the U.S. economy. We also acknowledge that the model makes the wealth middle class (quintiles 3 and 4) somewhat too wealth-poor and still misses the wealth concentration at the *very top* of the distribution. In the data the top 1% wealth holders account for over 30% of overall net worth in the economy, whereas the corresponding figure in the model is 22.0 %. A histogram of the model-implied wealth distribution can be found in figure 8 below. On the other hand, households between the 90th and the 99th percentile of the net worth distribution account for about 37% of wealth in the data, but 47% in the model.[33]

Finally, table 6 reproduces the (well-known, since Krusell and Smith, 1998) result that unemployment risk and incomplete markets alone are incapable of generating sufficient wealth dispersion. The problem relative to the data is two fold: households at the top of the wealth distribution are not nearly wealthy enough and (more importantly for the results to follow, as we will argue) households at the bottom of the distribution hold significantly too much wealth in the model, relative to SCF or PSID micro data. As a summary measure, whereas the wealth Gini in the data is well above 0.8, the original Krusell-Smith model delivers a number of only 0.35.In the next section we now confront the benchmark model with the empirical *joint* distribution of earnings, income, consumption and wealth in the data.

5.2 Inspecting the Mechanism I: What Accounts for Wealth Inequality in the Benchmark Economy?

A substantial literature, recently surveyed in De Nardi (2015), De Nardi, Fella and Yang (2015) and Benhabib and Bisin (2016), explores alternative mechanisms for generating the empirically observed high wealth concentration in the data. These mechanisms include the inclusion of very large but transient income realizations that the PSID misses out on (as in Castaneda, Diaz-Gimenez and Rios-Rull, 2003 or Kinderman and Krueger, 2015), large uninsured or only partially insured medical expenditure shocks in old age (see e.g. De Nardi, French and Jones, 2010, or Ameriks, Briggs, Caplin, Shapiro and Tonetti, 2015), the intergeneral transmission of wealth through accidental and intended bequests (as e.g. in De Nardi, 2004), the interaction between wealth accumulation and entrepreneurship (see Quadrini, 1999, Cagetti and De Nardi 2006, Buera 2009) or idiosyncratic shocks to investment opportunities or its returns, as in Benhabib, Bisin and Zhu (2011).

In our benchmark model we instead follow the sizeable literature that has explored the

[33] Although this is clearly a shortcoming, note that in this range of wealth levels the consumption function is essentially linear (as we will display below) and thus mechanically reshuffling wealth between the top 1% and the top 10% through top 1% would not alter aggregate consumption significantly.

potential importance of preference heterogeneity in general, and cross-sectional dispersion in patience specifically, for generating an empirically plausible cross-sectional wealth distribution. This mechanism had already been explored by the original Krusell and Smith (1998) paper, and has been further analysed by Hendricks (2007) and Carroll et al. (2015). In the previous section we argued that preference heterogeneity, when combined with idiosyncratic unemployment and earnings shocks as well as rudimentary life cycle elements[34] and social insurance policies, generates a wealth distribution that resembles the data in 2006 well, both at the bottom as well as at the top of the distribution. In table 11 we show which model element precisely is responsible for this finding.

Table 7: Net Worth Distributions and Consumption Decline: Different Versions of the Model

	Models*				
% Share:	KS	$+\sigma(y)$	+Ret.	$+\sigma(\beta)$	+UI
$Q1$	6.9	0.7	0.7	0.7	0.3
$Q2$	11.7	2.2	2.4	2.0	1.2
$Q3$	16.0	6.1	6.7	5.3	4.7
$Q4$	22.3	17.8	19.0	15.9	16.0
$Q5$	43.0	73.3	71.1	76.1	77.8
$90-95$	10.5	17.5	17.1	17.5	17.9
$95-99$	11.8	23.7	22.6	25.4	26.0
$T1\%$	5.0	11.2	10.7	13.9	14.2
Wealth Gini	0.350	0.699	0.703	0.745	0.767

*The KS model only has unemployment risk and incomplete markets, and thus the first column repeats information from table 6. The column $+\sigma(y)$ adds idiosyncratic earnings shocks (transitory and permanent) while employed. The column +Ret. adds the basic life cycle structure (positive probability of retirement and positive probability of death, plus social security in retirement); the column $+\sigma(\beta)$ incorporates preference heterogeneity into the model, and finally +UI raises the replacement of the unemployment insurance system from 1% to 50%; the resulting model is therefore the benchmark model, with results already documented in table 6. In all models the (mean) discount factor is calibrated so that all versions have the same capital-output ratio.

The table (which partially repeats information from table 6 to facilitate comparisons across different model economies) displays the share of net worth held by the five wealth quintiles, the wealth Gini and more detailed information about the top of the net worth distribution, in the data and in a sequence of models, ranging from the original Krusell-Smith (1998) economy to our benchmark economy. It contains two basic quantitative lessons. First, comparing the first and the second model columns, the inclusion of highly persistent earnings risk, in addition to unemployment risk, increases

[34] The literature on quantitative studies of the cross-sectional wealth distributions in general equilibrium life cycle economies with uninsurable idiosyncratic income risk starts with Huggett (1996)

wealth dispersion very significantly, relative to the economy with only unemployment risk. Given that we estimate earnings persistence, conditional on being employed, to be 0.97 on an annual level, the economy has a share of households with close to permanently low earnings, even in the absence of unemployment. These households, located predominantly in the lowest wealth quintile have had no opportunity[35] to accumulate significant wealth. Consequently the share of wealth held by the poorest household shrinks to very close to zero, as observed in the data. At the same time, the top wealth quintile is populated with households with high earnings realizations for whom the risk of a persistent fall of earnings provides motivation to accumulate substantial wealth. As a result, the wealth Gini doubles in the economy with earnings risk, relative to the original Krusell-Smith unemployment-only model. Adding a more explicit life cycle structure does not change the wealth distribution (of the working age population) much, but as we will see in the next section, will imply a more plausible joint wealth-consumption distribution. Incorporating a more generous unemployment insurance system further reduces the wealth held by the bottom two quintiles of the distribution since now losing a job with little net worth is not nearly as harmful. We will argue in our welfare analysis that the size of the unemployment insurance system not only shapes the bottom of the wealth distribution, but also has a strong impact on the welfare losses from severe recessions.

Second, however, as the examination of the very top of the wealth distribution reveals, this mechanism is insufficient to generate the very high wealth concentration. This is where the discount factor heterogeneity in the benchmark model plays a crucial role. It creates a class of households that are patient and have a high propensity to save, and the fact that in addition to a precautionary saving motive they also save for retirement (a phase they value highly due to their patience) insures that they do not start to decumulate wealth even at high wealth levels. As table 11 displays (commparing the last two columns) the model with both features (the life cycle and preference heterogeneity) is able to generate the wealth concentration at the top of the distribution that is observed in U.S. micro data.

5.2.1 Income and Consumption at Different Points of the Wealth Distribution

Table 8 reports the share of earnings, disposable income, consumption expenditures and the expenditure rates for the five quintiles of the wealth distribution, both for the data (as already contained in table 2) and for the benchmark model. Overall the model fares well in replicating the joint distributions of these variables, but with two notable exceptions. First, in the data households in the top quintile of the wealth distribution hold a significantly larger share of earnings and disposable income (37.2% and 41.4%, respectively) than in the model (27.8% an 35%, respectively). One possible ex-

[35] And if an unemployment insurance system with replacement rate of $\rho = 50\%$ is in place, no motive either.

planation for this discrepancy is that in the model wealthy households are no different in terms of their earnings process than poor ones (except for having had good luck with earnings realizations in the past).[36] Modeling explicitly households with differential (observed or unobserved) fixed characteristics such as differential education or innate ability would likely generate a higher positive correlation between earnings and wealth in the model and bring it in closer alignment to the data along this dimension.

Table 8: Selected Variables by Net Worth: Data v/s Models

	% Share of:						%s Expend. Rate			
	Earnings		Disp Y		Expend.		Earnings		Disp Y	
NW Q	Data	Mod	Data	Mod	Data	Mod	Data	Mod	Data	Mod
*Q*1	9.6	6.5	8.6	6.0	11.3	6.6	98.1	96.5	92.2	90.4
*Q*2	12.3	11.8	10.7	10.5	12.4	11.3	84.1	90.3	81.3	86.9
*Q*3	18.0	18.2	16.6	16.6	16.8	16.6	77.9	86.0	70.9	81.1
*Q*4	22.8	25.5	22.6	24.3	22.4	23.6	81.6	87.3	69.6	78.5
*Q*5	37.2	38.0	41.4	42.7	37.2	42.0	83.0	104.5	63.1	79.6
	Correlation with net worth									
	0.26	0.46	0.39	0.67	0.16	0.76				

The second main discrepancy between model and the data concerns consumption expenditure rates. In the model expenditure rates vary with wealth in a U-shaped pattern, with low and high wealth households consuming at a higher rate than households in the middle of the wealth distribution. In the model low wealth households, which also tend to be low income households, spend a high fraction of their current earnings because they expect earnings to increase, due to the degree of imperfect earnings (and unemployment) persistence we have estimated from the data. Households in the middle of the wealth distribution tend to have intermediate earnings realizations, therefore do not expect earnings to rise, and save for precautionary reasons to self-insure against potential declines in earnings in the future. Finally, households at the very top of the wealth distribution have, due to good luck and patience, reached a level of assets that shields them from the risk of low consumption, at least in the absence of an atypically long unemployment spell. Therefore these households consume a larger share of their current earnings and income.[37]

Note, however, that in the model expenditure rates are overall fairly flat across the wealth distribution. In the data we observe more marked differences, with low wealth

[36] High wealth households also tend to be more patient than those with low wealth, but patience is not correlated with income in the model. It is possible that if high earnings are associated with costly investments in human capital, more impatient households may forgo such investments, which in turn would generate an endogenous correlation between the degree of patience and earnings.

[37] In fact, they tend to decumulate part of their wealth on average and thus this group displays expenditure rates in excess of 100%.

households (the first, but also the second wealth quintile) exhibiting significantly higher expenditure rates than the other quintiles. Consequently households in the top wealth quintile spend a much smaller share of their earnings and disposable income in the data than in the model. In other words, whereas the model captures well the expenditure rates of wealth-poor households, it cannot rationalize the low expenditure rates at the top of the wealth distribution.[38].

5.3 The Dynamics of Income, Consumption and Wealth in Normal Times and in a Recession

In this section we compare, in the model and in the data, the *dynamics* of wealth, income and expenditures, first across two non-recession years (we refer to these as normal times), and then across two years in between which a recession has occurred. In the data we are somewhat limited in our choices by the sparse time series dimension of the PSID (for which comprehensive consumption data are available). We take normal times in the data to be the period from 2004-2006; we map this period into the model by studying an episode of eight quarters of good productivity, $Z = Z_h$, which in turn followed a long sequence of good aggregate shocks so that aggregates and distributions have settled down prior to this episode.

Table 9: Changes in Selected Variables by Net Worth in Normal Times (2004-2006): Data v/s Model

	Net Worth (%)		Disp Y (%)		Expend.(%)		Exp. Rate (pp)	
NW Q	Data	Model	Data	Model	Data	Model	Data	Model
Q1	$+\infty$	88	14.3	14.3	12.7	13.4	-1.4	-0.8
Q2	140	65	13.8	6.1	11.8	7.2	-1.5	1.0
Q3	50.0	39	9.4	3.2	18.4	4.9	6.3	1.6
Q4	28.2	18	10.8	1.0	8.8	3.4	-1.3	2.5
Q5	21.5	5	3.4	-2.0	5.9	0.9	1.4	2.9

Table 9 reports the statistics for the data (same as in table 3 above) together with the model. [39] In terms of wealth, the model captures well the fast accumulation of wealth of poor households, but it understates the accumulation of wealth by wealth-rich households. This is not surprising since in the data we observed a strong appreci-

[38] As with earnings and income, this discrepancy might in part be explained by the presence of different fixed characteristics (besides discount factor heterogeneity) between the wealth-rich and the wealth-poor that impact their saving behavior. One obvious candidate would be age, as older households tend to have higher earnings and accumulate wealth for retirement

[39] Since for tables 9 and 10 statistics for earnings and disposable income are quite similar we only report those for disposable income.

ation of house prices and financial asset valuations, whereas in the model the relative price of wealth (capital) is constant at one.

In terms of earnings (not reported) and disposable income, the model displays strong mean reversion, with income of the lowest wealth quintile rising fast (29.6%) and income of the high wealth falling substantially (-14.6%). As we saw earlier this is qualitatively consistent with the data, but quantitatively the model implies differences in income growth between the top and the bottom of the wealth distribution that are too large. In other words the model implies too much downward and upward mobility in incomes when households are ranked by wealth.[40]

With respect to consumption expenditures, table 10 shows that the model predicts that during normal times households in the bottom wealth quintile reduce their expenditure rates (by -5.7 percentage points [pp]) in normal times, whereas households at the top of the wealth distribution increase their expenditure rates (+4.8pp). The reason is intuitive from the perspective of the model: low wealth households have had, on average, unfortunate earnings realizations and their wealth is below their target wealth. Therefore these households cut their expenditure to re-build their wealth buffers. The opposite logic applies to households at the top of the wealth distribution. This implication of the model matches the data, although quantitatively, the differences in changes in expenditure rates between the top and the bottom wealth quintiles is larger in the model than in the data.

After documenting the dynamics of wealth, income and consumption (ordered by wealth) in normal times, table 10 displays the same statistics during a recession period.[41] For wealth the model predicts that during the recession wealth accumulation slows across all quintiles, although the reduction predicted by the model is smaller than in the data. For example, the wealth of the top net worth quintile grows at 1%, relative to the 3% growth in normal times. For the same quintile wealth growth in the data slows down from 21.5% to -11% between 2006 and 2010. As discussed above, in the data a large part of this reduction in wealth at the top of the distribution is the consequence of asset *price* movements which are, by construction, absent in the model.

The two other empirical facts we have documented in section 2.3 were that income declines in the recession hit the top wealth quintiles more than the bottom quintiles, and that households in the bottom quintiles cut expenditure rates more than households in the top quintiles. As table 10 shows, the first fact is not captured well by the model, which predicts a fairly uniform decline in income growth across the wealth distribution. This is seen by comparing disposable income growth in table 9 and 10.

In contrast, the model does well in capturing the differential changes in expenditure rates between normal and recession periods. Note that change in the consumption

40 Ranking households by earnings or income would make this statement even stronger.

41 In the model the Great Recession hits in Q.I, 2009, consistent with our calibration. In that quarter Z switches from $Z = Z_h$ to $Z = Z_l$ and remains there until Q.IV, 2010.

Table 10: Changes in Selected Variables by Net Worth in a Recession (2006-2010): Data v/s Model

	Net Worth (%)		Disp Y (%)		Expend.(%)		Exp. Rate (pp)	
NW Q	Data	Model	Data	Model	Data	Model	Data	Model
*Q*1	+∞	177	12.3	25.4	2.0	23.5	-8.8	-1.5
*Q*2	35	114	9.1	6.8	3.8	9.6	-4.2	2.7
*Q*3	9	62	3.9	-1.7	1.7	4.2	-1.5	6.2
*Q*4	4	26	3.3	-7.2	-3.5	1.1	-4.1	9.0
*Q*5	-11	3	-2.3	-14.5	-8.3	-3.8	-3.2	12.5

expenditure rate of the bottom wealth quintile falls from -5.7pp in normal times to -6.1 in the recession. Compare this with the change in the expenditure rate at the top of the wealth distribution, which rises from 4.8pp to 14.3pp. This finding is explained by the change in the precautionary saving motive induced by the recession. During the recession high wealth households increase their consumption rates because they use their assets to smooth consumption. Low wealth households, in contrast, cannot draw down wealth (since they have none). In fact, those low-wealth households that held on to their job increase their savings rates for precautionary reasons since the recession is a very persistent even and brings with it a substantial increase in unemployment risk. do not that because reduce the additional risk brought in by the recession.

We conclude this section by briefly summarizing the strengths and shortcomings of our baseline model when confronted with the PSID earnings, income, consumption and wealth data. The model succeeds in replicating the observed cross-sectional wealth distribution (except at the very top) and does well in capturing the salient features of the joint distribution of wealth, income and expenditures. It also replicates the empirical observation that low-wealth households cut their expenditure to a larger degree during a recession. In contrast, the model fails to capture the large movements in wealth we see in the data during the years 2006-2010 since it abstracts from asset price movements, it fails to reproduce the fact that high wealth households exhibit a substantially lower consumption rate than low wealth households, and finally, it implies too much mean reversion in income and earnings.[42]

[42] This statement applies despite the fact that the earnings process is directly estimated from the PSID.

6 Cross-Sectional Household Heterogeneity and the Aggregate Dynamics of Consumption and Investment in a Severe Crisis

In this section we argue that the cross-sectional distribution of households across individual characteristics (primarily in wealth and impatience) is a crucial determinant of the aggregate consumption and investment response to a negative business cycle shock. In addition we show that in the presence of such significant household heterogeneity the generosity of social insurance polices strongly affects the dynamics of macroeconomic aggregates.

6.1 Benchmark Results

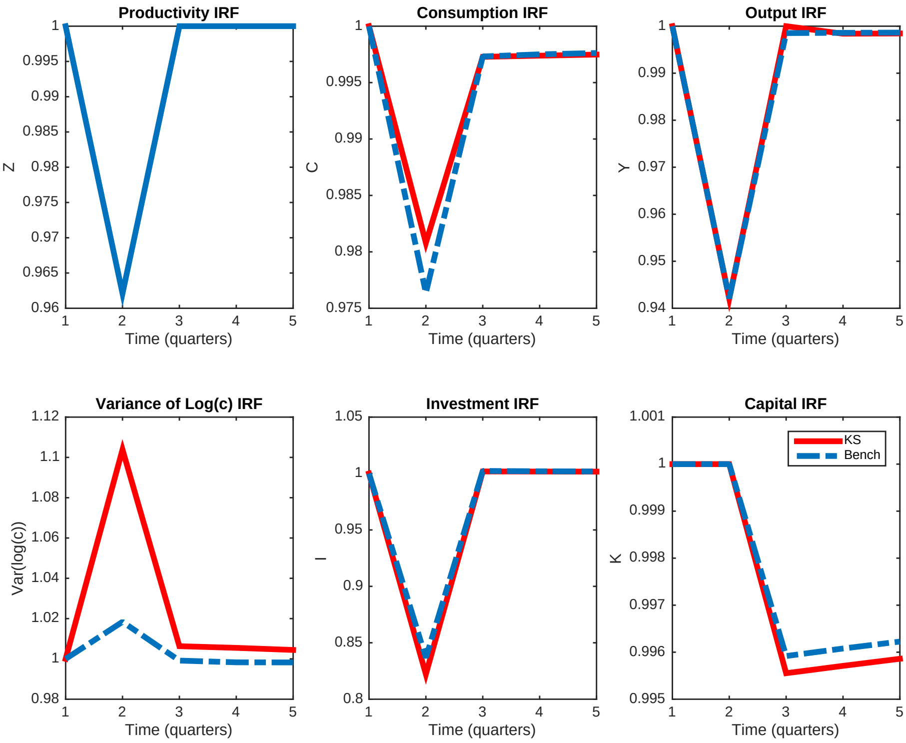
We consider two thought experiments, both of which take as initial condition the wealth distribution after a long sequence of good shocks so that the cross-sectional distribution has settled down. Then a severe recession hits. In the first thought experiment productivity returns to the normal state $Z = Z_h$ after one quarter (and remains there forever after). Although this thought experiment is not a good depiction of the actual great recession because of the short duration of the downturn, it displays the mechanics of the model most clearly.[43] In the second thought experiment we plot the response of the economy to a great recession of typical length (according to our calibration) that lasts for 5.5 years (22 quarters). In both cases we trace out the dynamic response of the key macroeconomic aggregates to the shocks. We are mainly interested in the extent to which the aggregate consumption and investment responses differ across two economies that differ fundamentally in their extent of household heterogeneity.

To make our main point we perform both experiments for two model economies: the original Krusell-Smith economy without preference heterogeneity, life cycle structure and only modest unemployment insurance, and our benchmark model that includes these features and therefore, as documented above, provides a model wealth distribution that matches its empirical counterpart very well. We will also show that the KS economy approximates an economy with representative agents (RA) very well, and thus when macroeconomic aggregates are concerned, the KS and the RA economy can be treated as quantitatively equivalent.

In figure 2 we plot the model impulse response to a one-time negative technology shock in which Z switches to Z_l after a long spell of good realizations Z_h. The upper left panel plots the time series of TFP Z fed into the model the remaining sub-plots show the model-implied dynamics of aggregate consumption, investment, output and capi-

[43] And of course households form expectations and make decisions based on the persistent Markov chain for Z driving the model.

Figure 2: Impulse Response to Aggregate Technology Shock in 2 Economies: One Time Technology Shock



tal, as well as the evolution of consumption inequality induced by the great recession-type TFP shock. By construction the time paths of exogenous TFP Z are identical in both economies in the short run; for output they are identical on impact and virtually identical over time. Since TFP and labor supply are exogenous in both economies and follow the same time path, and capital is predetermined on impact, and the one time shock is not sufficient to trigger a substantially different dynamics of the capital stock (see the units on the y-axis for the lower right hand panel), the time path of output is virtually identical in both economies. Thus the key distinction between both economies is the extent to which a very similar decline and recovery in output is reflected in lower aggregate consumption rather than aggregate investment.

The key observation we want to highlight is that the aggregate consumption (and thus investment) response to the negative productivity shock differs substantially between the two economies. In the benchmark model consumption falls by 4.2% in response to a technology shock that induces a decline in output by 6% on impact. The same fall in output only triggers a decline of 1.9% in the original Krusell-Smith (labeled as KS) economy. Thus the impact of the recession on aggregate consumption more than twice as large in the economy with empirically plausible wealth heterogeneity. The dynamics of investment follows the reverse pattern; consequently the investment recession is much less severe in the benchmark economy, and the output recovery more potent, although this effect is quantitatively minor, as the upper right plot shows. Given that output is exogenous in the short run, and used for consumption and investment only in this closed economy, the investment impulse response necessarily shows the reverse pattern: the decline in investment is much weaker in the high wealth inequality economy. This in turn triggers a less significant decline and more rapid recovery of the macro economy once the recession has ended. However, given that new investment is only a small fraction of the capital stock, these differential effects on capital, and thus output, are quantitatively minor (notice the units on the axis of the capital impulse response), at least in the case where the recession is short-lived.

Note that for all practical purposes, in what follows the Krusell-Smith economy displays an aggregate consumption-investment dynamics very close to a representative agent (RA) economy. Figure 3 shows this fact. For example, the aggregate consumption decline in the RA economy amounts to 1.78%, relative to a fall in aggregate consumption of 1.9% in the KS economy. Of course, one exception is the cross-sectional variance of consumption which by construction is zero (and stays at zero throughout the recession) in the RA economy (and thus is not displayed in the lower left panel of figure 3).

Figure 3: Impulse Response to Aggregate Technology Shock in Krusell-Smith and RA Economy

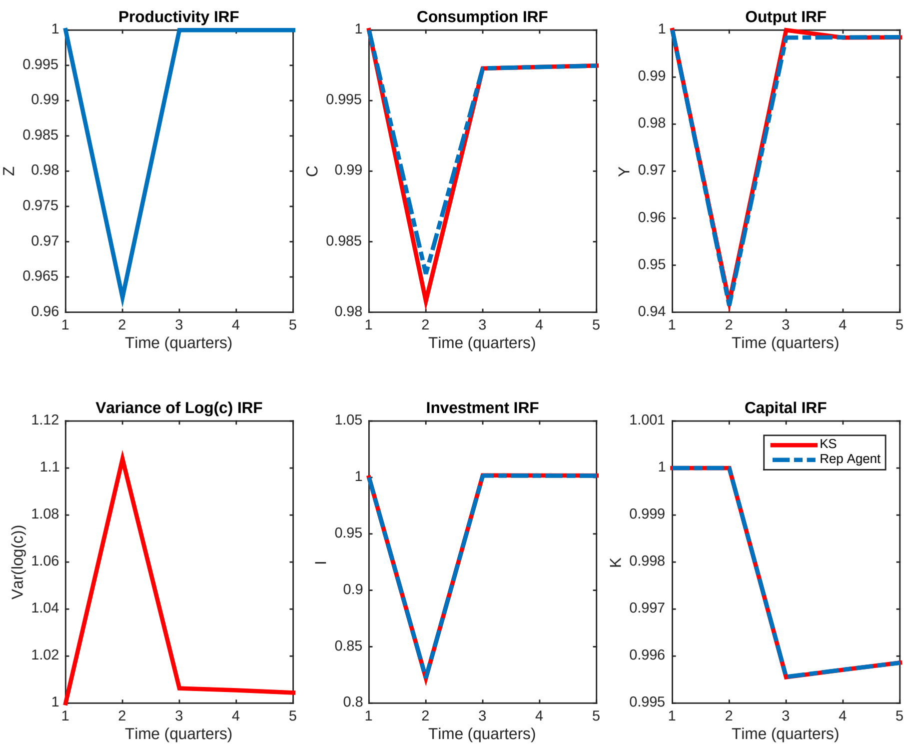
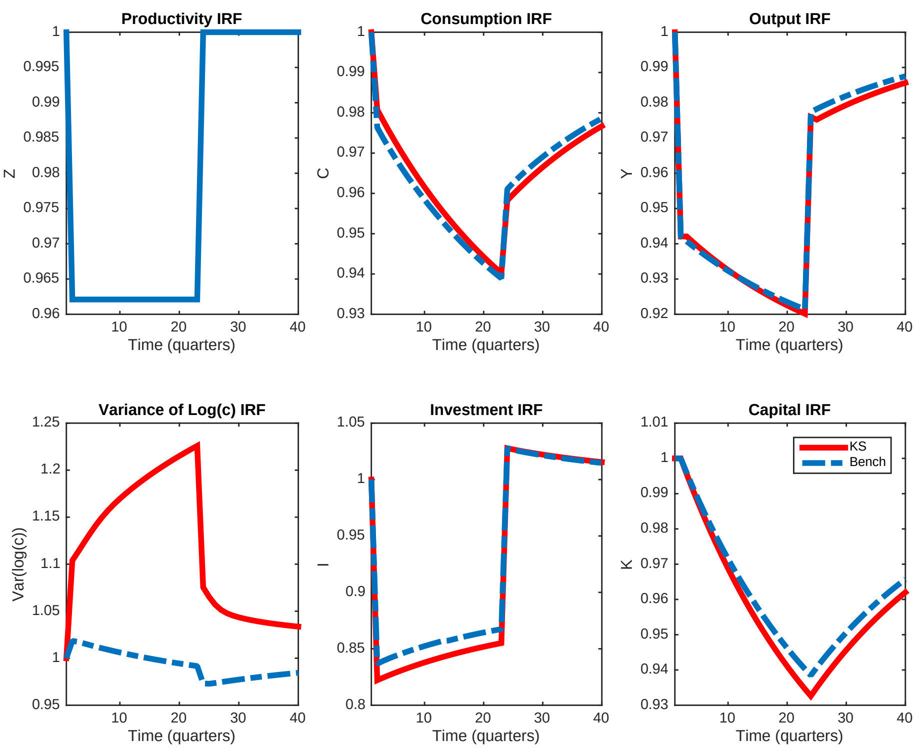


Figure 4 below demonstrates that a in great recession lasting several years as in the data, the differences in capital and output dynamics across the low-wealth inequality KS economy and the high inequality benchmark are now more noticeable, especially towards the end of the recession. As a result, the recovery after TFP has turned back up again is substantially stronger, but approximately one percentage point in the period in which the recession ends. Also, strongly noticeable in this figure (as was already in the one-period recession in figure 2) is the more potent increase in consumption inequality (albeit from a much lower initial level) in the KS economy as more and more individuals are driven down towards low asset and thus low consumption levels; in the high wealth inequality economy most wealth-poor households are already there.

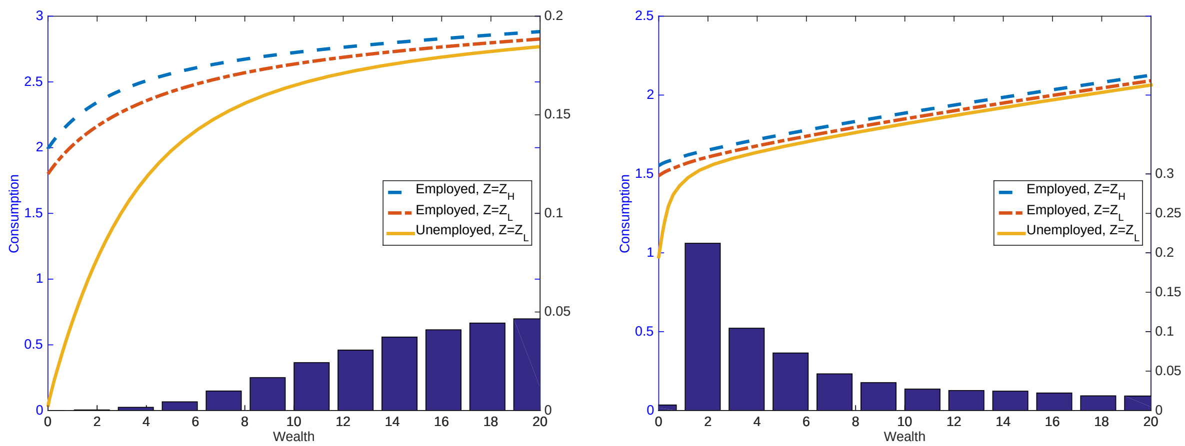
Figure 4: Impulse Response to Aggregate Technology Shock in 2 Economies: "Typical" Severe Recession Technology Shock



6.2 Inspecting the Mechanism II: What Accounts for the Size of the Aggregate Consumption Recession

The key finding from the last section is that the aggregate consumption recession in our benchmark economy with preference and realistic wealth heterogeneity is more than twice as deep as in the corresponding representative agent economy (which in turns displays aggregate time series very close to those in the original Krusell and Smith economy). In this section we dissect the reasons behind this finding. To start, in figure 7 we display the consumption functions and wealth distributions both for the KS and the benchmark economy. The left panel shows the consumption functions (plotted against individual wealth on the x-axis) in the original Krusell-Smith economy for three combinations of idiosyncratic employment and aggregate productivity states. For a given wealth level, the vertical difference between the consumption functions for the employed in aggregate state $Z = Z_h$ (blue dashed line) and the employed in aggregate state $Z = Z_l$ (red, dot-dashed line) gives the consumption drop in the great recession, conditional on not losing a job. In the same way, the vertical distance between the blue-dashed consumption function and the orange sold consumption function (for the unemployed in the recession) gives the consumption decline for those households that lose their jobs in a recession. The figure also contains the pre-recession wealth distribution, displayed as a histogram.[44] The right panel displays the same information, but for our benchmark economy, for working age households with mean earnings state y and mean discount factor $\bar{\beta}$.

Figure 5: Consumption Function, Wealth Distribution, Krusell-Smith (left panel) and Benchmark (right panel) Economies



The first observation we make is that, for a given level of wealth, the drop in individual consumption as the KS economy falls into a great recession is substantially *larger* than in our benchmark economy.[45] This is especially true for households with little wealth that lose their jobs at the onset of the recession, due to the absence of unemployment

[44] The aggregate capital stock on which these plots are based is the pre-recession capital stock; note that both economies, by virtue of the calibration, have the same average (over the cycle) capital stock.

[45] The figure displays the consumption functions in the benchmark economy for individuals with median (y, β), but (as we will show below), the same statement will apply, qualitatively, to the consumption functions for households with other (y, β) characteristics. Recall that there is no (y, β) heterogeneity in the original KS economy.

insurance.[46]

The observation of larger individual consumption declines in the KS economy would suggest that the aggregate consumption recession is actually larger than in the benchmark economy, in contrast to the result documented in the previous section. However, as the figure 7 (and table **??**) display clearly, the cross-sectional wealth distribution places almost no mass on households with very little net worth, exactly the households with the largest consumption declines. In contrast, the benchmark model with realistic wealth inequality places substantial probability mass at zero or close to zero wealth where the individual consumption losses are significant, especially (but not only) for newly unemployed households.[47] Note that average net worth is the same in both economies: we truncate the plots at net worth twenty times average income in order to make the individual consumption declines at the low end more clearly visible, but the benchmark economy has a fat right tailed wealth distribution that is well-approximated by a Pareto distribution (as is the data, see e.g. Benhabib and Bisin, 2016), whereas the original KS economy displays a wealth distribution whose right tail more closely resembles a log-normal distribution. Thus both distributions have the same mean despite the fact that, as clearly visible from the figure, the benchmark economy places substantially more weight on low net-worth households.

As we will see in subsection 6.3, the same twofold impact of social insurance (on the consumption response to aggregate shocks for given wealth level, and on the wealth distribution itself) is also crucial when determining the overall impact of unemployment insurance policies on the consumption and investment dynamics over the business cycle. First, however, we further explore the precise reasons behind the significant differences in aggregate and distributional characteristics between the original KS economy and our benchmark, thereby pinpointing precisely what model elements (and their interaction) are responsible for the main point of this paper that modelling household heterogeneity explicitly is a crucial quantitative determinant of aggregate consumption in the great recession.

Relative to the KS model our benchmark includes idiosyncratic earnings shocks[48], a rudimentary life cycle structure with social security system[49], a (more generous) un-

[46] In order to avoid numerical problems with zero consumption we include a minimal unemployment insurance system with a replacement rate of $\rho = 1\%$ in the KS economy.

[47] The right panel shows the wealth distribution for the entire working age population, rather than conditioning on the specific (y, β) types for which the consumption functions are displayed.

[48] As in Aiyagari (1994) and the large subsequent quantitative literature employing the standard incomplete markets model, reviewed recently by Heathcote, Storesletten and Violante (2009).

[49] See Huggett (1996) for an early study of the cross-sectional wealth distributions in general equilibrium life cycle economies with uninsurable idiosyncratic income risk.

employment insurance system and permanent preference heterogeneity.[50] In table 11 we add these model elements one by one, and show how they shape the cross-sectional wealth distribution[51] as well as the magnitude of the aggregate consumption recession when the economy is hit by a great recession type TFP shock of equal magnitude as in the benchmark model.

Table 11: Net Worth Distributions and Consumption Decline: Different Versions of the Model

	Models*					
% Share:	KS	$+\sigma(y)$	+Ret.	$+\sigma(\beta)$	+UI	KS + Top 1%
$Q1$	6.9	0.7	0.7	0.7	0.3	5.0
$Q2$	11.7	2.2	2.4	2.0	1.2	8.6
$Q3$	16.0	6.1	6.7	5.3	4.7	11.9
$Q4$	22.3	17.8	19.0	15.9	16.0	16.5
$Q5$	43.0	73.3	71.1	76.1	77.8	57.9
$90-95$	10.5	17.5	17.1	17.5	17.9	7.4
$95-99$	11.8	23.7	22.6	25.4	26.0	8.8
$T1\%$	5.0	11.2	10.7	13.9	77.8	30.4
Wealth Gini	0.350	0.699	0.703	0.745	0.767	0.525
ΔC	-1.9%	-2.5%	-2.6%	-2.9%	-2.4%	-2.0%

*The KS model only has unemployment risk and incomplete markets, and thus the first column repeats information from table 6. The column $+\sigma(y)$ adds idiosyncratic earnings shocks (transitory and permanent) while employed. The column +Ret. adds the basic life cycle structure (positive probability of retirement and positive probability of death, plus social security in retirement); the column $+\sigma(\beta)$ incorporates preference heterogeneity into the model, and finally +UI raises the replacement of the unemployment insurance system from 1% to 50%; the resulting model is therefore the benchmark model, with results already documented in table 6. In all models the (mean) discount factor is calibrated so that all versions have the same capital-output ratio.

The table (which partially repeats information from table 6 to facilitate comparisons across different model economies) displays the share of net worth held by the five wealth quintiles, the wealth Gini and more detailed information about the top of the net worth distribution in a sequence of models, ranging from the original Krusell-Smith (1998) economy to our benchmark economy. It contains several important insights that we will explore in greater detail below, but two broad findings stand out.

[50] The potential importance of preference heterogeneity for explaining the empirically observed wealth concentration has already been stressed by the original Krusell and Smith (1998) paper, and is further explored by Hendricks (2007) and Carroll et al. (2015).

[51] For a discussion of alternative mechanisms leading to plausible cross-sectional wealth distributions, see the recent surveys by De Nardi (2015) and Benhabib and Bisin (2016).

First, adding to the model with idiosyncratic unemployment risk highly persistent earnings risk[52] upon being employed, not only permits the model to deliver a plausible cross-sectional earnings distribution (as documented in the last subsection), but also brings the

wealth dispersion very significantly, relative to the economy with only unemployment risk. Given that we estimate earnings persistence, conditional on being employed, to be 0.97 on an annual level, the economy has a share of households with close to permanently low earnings, even in the absence of unemployment. These households, located predominantly in the lowest wealth quintile have had no opportunity[53] to accumulate significant wealth. Consequently the share of wealth held by the poorest household shrinks to very close to zero, as observed in the data. At the same time, the top wealth quintile is populated with households with high earnings realizations for whom the risk of a persistent fall of earnings provides motivation to accumulate substantial wealth. As a result, the wealth Gini doubles in the economy with earnings risk, relative to the original Krusell-Smith unemployment-only model. However, as the examination of the very top of the wealth distribution reveals, this mechanism is insufficient to generate the very high wealth concentration

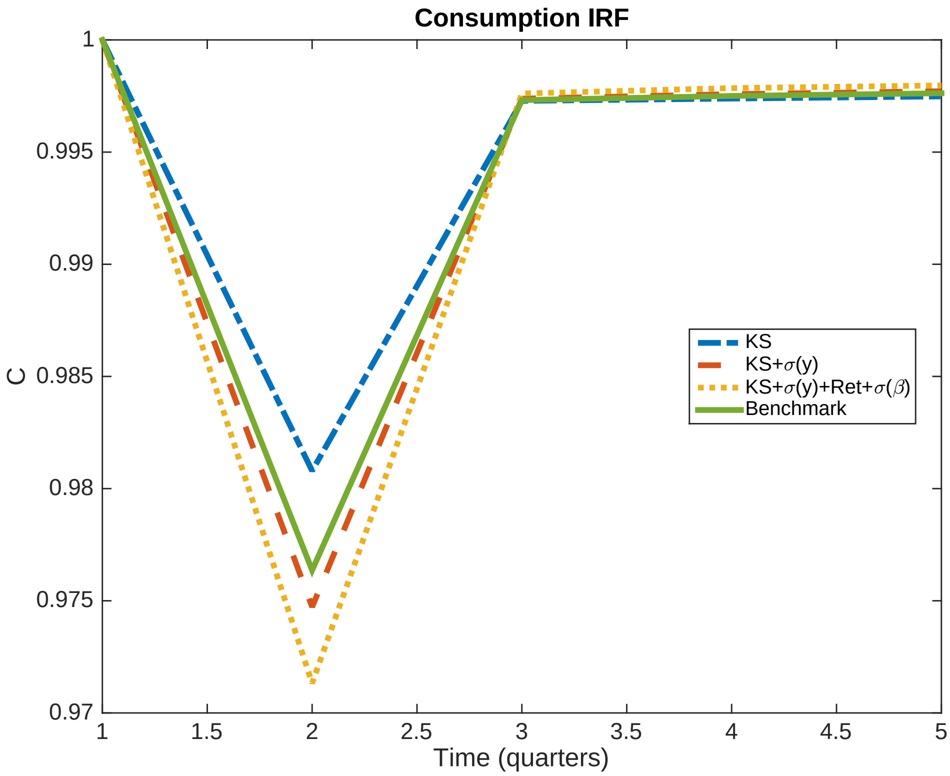
Why we get such a big drop:

- Rep Agent is basically the same as KS, het doesn't really matter there
- Adding in income risk does increase inequality, which gives us some poor guys who have higher MPCs, which increases the response of consumption. However, these guys aren't as poor as in the data, so they are still less responsive to the aggregate shocks
- So why does adding in preference het along with incoem risk do so much? It's all in the impatient, high income types. For them, the consumption drop, *conditional on staying employed* is still almost 20% when we fall into a recession. The idea is that these households had high consumption rates during normal times because they don't care very much about saving for retirement and unemployment risk is low. When we fall into a recession, the idiosyncratic risk goes up significantly for the "foreseeable future" from the point of view of impatient guys. Faced with the much higher chanced of becoming unemployed, these households cut consumption and start saving a lot more. So there's a big precautionary savings motive. The reason why we don't see similar behavior from the patient guys with high income, is that they were already saving a large fraction of their income even in good times, since they are more focused on the long horizon. Because the persistent income component is more persistent than the recession, for the patient

[52] Recall the AR(1) coefficient of the persistent component was estimated to be 0.97 on an annual level.

[53] and if an unemployment insurance system with replacement rate of $\rho = 50\%$ is in place, no motive either.

Figure 6: Consumption Recessions in Various Versions of the Model

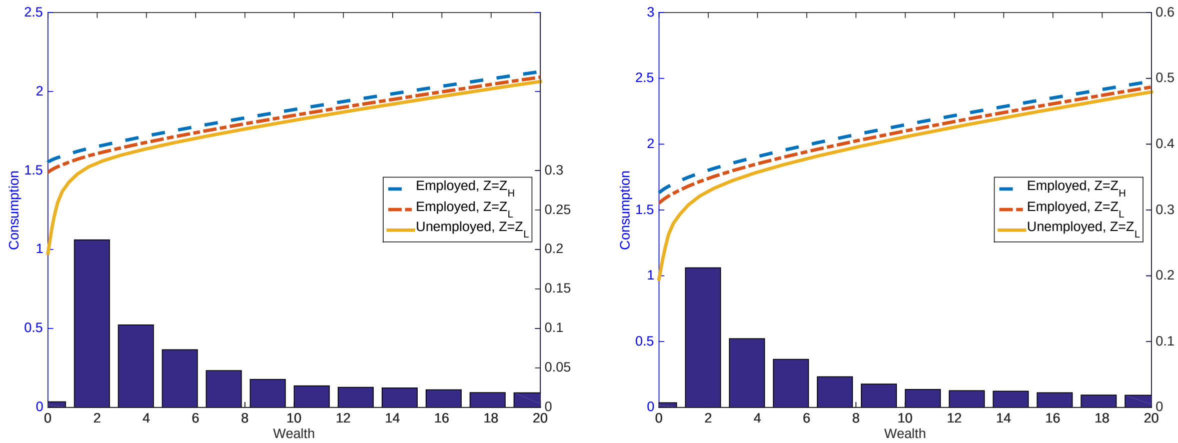


high income guys, they expect to have high income even exiting the recession, so the short run possibility of increased unemployment is not as big of a concern to them.

- low beta people care about the current cycle and thus precautionary saving for unemployment spell in the great recession is important relative to life cycle saving. Strong consumption response to persistent recession
- high beta people mainly save for life cycle reasons, and thus the change in precautionary saving as economy falls into recession not at important

There are two key differences between our benchmark, the high wealth inequality economy, and the original Krusell-Smith (1998) low wealth inequality economy. First, in addition to unemployment risk, we model earnings risk even conditional on being employed. Second, in our model, as in Carroll et al. (2015), households differ permanently (conditional on survival) in their time discount factors. As documented above, the wealth distribution is more dispersed, with a larger fraction of households at or close to the borrowing constraint, and the aggregate consumption response is about 50% larger in the in the high-wealth inequality economy. We now document which of these two novel (relative to the original Krusell-Smith economy) model elements is

Figure 7: Consumption Function, Wealth Distribution, Patient Households (left panel) and Impatient Households (right panel)



mainly responsible for these findings. Table **??** summarizes the results.

In figure **??** we shut down earnings risk (but retain unemployment risk), and compare the impulse responses to a one-time productivity shock to those in the benchmark model. The cross-sectional distribution of time discount factors (and all other parameters) is held constant, rather than recalibrated.[54] As a consequence of reduced earnings risk, fewer households save for precautionary reasons. Relative to the benchmark, the wealth Gini rises to 0.912 and the capital-output ration K/Y falls to 9.844. As figure **??** displays, the collapse in aggregate consumption is now even more severe, amounting to 3.87% rather than 2.64% as in the benchmark economy.[55]

In figure **??** we instead eliminate heterogeneity in time discount factors β endow all households with the same discount factor as in the KS model[56], and compare the results to the benchmark economy with β heterogeneity. Again all other parameters remain unchanged. In the model without β heterogeneity, the Gini falls sharply, to 0.552 and K/Y rises to 11.84. With less dispersion in discount factors, wealth dispersion declines massively, the share of households at or close to the borrowing constraint (with strong consumption response to earnings) falls, resulting in an aggregate consumption decline that is close to that of the original Krusell-Smith economy.

[54] The objective of this section is to understand what drives the results in the previous section, not to investigate which version of the model is most successful empirically.

[55] Essentially, absent significant income risk, only the very patient employed people save, and everyone else is content to behave as hand to mouth consumers facing mostly aggregate risk, which the economy cannot insure against.

[56] Note that the results are unchanged qualitatively and essentially unchanged quantitatively if instead the average β from the baseline economy was used. Using the average β from the benchmark economy the Gini falls to 0.565 and K/Y falls to 9.71

Thus to a very good first approximation, the additional earnings risk helps to increase the earnings Gini relative to the original Krusell-Smith benchmark[57]. But it mainly affects the top end of the wealth distribution, and to a lesser extent the bottom of the wealth distribution. Discount factor heterogeneity, instead, introduces some very impatient households into the economy who hold little or no wealth, independent of whether they are subject only to unemployment risk or also additional earnings risk. These households are in turn crucial for reproducing the close to 40% households in the data with little net worth; it is in turn this group who is the most important determining factor in the *aggregate* consumption response to a great recession shock that sees the incomes of all fall, and induces higher unemployment in the economy.

6.2.1 The Importance of "Hand to Mouth" Consumers

Given the importance we assigned to households with little net worth in our discussion above, in this subsection we briefly ask whether a model with a *fixed* fraction of households κthat always have zero wealth and thus simply consume their income in every period has the same implications for the consumption dynamics as our benchmark model.[58] We resolved our model under the assumption that the bottom $\kappa = 40\%$ of the wealth distribution in model period $t-1$ just consumes their earnings and unemployment benefits (if applicable) from period t on, whereas the remainder of the distribution (in period $t-1$) continues to follow the intertemporally optimal decision rules from the benchmark economy.

The drop of consumption in a one-period great recession (the counterpart of figures **??** and **??**) now amounts to 2.30%, relative to the decline in the benchmark economy of 2.64%. The drop is larger in the benchmark economy since households at the bottom of the wealth distribution on average (and especially those not currently experiencing unemployment benefits) find it optimal to *reduce* consumption rates for precautionary reasons: the great recession is expected to last a long time, and those not yet affected by a job loss try to build a buffer to hedge against the increased risk of being laid off in the future. This precautionary saving motive is absent among households that follow a mechanical hand-to-mouth consumption rule and is responsible for the deeper (by

57 It also helps the model to generate a more realistic correlation between earnings and wealth.

58 This question is interesting from a modeling perspective since a model in which a fixed fraction κ of households act as hand-to-mouth and the remaining fraction employs permanent income consumption and savings functions (which are linear in wealth with identical marginal propensities to consume out of wealth, given our model) would give rise to easy aggregation: aggregate consumption would simply be κ times aggregate income of the hand to mouth consumers plus $1-\kappa$ times aggregate wealth times the common (and small) MPC out of wealth times total wealth in the economy (which equals the aggregate capital stock). If one assumes that hand to mouth consumers and PIH consumers face the same income rpocess, then the consumption drop is $[\kappa + (1-\kappa) * \frac{r}{1+r}] \approx \kappa$ times the drop in income. This model is clearly much more tractable than our benchmark model, and therefore a relevant question is whether this model has significantly different predictions than our benchmark economy.

34 basis points) recession in the benchmark economy.[59]

Add reference to wealthy hand to mouth consumers, Kaplan and Violante and Bayer, Luetticke, Pham-Daoz and Tjadden (2015). Here also good place to discuss the precautionary savings motive of the poor, relate to DenHaan and Rendahl. In terms of earnings, the feature of unemployment risk going up in recessions generates exactly the time-varying skewness that Guvenen and co-authors have stressed.

6.3 The Impact of Social Insurance Policies

In this section we ask how the presence of public social insurance programs affects the response of the macro economy to aggregate shocks in a world with household heterogeneity. We focus specifically on the effects of government-provided, and tax-financed unemployment insurance. We will argue that the impact of this policy is two-fold: it changes the consumption-savings response of a household with a given wealth level to income shocks, and it changes the cross-sectional wealth distribution in society, at least in the medium to long run.

In the left panel of figure 8 we plot, against wealth, the consumption functions (for the unemployed in the low and the employed in the high aggregate shock, with the mean discount factor) as well as the wealth histogram in the benchmark economy (with a replacement rate of 50%). This was the right panel of figure 7. The right panel of Figure 8 does the same for an economy with an unemployment insurance system of only 10%. The reason we chose to display the consumption function for the employed in an expansion and the unemployed in a recession is that this helps us best to understand what drives the aggregate consumption impulse response below.[60]

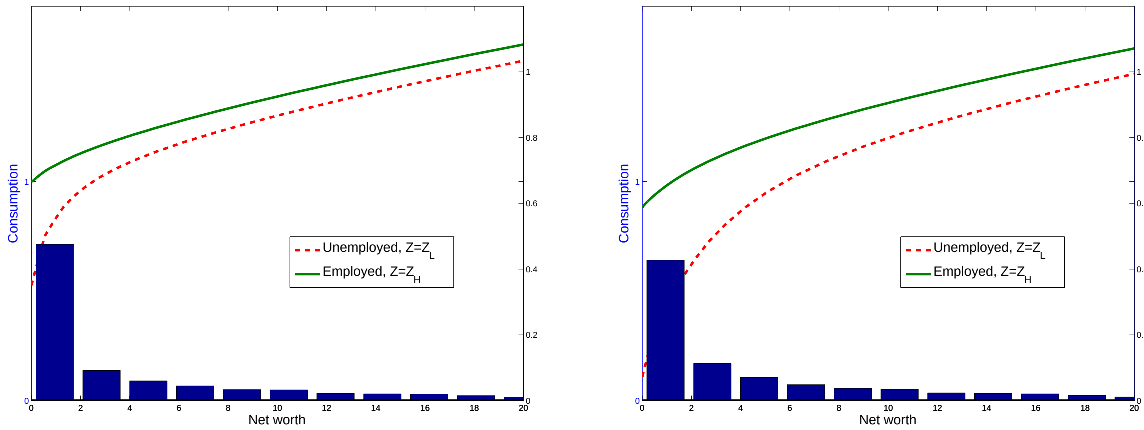
We want to highlight three observations. First, in the high unemployment insurance economy households with low wealth consume much more than in the economy with small unemployment insurance. Second, and related, the decline in consumption for low wealth households from experiencing a recession with job loss is much more severe in the low-benefit economy. However, and third, the size of the social insurance system, by affecting the extent to which households engage in precautionary saving, is a crucial determinant of the equilibrium wealth distribution. In the benchmark economy (as in the data) a sizable mass of households has little or no wealth, whereas in the no-benefit economy this share of the population declines notably[61].

[59] Obviously, the magnitude of this effect depends on the share of hand to mouth consumers κ. In the limit, as $\kappa = 0$ we are back in the benchmark economy. For $\kappa = 20\%$ the fall in aggregate consumption is 2.48%, about halfway between the $\kappa = 40\%$ economy and the benchmark.

[60] Setting $\rho = 0$ would create the problem of zero consumption is some of the decomposition analyses we conduct below.

[61] Average assets increase by 0.5% relative to the benchmark economy, but only 1% of the population holds exactly zero assets, relative to 8.2% in the benchmark economy.

Figure 8: Consumption Function, Wealth Distribution, High and Low UI



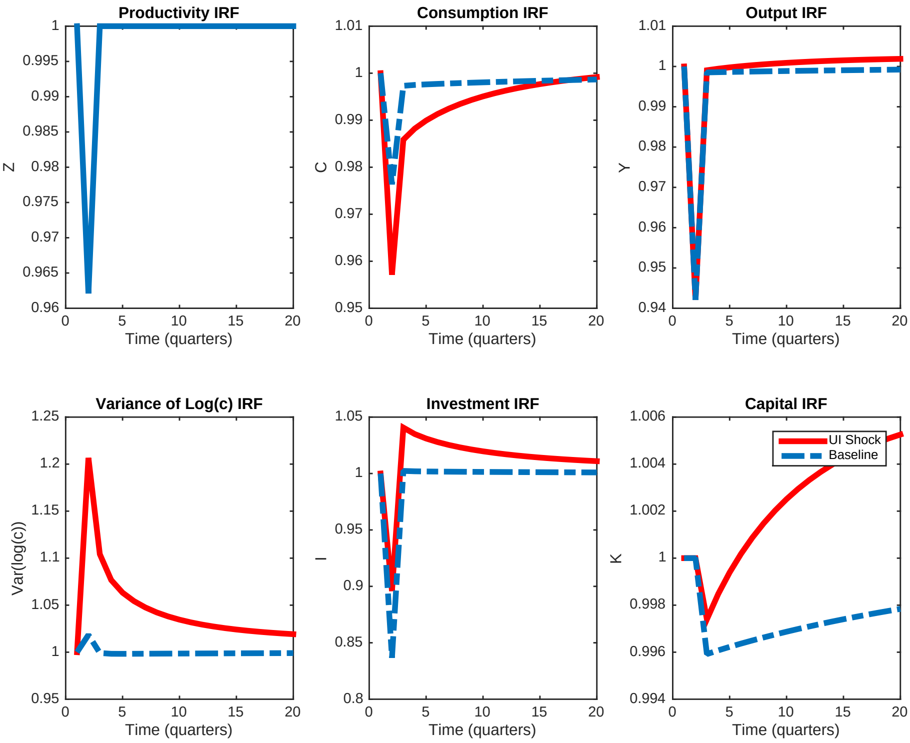
The difference in the consumption decline in a recession across the two economies can then be decomposed into the differential consumption response of households, integrated with respect to the *same* cross-sectional wealth distribution (which is a counterfactual distribution for one of the two economies), and the effect on the consumption response stemming from a policy-induced difference in the wealth distribution coming into the recession. As it turns out, both effects (the change in the consumption functions and the change in the wealth distribution) are quantitatively large, but partially offset each other.

In order to isolate the first effect we now plot, in figure 9, the recession impulse response for the benchmark economy and the economy with low unemployment insurance, but starting at the *same pre-recession wealth distribution* as in the benchmark economy.[62] Under this fixed wealth distribution scenario the consumption response in both cases is given by the difference in the consumption functions (in both panels) integrated with the wealth distribution of the high UE insurance economy. We find that consumption declines much more substantially in the economy with low replacement rate, by 6.24%, relative to 2.64% in the benchmark economy. This is of course exactly what the consumption functions in figure 8 predict.

To further quantify what drives this differential magnitude in the consumption response, in table 12 we display the fall in consumption for 4 groups in the population that differ in their transitions between idiosyncratic employment states as the aggregate economy slips into a recession. The share of households undergoing a specific transition is exogenous and the same across both economies, and is given in the second column of the table. Most households, 88.1% retain their job even though the aggregate economy turns bad. Of particular interest are those households that transition

[62] One can interpret this thought experiment as a surprise permanent removal of the unemployment insurance system exactly in the period in which the recession hits.

Figure 9: Impulse Response to Aggregate Technology Shock without and with Generous Unemployment Insurance, Fixed Wealth Distribution: One Time Technology Shock



from employment into unemployment. Even though the share of these households is relatively small in the population 6.6%, this group accounts for a disproportionately large fraction of the overall consumption collapse in both economies, as the third and forth column of table 12 highlight.

The aggregate consumption decline documented in the last row of the table corresponds to the impulse responses of figure 9. The rows above give the consumption declines accounted for by each of the 4 groups, so that the sum of the rows adds up to the total fall in consumption.

In the benchmark economy (column 3) households with $s = e, s' = u$ comprise 6.6% of the population, but account for 22.7% of the consumption drop and the $s = u, s' = u$ group makes up 1.8% of the population but accounts for 4.5% of the consumption drop. Carrying out the same decomposition for the economy a small unemployment insurance system (column 4) we observe that the total drop in consumption is about 2.4 times as large now, as already displayed in the impulse response plot. Now the $s = u, s' = e$ group accounts for a relatively larger component in the drop in consumption since this group now cuts consumption despite having found a job, in order to avoid hitting the borrowing constraint upon becoming unemployed again in the future.

Table 12: Consumption Response by Group in 3 Economies: Share of Total Decline

Transitions	Share	$\rho = 50\%, \Phi^{\rho=0.5}$	$\rho = 10\%, \Phi^{\rho=0.5}$	$\rho = 10\%, \Phi^{\rho=0.1}$
$s = e, s' = e$	88.1%	76.8%	72.9%	69.9%
$s = e, s' = u$	6.6%	22.7%	21.3%	28.0%
$s = u, s' = e$	3.5%	-3.9%	0.2%	-4.9%
$s = u, s' = u$	1.8%	4.5%	5.6%	7.0%
Total Decline	100%	-2.64%	-6.24%	-3.26%

Finally we document happens if the wealth distribution is determined endogenously and responds to the absence of an unemployment insurance system. Figure 10 displays the impulse responses for the benchmark economy (again) and the no-benefits economy with a pre-recession wealth distribution that emerges in *that economy* after a long period of economic prosperity.[63] Column 5 of table 12 breaks down the consumption response by subgroups. Overall we observe that the endogenous shift in the wealth distribution to the right due to the less generous unemployment insurance largely offsets the larger individual consumption declines in the no-benefits economy for a given wealth level. The end effect is an aggregate consumption (and thus investment) dynamics that is fairly similar between both economies despite the fact that *individual consumption* responses to the crisis differ markedly across the two economies.

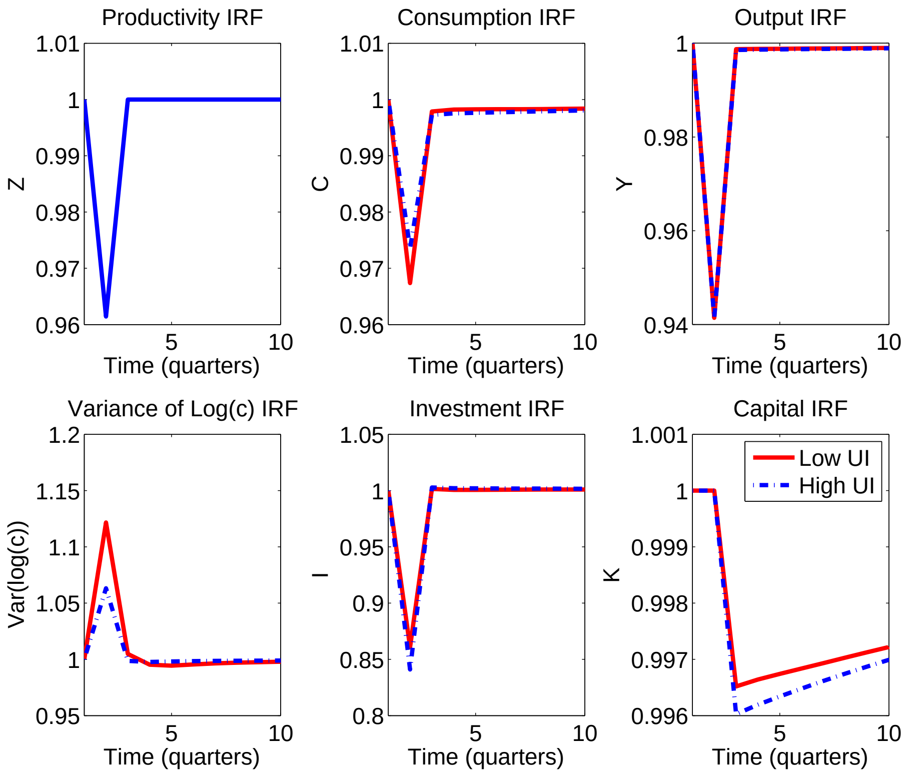
To see this more precisely, compare the third and fifth column of table 12. The aggregate consumption decline in the economy with little unemployment insurance is somewhat larger than in the benchmark economy (by 0.62 percentage points). But very notably, in this economy the unemployed (both newly and already existing ones) account for a substantially larger share of the reduction in consumption, despite the fact that this group understands the possibility of a great recession and has access to self-insurance opportunities to prepare for it.

7 Inequality and Aggregate Economic Activity

In the model studied so far the wealth distribution did potentially have an important impact on the dynamics of aggregate consumption and investment, but by construction a fairly negligible effect on aggregate economic activity. Output depends on capital, labor input and aggregate TFP, and in the previous model the latter two are exogenously given. The capital stock is predetermined in the short run, and even in the medium run only responds to net investment, which is a small fraction of the overall capital stock. So the output response to a negative productivity shock is exogenous on impact and to a first approximation exogenous (to the wealth distribution and to

[63] That wealth distribution was displayed in the right panel of figure 8.

Figure 10: Impulse Response to Aggregate Technology Shock without and with Generous Unemployment Insurance: One Time Technology Shock



social insurance policies) even in the medium run; that is why in the previous section we focused on the distribution of the output decline between aggregate consumption and investment.

We now present a version of the model in which the output response to a negative shock is endogenous even in the short run, and thus potentially depends on the wealth distribution in the economy as well as policies that shape this distribution. The model retains the focus on real, as opposed to nominal, factors. We now consider a world in which $\omega > 0$ and thus TFP $Z^* = ZC^{\omega}$ endogenously responds to the level of aggregate demand. A decline in aggregate consumption triggered by a fall in Z, an ensuing reduction of aggregate wages and household incomes endogenously reduces TFP and thus output further.

In the models discussed so far aggregate demand played no independent role in shaping business cycle dynamics and, by construction, government demand management is ineffective. In this section we present a simple extension of the baseline model with aggregate demand externalities in the spirit of Bai, Rios-Rull and Storesletten (2012), Huo and Rios-Rull (2013), Kaplan and Menzio (2014) and also Den Haan, Rendahl and Riegler (2014), who provide micro foundations for the aggregate productivity process we are assuming here[64]. We now assume that the aggregate production function takes the form: [65]

$$Y = Z^* F(K, N)$$

with $Z^* = ZC^{\omega}$ and $\omega > 0$.

In this model a reduction in aggregate consumption C (say, induced by a negative Z shock) feeds back into lower TFP and thus lower output, deepening the crisis. Thus, in this model, government "demand management" might be called for even in the absence of incomplete insurance markets against idiosyncratic risk. In addition, a social insurance program that stabilizes consumption demand of those adversely affected by idiosyncratic shocks in a crisis might be desirable not just from a distributional and insurance perspective, but also from an aggregate point of view. In the model with consumption externalities, in addition to providing consumption insurance it increases productivity and accelerates the recovery.[66]

We now first discuss the calibration of the extended model before documenting how the presence of the demand externality impacts out benchmark results.

7.1 Calibration Strategy

We retain all model parameters governing the idiosyncratic shock processes (s, y), but recalibrate the *exogenous* part of aggregate productivity Z. In addition we need to specify the strength of the externality ω. Our basic approach is to use direct observations on TFP to calibrate the exogenous process Z and then choose the magnitude of the externaility ω such that the demand externality model displays the same volatility of output as the benchmark model (which, as the reader might recall) was calibrated to match the severity of the two severe recession episode we identified in the data.[67]

Exogenous TFP Process Z For comparability with the benchmark results we retain the transition matrix $\pi(Z'|Z)$ but recalibrate the states (Z_l, Z_h) of the process. To do so we

64 We are certainly not claiming that our and their formulations are isomorphic on the aggregate level; rather, their work provides the structural motivation for the reduced form approach we are taking here.

65 In this paper we abstract completely from nominal frictions that make output partially demand-determined. A representative paper that contains a lucid discussion of the demand- and supply-side determinants of aggregate *output* fluctuations in heterogeneous agent New Keynesian models is Challe, Matheron, Ragot and Rubio-Ramirez (2015)

66 We view this model as the simplest structure embedding a channel through which redistribution affects output directly and in the short run.

67 An alternative approach would have been to retain the original calibration of the Z process, choose a variety of ω values and document how much amplification, relative to the benchmark model, the externality generates. The drawback of this strategy is that output is counterfactually volatile in these thought experiments unless $\omega = 0$.

HP-filter the Fernald (2012) (non-adjusted for capital utilization) data for total factor productivity, identify as severe recessions the empirical episodes with high unemployment as in the benchmark analysis, and then compute average TFP (average % deviations relative to the HP-trend) in the severe recession periods as well as in normal times. This delivers

$$\frac{Z_l}{Z_h} = \frac{1 - 1.84\%}{1 + 0.36\%} = 0.9781$$

Thus, the newly calibrated exogenous TFP process is less volatile than in the benchmark economy, where the corresponding dispersion of TFP was given by $\frac{Z_l}{Z_h} = 0.9614$.

7.1.1 Size of the Spillover ω

Given the exogenous TFP process we now choose ω such that the externality economy has exactly the same output volatility as the benchmark economy. This requires $\omega = 0.30$.

7.2 Results

7.2.1 Aggregate Dynamics

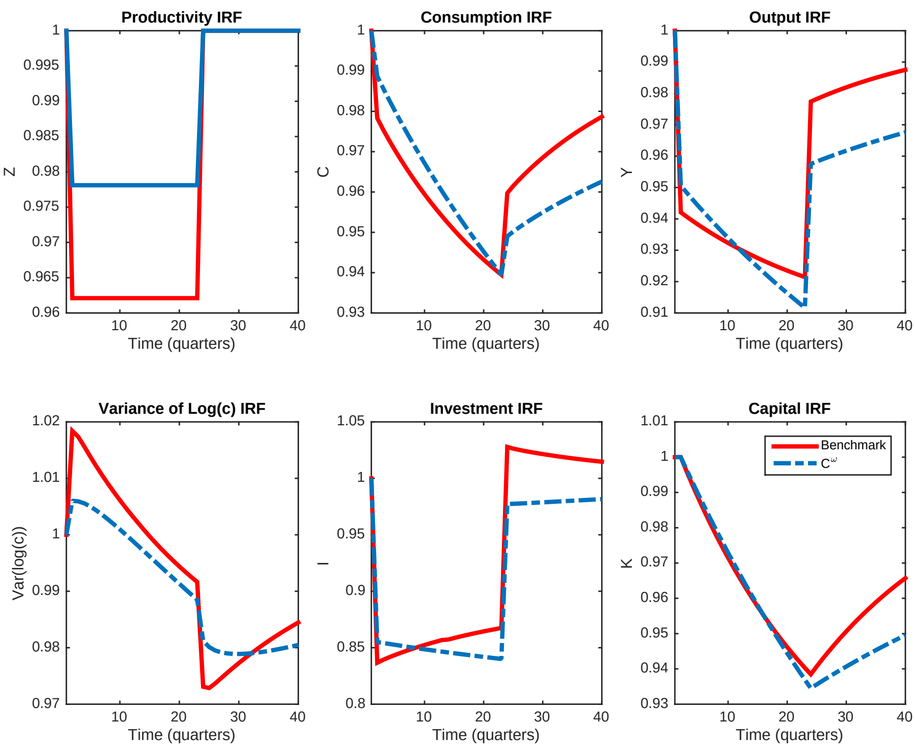
In figure 11 we display the dynamics of a typical great recession (22 quarters of low TFP) in both the baseline economy and the demand externality economy (labeled C^{ω}).[68] The upper right panel shows that, as determined in the calibration section, a significantly smaller exogenous shock is needed in the externality economy to generate a decline in output (and thus consumption and investment) of a given size. The impulse response functions are qualitatively similar in both economies, but with important quantitative differences.

First, the average decline in output in a great recession is the same across both economy since this is how $\frac{Z_l}{Z_h}$ was calibrated in the externality economy. However, since aggregate consumption declines during the course of a great recession and aggregate consumption demand impacts productivity, the decline in output is more pronounced and the recovery slower in the externality economy. Thus, the consumption externality adds endogenous persistence to the model, over and above the one already present through endogenous capital accumulation.

Of course the demand externality mechanism also adds endogenous volatility to the model, but our desire to insure both models have the same output volatility via calibra-

[68] The figure for a one quarter great recession is qualitatively similar, but less useful in highlighting the differences between both economies.

Figure 11: Impulse Response to Aggregate Technology Shock: Comparison between Benchmark and Demand Externality Economy



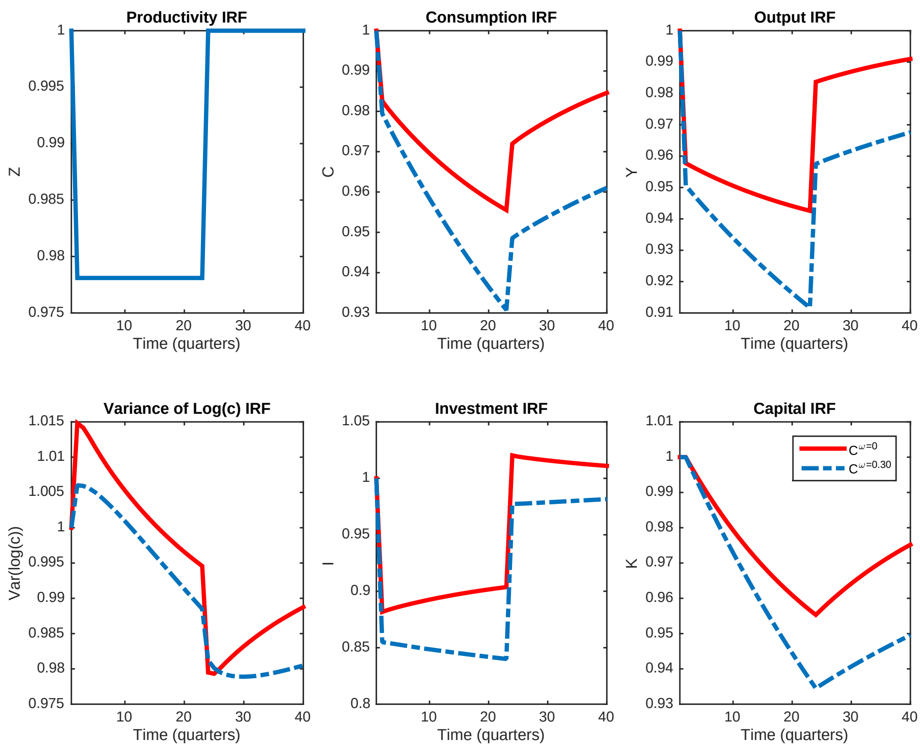
tion obscures this fact. In figure 12 we display the magnitude of this amplification by comparing the impulse responses in two economies with the *same* exogenous TFP process (the one recalibrated for the demand externality model), but with varying degrees of the externality ($\omega = 0$ and $\omega = 0.30$).

In contrast to figure 11, now the differences in the dynamics of the time series are purely driven by the presence of the demand externality. The amplification of the exogenous shock is economically important: the initial fall in output, consumption and investment is substantially larger (5.16%, 2.64% and 13.02% versus 4.23%, 1.98% and 11.23%, respectively). In addition, and consistent with figure 11, these larger output and consumption losses are more persistent in the economy with negative feedback effects from aggregate demand on productivity and production.

7.2.2 On the Interaction of Social Insurance and Wealth Inequality with Demand Externalities

In section 6.3 we demonstrated that the presence of social insurance policies had a strong impact on the aggregate consumption response to an adverse aggregate shock

Figure 12: Impulse Response to Identical Aggregate Technology Shock: Comparison between Economies with and without Demand Externality



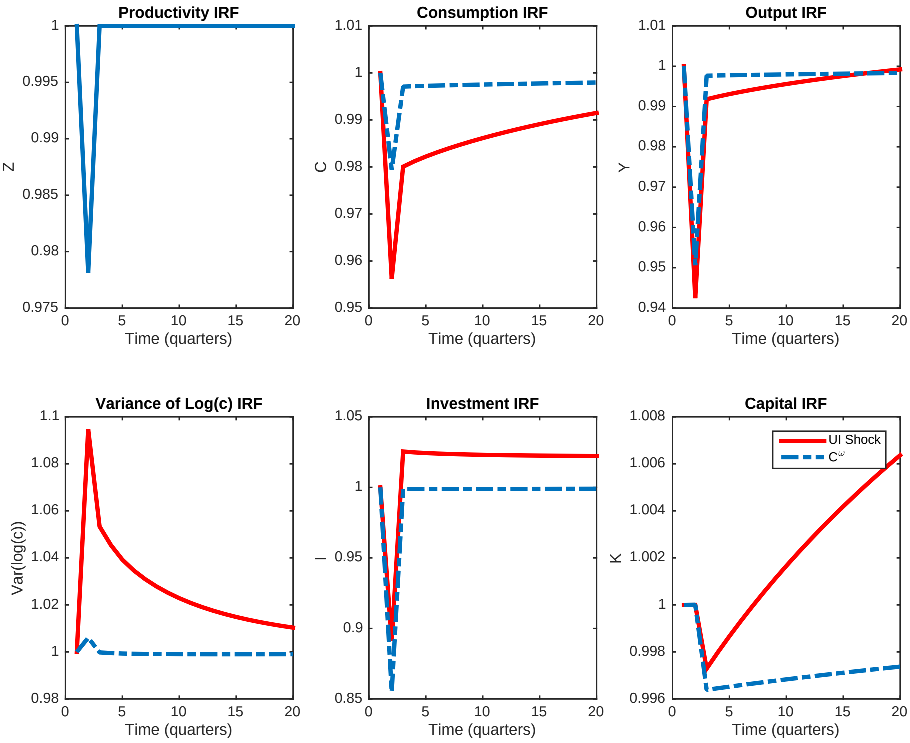
for a given wealth distribution, but also alters the long-run wealth distribution in the economy. With output partially demand-determined, now these policies indirectly impact aggregate productivity and thus output. As the previous figures suggested, the effects are particularly important in the medium run due to the added persistence in the demand externality economy.

In figure 9 above we documented that, holding the wealth distribution fixed, the size of the social insurance system mattered greatly for the aggregate consumption (and thus investment) response to an aggregate productivity shock. Figure 13 repeats the same thought experiment (impulse response to a TFP shock in economies with $\rho = 50\%$ and $\rho = 10\%$ with same pre-recession wealth distribution), but now in the consumption externality model.

The key observations from 13 are that now, in the consumption externality model the size of the unemployment insurance system not only affects the magnitude of the aggregate consumption decline on impact, but also aggregate output, and the latter effect is quite persistent.

This can perhaps more clearly be seen from figure 14 which displays the difference in the impulse response functions for output and consumption between economies with

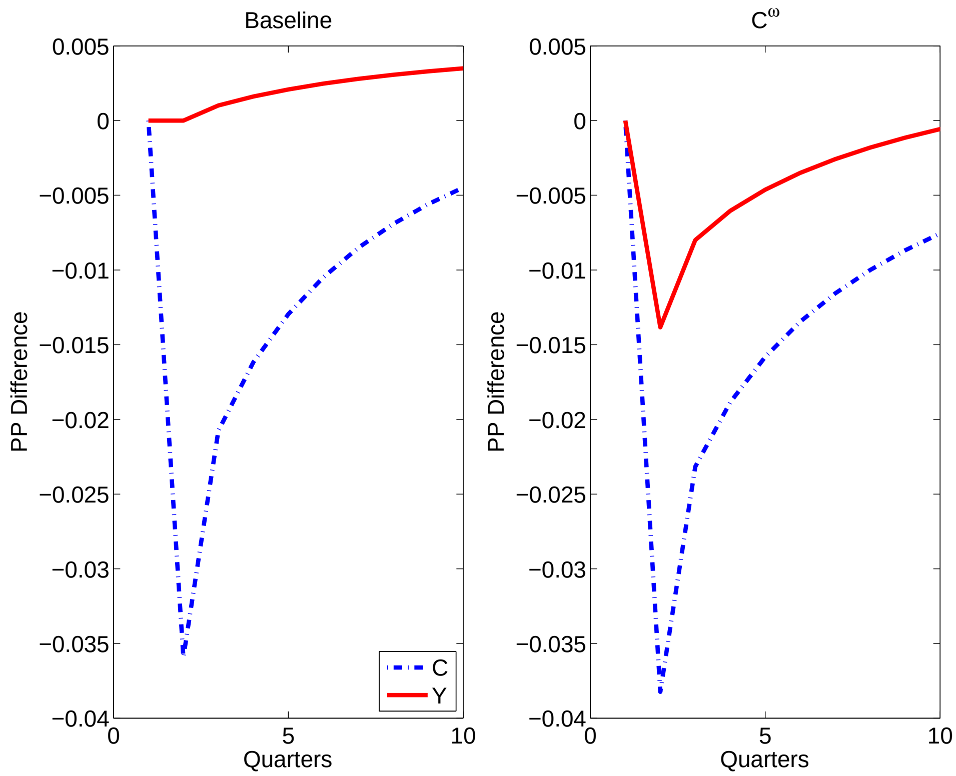
Figure 13: Impulse Response to Aggregate Technology Shock without and with Generous Unemployment Insurance in Consumption Externality Model, Fixed Wealth Distribution



$\rho = 50\%$ and $\rho = 10\%$, both for the benchmark model and the demand externality model. Not only does the presence of sizable unemployment insurance stabilize aggregate consumption more in the externality economy (the UI-induced reduction in the fall of C is 3.9% on impact and 0.8% after ten quarters of the initial shock in the externality economy, relative to 3.6% and 0.5% in the benchmark economy).

In addition, whereas in the benchmark economy more generous social insurance has no impact on output in the short run (by construction) and a moderately negative impact in the medium run (since investment recovers more slowly in the presence of more generous UI), with partially demand-determined output UI stabilizes output significantly (close to 1.5% on impact, with the effect fading away only after 10 quarters -despite the fact that the shock itself only lasts for one quarter in this thought experiment.

Figure 14: Difference in IRF between $\rho = 50\%$ and $\rho = 10\%$ without and with Consumption Externality



8 Related Literature

The literature on macroeconomics with heterogeneous households (or firms) is too large, at this point, to discuss exhaustively.[69] Excellent surveys of different aspects of this literature are contained in Deaton (1992), Attanasio (1999), Krusell and Smith (2006), Heathcote, Storesletten and Violante (2009), Attanasio and Weber (2010), Guvenen (2011) as well as Quadrini and Rios-Rull (2013).

In the paper we have focused on the impact of household heterogeneity in wealth on the aggregate consumption dynamics in large recession, a focus we share with Guerrieri and Lorenzoni (2012), Glover, Heathcote, Krueger and Rios-Rull (2014), Heathcote and Perri (2015) as well as Berger and Vavra (2015). In order to conduct such a study in a quantitatively meaningful way we required model elements that led to an empirically plausible wealth distribution. As Krusell and Smith (1998) and Carroll, Slacalek, Tokuoka and White (2015) we used heterogeneity in time discount factors for this purpose, complemented by a non-standard labor earnings (or labor productivity) process,

[69] This section will be significantly revised and expanded in future versions of this paper. We are grateful for suggestions about papers we might have missed.

as advocated by Castaneda, Diaz-Gimenez and Rios-Rull (2003). Alternative mechanisms include the explicit consideration of entrepreneurial activity, as in the models by Quadrini (1997) and Cagetti and De Nardi (2006), or heterogeneity in investment opportunities or returns as in Benhabib, Bisin and Zhu (2011).

We have also explored the role social insurance policies can play in shaping the aggregate consumption and output response to adverse business cycle shocks in economies with household heterogeneity. As Krusell and Smith (2006) we focused on income insurance programs (and unemployment insurance, more concretely). McKay and Reis (2016) conduct a comprehensive study of automatic stabilization programs on business cycle dynamics, whereas Kaplan and Violante (2014) as well as Jappelli and Pistaferri (2014) study the role of discretionary changes in income taxation on aggregate consumption.[70]

Finally we have explored a class of heterogeneous household models in which output is partially demand determined, but in which there are no nominal frictions, the main focus of the New Keynesian literature.[71] This part of our paper is motivated and builds on the work of Bai et al. (2012), Huo and Rios-Rull (2013), Kaplan and Menzio (2014) and also Den Haan, Rendahl and Riegler (2016).

9 Conclusion

In this chapter we have used PSID data on earnings, income, consumption and wealth as well as different versions of a canonical business cycle model with household earnings and wealth heterogeneity to study under which conditions the cross-sectional wealth distribution shapes the business cycle dynamics of aggregate output, consumption and investment in a quantitatively meaningful way. We have argued that the low end of the wealth distribution is crucial for the answer to this question and have studied mechanisms that helped to generate close to 40% of households without significantly positive net worth, including preference heterogeneity and publicly provided social insurance programs.

---

70 As we do, Auclert (2014) stresses the importance of the heterogeneity in the marginal propensity to consume across households for the impact of redistributive policies. His focus is on monetary policy, however.

71 A representative paper that contains a lucid discussion of the demand- and supply-side determinants of aggregate *output* fluctuations in heterogeneous agent New Keynesian models is Challe, Matheron, Ragot and Rubio-Ramirez (2015). Ravn and Sterk (2013) present a model with labor market frictions and nominal rigidities to study the labor market in the great recession.

References

[1] Aiyagari, R. (1994), "Uninsured Risk and Aggregate Saving", *Quarterly Journal of Economics*, 109, 659-684.

[2] Arellano, M., R. Blundell and S. Bonhomme (2015), "Household Earnings and Consumption: A Nonlinear Framework", *Working paper*.

[3] Ameriks, J., J. Briggs, A. Caplin, M. Shapiro and C. Tonetti (2015), "Long-Term Care Utility and Late in Life Saving, "textitWorking paper.

[4] Athreya, K., A. Owens and F. Schwartzman (2015), "Does Redistribution Increase Output? The Centrality of Labor Supply, ", *Working paper*.

[5] Attanasio, O. (1999), "Consumption", in *Handbook of Macroeconomics*, J. Taylor and M. Woodford (eds.), Elsevier Science.

[6] Attanasio, O. and G. Weber (2010). "Consumption and Saving: Models of Intertemporal Allocation and Their Implications for Public Policy", *Journal of Economic Literature* 48 September: 693-751.

[7] Auclert, A. (2014), "Monetray Policy and the Redistribution Channel," *Working paper*.

[8] Bachmann, R., R. Caballero and E. Engel (2013), "Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model,"*American Journal of Economics: Macroeconomics*, 5, 29-67.

[9] Bai, Y., V. Rios Rull and K. Storesletten (2012), "Demand Shocks as Productivity Shocks", *Working Paper*.

[10] Bayer, C., R. Luetticke, L. Pham-Daoz and V. Tjaden (2015), "Precautionary Savings, Illiquid Assets, and the Aggregate Consequences of Shocks to Household Income Risk,". *Working paper*.

[11] Benhabib, J. and A. Bisin (2016), "Skewed Wealth Distributions: Theory and Empirics", *NBER working paper* 21924.

[12] Benhabib, J., A. Bisin, and S. Zhu, S (2011), "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents ", *Econometrica*, 79(1), 123-157.

[13] Benigno, G. and L. Fornaro , "The Financial Resource Curse ", *Working paper*.

[14] Berger, D. and J. Vavra (2015) "Consumption Dynamics During Recessions," *Econometrica*, 83, 101-154.

[15] Bewley, T. (1986), "Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers", in Werner Hildenbrand and Andreu Mas-Colell (eds.) *Contributions to Mathematical Economics in Honor of Gerard Debreu*, North-Holland, 79-102.

[16] Brinca, P., H. Holter, P. Krusell and L. Malafry (2015), "Fiscal Multipliers in the 21st Century," *Journal of Monetary Economics*, forthcoming.

[17] Buera, F. (2009), "A Dynamic Model of Entrepreneurship with Borrowing Constraints: Theory and Evidence,"textitAnnals of Finance, 5, 443-464.

[18] Cagetti, M., and De Nardi, M. (2006), "Entrepreneurship, Frictions, and Wealth", *Journal of Political Economy*, 114, 835-870.

[19] Carroll, C., J. Slacalek, K. Tokuoka and M. White (2015), "The Distribution of Wealth and the Marginal Propensity to Consume", *Working paper*.

[20] Castaneda, A., J. Diaz-Gimenez and V. Rios-Rull (2003), "Accounting for the U.S. Earnings and Wealth Inequality", *Journal of Political Economy*, 111, 818-857

[21] Challe, E., J. Matheron, X. Ragot and J. Rubio-Ramirez (2015), "Precautionary Saving and Aggregate Demand", *Working Paper*.

[22] Chang, Y. and S.-B. Kim (2007), "Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations", *American Economic Review*, 97, 1939-1956.

[23] Chatterjee, S., D. Corbae, M. Nakajima and V. Rios-Rull (2007), "A Quantitative Theory of Unsecured Consumer Credit with Risk of Default", *Econometrica*, 75, 1525-1590.

[24] Deaton, A. (1992), *Understanding Consumption*, Oxford University Press.

[25] De Nardi, M. (2004) "Wealth Inequality and Intergenerational Links,"*Review of Economic Studies*, 71, 743-768.

[26] De Nardi, M. (2015) "Quantitative Models of Wealth Inequality: A Survey", *NBER working paper* 21106.

[27] De Nardi, M., G. Fella and F. Yang (2015) "Piketty's Book and Macro Models of Wealth Inequality", *NBER working paper* 21730.

[28] De Nardi, M., E. French and J. Jones (2010) "Why Do the Elderly Save? The Role of Medical Expenses, ", *Journal of Political Economy*, 118, 39-75.

[29] den Haan, W. (2010) "Assessing the Accuracy of the Aggregate Law of Motion in Models with Heterogeneous Agents ", *Journal of Economic Dynamics & Control*, 34, 79-99.

[30] den Haan, W., P. Rendahl and M. Riegler (2016), "Unemployment (Fears), Precautionary Savings, and Aggregate Demand", *Working paper*.

[31] Fernald, J. (2012), "Productivity and Potential Output before, during, and after the Great Recession", *Federal Reserve Bank of San Francisco Working Paper* 2012-18.

[32] Fisher, J., D. Johnson, T. Smeeding and J. Thompson (2015), "Inequality in 3D: Income, Consumption and Wealth,", *Working paper*.

[33] Glover, A., J. Heathcote, D. Krueger and V. Rios-Rull (2014), "Inter-generational Redistribution in the Great Recession", *Working paper*.

[34] Gomes, F. and A. Michaelides (2008), "Asset Pricing with Limited Risk Sharing and Heterogeneous Agents", *Review of Financial Studies*, 21, 415-448.

[35] Guerrieri, V. and G. Lorenzoni (2012), "Credit Crises, Precautionary Savings, and the Liquidity Trap", *Working Paper*.

[36] Guvenen, F. (2011), "Macroeconomics with Heterogeneity: A Practical Guide", *Economic Quarterly*, 97, 255-326.

[37] Guvenen, F., S. Ozkan and J. Song (2014), "The Nature of Countercyclical Income Risk", *Journal of Political Economy*, 122, 621-660.

[38] Hagedorn, M., F. Karahan, I. Manovskii and K. Mitman (2013), "Unemployment Benefits and Unemployment in the Great Recession: The Role of Macro Effects", *NBER Working Paper 19499*.

[39] Hagedorn, M., I. Manovskii and K. Mitman (2015), "The Impact of Unemployment Benefit Extensions on Employment: The 2014 Employment Miracle?", *NBER Working Paper 20884*.

[40] Heathcote, J. (2005) "Fiscal Policy with Heterogeneous Agents and Incomplete Markets", textitReview of Economic Studies, 72, 161-188.

[41] Heathcote, J., and F. Perri (2015), "Wealth and Volatility", *NBER Working paper 20994*.

[42] Heathcote, J., Perri, F., and Violante, G. L. (2010), "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967-2006", *Review of Economic Dynamics*, 13, 15-51.

[43] Heathcote, J., K. Storesletten and G. Violante (2009), "Quantitative Macroeconomics with Heterogeneous Households", *Annual Review of Economics*, 1, 319-354.

[44] Heathcote, J., S. Storesletten and G. Violante (2014), "Optimal Tax Progressivity: An Analytical Framework,"*Working paper*.

[45] Hendricks, L. (2007), "How Important Is Discount Rate Heterogeneity for Wealth Inequality?, "*Journal of Economic Dynamics and Control*, 31, 3042-68.

[46] Herkenhoff, K. (2015), "The Impact of Consumer Credit Access on Unemployment,"*Working paper*.

[47] Huggett, M. (1993), "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies", *Journal of Economic Dynamics and Control*, 17, 953-969.

[48] Huggett, M. (1996), "Wealth Distribution in Life-Cycle Economies,"*Journal of Monetary Economics*, 38, 469-494

[49] Huggett, M. (1997), "The One-Sector Growth Model with Idiosyncratic Shocks: Steady States and Dynamics", *Journal of Monetary Economics*, 39, 385-403.

[50] Huo, Z. and V. Rios-Rull (2013), "Paradox of Thrift Recessions ", *Working Paper*.

[51] Imrohoroglu, A. (1989), "Cost of Business Cycles with Indivisibilities and Liquidity Constraints", *Journal of Political Economy*, 97, 1364-1383.

[52] Jappelli, T. and L. Pistaferri (2014), "Fiscal Policy and MPC Heterogeneity", *American Economic Journal: Macroeconomics*, 6, 107-136.

[53] Kaplan, G. and G. Menzio (2014), "Shopping Externalities And Self-Fulfilling Unemployment Fluctuations", *Journal Of Political Economy*, forthcoming.

[54] Kaplan, G., B. Moll and G. Violante (2015), "The Macroeconomy According to HANK ", *NBER Working paper*

[55] Kaplan, G. and G. Violante (2014), "A Model of the Consumption Response to Fiscal Stimulus Payments", *Econometrica*, 82, 1199-1239.

[56] Khan, A. and J. Thomas (2008), "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics, ", *Econometrica*, 76, 395-436.

[57] Kopecky, K. and R. Suen (2010), "Finite State Markov-Chain Approximations to Highly Persistent Processes, ", *Review of Economic Dynamics*, 13, 701-714.

[58] Krueger, D., K. Mitman and F. Perri (20106) "Inequality and the Macro Economy: Data, Models and Policy Implications ", *Proceedings of the Econometric Society World Congress 2015*, forthcoming.

[59] Krusell, P, T. Mukoyama and A. Sahin (2010), "Labour-Market Matching with Precautionary Savings and Aggregate Fluctuations", *Review of Economic Studies*, 77, 1477-1507.

[60] Krusell, P, T. Mukoyama, A. Sahin and A. Smith (2009), "Revisiting the Welfare Effects of Eliminating Business Cycles, ", *Review of Economic Dynamics*, 12 (3), 393-404

[61] Krusell, P. and A. Smith (1997),"Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns", *Macroeconomic Dynamics*, 1, 387-422.

[62] Krusell, P. and A. Smith (1998), "Income and Wealth Heterogeneity in the Macroeconomy", *Journal of Political Economy*, 106, 867-896.

[63] Krusell, P. and T. Smith (2006), "Quantitative Macroeconomic Models with Heterogeneous Agents", in *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*.

[64] Livshits, I., J. MacGee and M. Tertilt (2007), "Consumer Bankruptcy - A Fresh Start", *American Economic Review*, 97, 402-418..

[65] Maliar, L., S. Maliar and F. Valli (2010), "Solving the Incomplete Markets Model with Aggregate Uncertainty using the Krusell-Smith Algorithm", *Journal of Economic Dynamics and Control*, 34, 42-49

[66] McKay, A. and R. Reis (2016), "The Role of Automatic Stabilizers in the U.S. Business Cycle", *Econometrica*, forthcoming.

[67] Mitman, K. and S. Rabinovich (2014), "Do Unemployment Benefit Extensions Explain the Emergence of Jobless Recoveries?", *Working Paper*.

[68] Quadrini, V. (1999), "Entrepreneurship, Saving and Social Mobility", *Review of Economic Dynamics*, 3, 1-40.

[69] Quadrini, V. and V. Rios-Rull (2013), "Inequality in Macroeconomics", in A.B. Atkinson and F.J. Bourguignon (Eds.), *Handbook of Income Distribution*, Volume 2B, North Holland.

[70] Ravn, M. and V. Sterk (2013), "Job Uncertainty and Deep Recessions", *Working Paper*.

[71] Schaal, E. and M. Taschereau-Dumouchel (2014), "Coordinating Business Cycles", *Working Paper*.

[72] Shimer, R. (2005), "The Cyclical Behavior of Equilibrium Unemployment and Vacancies", *American Economic Review*, 95, 25-49.

[73] Smith, M. and C. Tonetti (2014), "A Bayesian Approach to Imputing a Consumption-Income Panel Using the PSID and CEX", *Working Paper*

[74] Storesletten, K., C. Telmer and A. Yaron (2004), "Cyclical Dynamics in Idiosyncratic Labor-Market Risk", *Journal of Political Economy*, 112, 695-717.

[75] Storesletten, K., C. Telmer and A. Yaron (2007), "Asset Pricing with Idiosyncratic Risk and Overlapping Generations," *Review of Economic Dynamics*, 10, 519-548.

[76] Young, E. (2010), "Solving the Incomplete Markets Model with Aggregate Uncertainty using the Krusell-Smith Algorithm and Non-Stochastic Simulations, "*Journal of Economic Dynamics and Control*, 34, 36-41

A Data and Estimation Appendix

A.1 Construction of Facts from Section 2

The series for disposable income from the the BEA is Disposable Personal Income minus medicare and medicaid transfers, which are not reported in PSID. The disposable income series from PSID is constructed adding, for each household and from all members, wage and salary income, income from business and farm, income form assets (including the rental equivalent for the main residence for home owners), all money transfers minus taxes (computed using the NBER TAXSIM calculator). The series for consumption expenditures (both from the BEA and PSID) include the following expenditures categories: cars and other vehicles purchases, food (at home and away), clothing and apparel, housing (including rent and imputed rental services for owners), household equipment, utilities, transportation expenses (such us public transportation and gasoline), recreation and accommodation services. In PSID imputed rental services from owners are computed using the value of the main residence times an interest rate of 4%. Total consumption expenditures are reported for a two year period because of the timing of reporting in PSID. In PSID some expenditures categories (food, utilities) are reported for the year of the interview, while others are reported for the year preceding the interview, so total expenditures span a two year period. The measure of total consumption the BEA is constructed aggregating using the different categories using PSID timing, so, for example, total expenditures in 2004-2005 include car purchases from 2004 and food expenditures from 2005. We have excluded health services as PSID only report out of pocket expenditures and insurance premia. All PSID observations are aggregated using sample weights. Table A1 reports the 2004 levels of the per capital variables plotted in figure 1, along side, for comparison purposes, with the level of food expenditures from both sources and of total houshold personal consumption expenditures from the the BEA

Table A1. Per capita levels in 2004: BEA v/s PSID

	BEA	PSID
1.Disposable income	$24120	$21364
2.Personal Consumption (PSID aggregate)	$18705	$15889
3.Food Expenditures	$3592	$2707
4.Personal Consumption (Total)	$27642	-

The table suggests that the levels from PSID and from the BEA are not too far off, although there are differences. In particular the aggregated PSID data is different from the aggregates from BEA for two reasons. Comparing lines 2-3 across columns we see that, for a given category the average from PSID is different (typically lower) than what reported from the BEA. This discrepancy between aggregate and aggregate survey data has been widely documented before. The second reason is that some categories are just not included in our PSID aggregate, either because mis-measured in

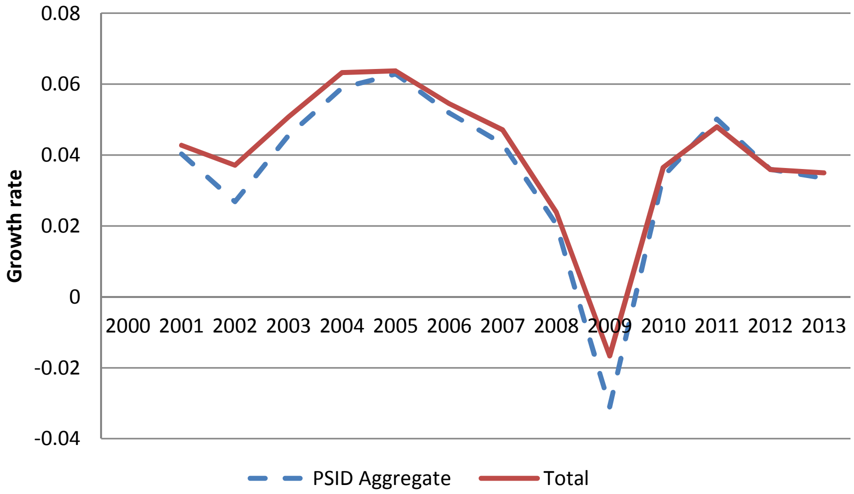


Figure 15: BEA Consumption growth for two different aggregates

PSID (Health expenditures) or because not reported by PSID (Expenditure in Financial Services). One might wonder whether this omitted categories matter for the aggregate pattern of expenditures. Figure A1 reports the growth rate of total household personal consumption expenditures from the the BEA, along with the growth rate for the BEA consumption expenditures that are included in the PSID aggregate defined above. The table above suggest that categories included in PSID aggregate only cover about 65% of the total consumption expenditures; the figure though shows that the cyclical pattern of total expenditures is similar to the one in the PSID.aggregate, suggesting that the missing consumption categories in the the PSID aggregate should not make a difference for our results.

Table A2. Changes in selected variables across the PSID net worth (2006-2008)

		%			% Expend. Rate (pp)	
NW Q	Net Worth[a]	Earnings	Disp Y	Expend.	Earnings	Disp Y
Q1	15.5k (NA)	17.8	17.1	-1.1	-16.5	-15.0
Q2	34.4k (140%)	15.4	16.4	5.2	-8.0	-8.3
Q3	31.1k (39%)	9.9	7.6	-4.5	-10.0	-7.7
Q4	42.2k (19.0%)	7.8	6.4	-3.6	-8.3	-6.2
Q5	-265.7k (-23.9%)	1.5	-0.9	-15.4	-12.2	-8.0

[a] In 000s of dollars. Percentage change (when possible to calculate it) in parenthesis

Table A3. Changes in selected variables across the PSID net worth (2008-2010)

		%			% Expend. Rate (pp)	
NW Q	Net Worth[a]	Earnings	Disp Y	Expend.	Earnings	Disp Y
$Q1$	+28.8k (NA)	9.8	10.4	3.5	-4.9	-5.0
$Q2$	+13.1k (+138%)	-0.1	2.7	7.3	6.3	3.5
$Q3$	+20.7k (+41%)	-0.9	2.1	2.2	2.1	0.0
$Q4$	+41.5k (+25%)	-0.5	0.1	5.3	3.9	3.0
$Q5$	-77.2k (-8%)	-5.6	-4.9	-1.4	2.7	1.8

[a] In 000s of dollars. Percentage change (when possible to calculate it) in parenthesis

A.2 Estimation of Earnings Process

TBC

B Theoretical Appendix

B.1 Explicit Statement of Aggregate Law of Motion for Distribution

Since the extent of heterogeneity and the choice problem of young and old households differ significantly it is easiest to separate the cross-sectional probability measure Φ into two components (Φ_W, Φ_R), and note that the measures integrate to Π_W and Π_R, respectively. First define the Markov transition function, conditional on staying in the young age group $j = W$ as

$$Q_{W,(Z,\Phi,Z')}((s,y,a,\beta),(\mathcal{S},\mathcal{Y},\mathcal{A},\mathcal{B})) = \sum_{s'\in\mathcal{S},y'\in\mathcal{Y}} \begin{cases} \pi(s'|s,Z',Z)\pi(y'|y): & a'_W(s,y,a,\beta;Z,\Phi)\in\mathcal{A},\beta\in\mathcal{B} \\ 0 & else \end{cases}$$

and for the old, retired age group, as

$$Q_{R,(Z,\Phi,Z')}((a,\beta),(\mathcal{A},\mathcal{B})) = \begin{cases} 1: & a'_R(a,\beta;Z,\Phi)\in\mathcal{A},\beta\in\mathcal{B} \\ 0 & else \end{cases}$$

For each Borel sets $(\mathcal{S},\mathcal{Y},\mathcal{A},\mathcal{B}) \in P(\mathcal{S}) \times P(\mathcal{Y}) \times B(\mathcal{A}) \times P(\mathcal{B})$, the cross-sectional probability measures of the young and old tomorrow are then given by[72]

[72] These expressions captures the assumption that in each period a measure $1 - nu$ of newborn households enter the economy as workers, with zero assets and with idiosyncratic productivities and discount factors drawn from the stationary distributions, and that a fraction $1 - \theta$ of working households retire, and that the retirement probability is independent of all other characteristics.

$$H_W(Z,\Phi,Z')(\mathcal{S},\mathcal{Y},\mathcal{A},\mathcal{B}) = \theta \int Q_{W,(Z,\Phi,Z')}((s,y,a,\beta),(\mathcal{S},\mathcal{Y},\mathcal{A},\mathcal{B}))d\Phi_W + (1-\nu)\mathbf{1}_{\{0\in\mathcal{A}\}} \sum_{s'\in\mathcal{S}} \Pi_Z(s') \sum_{y'\in\mathcal{Y}} \Pi(y') \sum_{\beta'\in\mathcal{B}} \Pi(\beta')$$

and

$$H_R(Z,\Phi,Z')(\mathcal{A},\mathcal{B}) = \nu \int Q_{R,(Z,\Phi,Z')}((a,\beta),(\mathcal{A},\mathcal{B}))d\Phi_R + (1-\theta) \int Q_{W,(Z,\Phi,Z')}((s,y,a,\beta),(S,Y,\mathcal{A},\mathcal{B}))d\Phi_W$$

C Computational Appendix

The computational strategy follows the framework developed initially in Krusell and Smith (1998), which was further adapted by Storesletten, Telmer and Taron (2007) and Gomes and Michaelides (2008). In particular we employ the computational strategy outlined in Maliar, Maliar and Valli (2010), focusing on the non-stochastic simulation algorithm first introduced by Young (2010).

C.1 The individual problem

We approximate the true aggregate state (S=(Z,Φ)) by $\hat{S}$, whose specific form depends on which version of the model we solve, which is detailed explicitly below. Thus, the household state is determined by $(s,y,a,\beta;\hat{S})$ in working life and $(a,\beta;\hat{S})$ when retired.

The solution method from Maliar, Maliar and Valli (2010) is an Euler-equation algorithm which takes into account occasionally-binding borrowing constraints. The problem to be solved is:

Retired:

$$\begin{aligned}
c_R(a,\beta;\hat{S})^{-\sigma} - \lambda &= \nu\beta\mathbb{E}[(1-\delta+r'(\hat{S}'))c'_R(a'_R,\beta;\hat{S}')^{-\sigma}] \\
a'_R(a,\beta;\hat{S}) + c_R(a,\beta;\hat{S}) &= b_{SS}(\hat{S}) + (1+r(\hat{S})-\delta)a/\nu \\
a'_R(a,\beta;\hat{S}) &\geq 0 \\
\lambda &\geq 0, \qquad \lambda a'_R(a,\beta;\hat{S}) = 0
\end{aligned}$$

Working:

$$c_W(s,y,a,\beta;\hat{S})^{-\sigma} - \lambda = \theta\beta\mathbb{E}[(1-\delta+r'(\hat{S}'))c'_W(s',y',a'_W,\beta;\hat{S}')^{-\sigma}]$$

$$
\begin{aligned}
& +(1-\theta)\beta\mathbb{E}[(1-\delta+r'(\hat{S}'))c'_R(a'_W,\beta;\hat{S}')^{-\sigma}] \\
a'_W(s,y,a,\beta;\hat{S})+c(s,y,a,\beta;\hat{S}) &= (1-\tau(Z;\rho))w(\hat{S})y\,[1-(1-\rho)1_{s=u}]+(1+r(\hat{S})-\delta)a \\
a'_W(s,y,a,\beta;\hat{S}) &\geq 0 \\
\lambda &\geq 0, \qquad \lambda a'_W(s,y,a,\beta;\hat{S})=0
\end{aligned}
$$

where λ is the Lagrange multiplier on the borrowing constraint.

We eliminate consumption via the budget constraint and then guess a policy rule for $a'_W(s,y,a,\beta;\hat{S})$ and $a'_R(a,\beta;\hat{S})$. We then substitute the policy rule to compute $a''_W(s',y',a'_W,\beta;\hat{S}')$, $a''_R(a'_W,\beta;\hat{S}')$ and $a''_R(a'_R,\beta;\hat{S}')$, and use the Euler equation to back out the implied policy rule for a'. If the implied policy rule is the same as the conjectured policy rule, we have computed the optimal policy, if not we update the guess and repeat.

C.2 The simulation algorithm

In order to simulate the model we pick a grid on $\mathcal{A}$ and fix a distribution of workers $\Phi_0 \in S \times Y \times A \times B$ space. We fix a long time series for the realization of the aggregate shock, Z. Using the realization Z_t and Φ_t we can compute $\hat{S}_t$ and then apply the policy rules from the individual problem, and the markov transition matrices associated with s and y, to compute Φ_{t+1} by interpolating onto the grid points in $\mathcal{A}$.

C.3 Approximating the Aggregate Law of Motion

C.3.1 KS and Benchmark Economies

For the KS and benchmark economies we approximate the true aggregate state with $\hat{S}_t = (Z,\bar{K}_t)$ where $\bar{K}_t$ is the average capital in the economy. Agents need to forecast the evolution of the capital stock. We conjecture that that law of motion in capital depends only on the Z and $\bar{K}$:

$$\log(\bar{K}_{t+1}) = a_0(Z_t) + a_1(Z_t)\log(\bar{K}_t)$$

We conjecture coefficients a_0 and a_1, solve the household problem and simulate the economy. Then, using the realized sequence of $\hat{S}$ we perform the regression above and check whether the implied coefficients are the same as the conjectured ones. If they are we have found the law of motion, if not we update our guess and repeat.

For the KS economy, the computed law of motion is:

$$
\begin{aligned}
\log(\bar{K}_{t+1}) &= 0.1239 + 0.9652\log(\bar{K}_t) \qquad \text{if } Z_t = Z_l \\
\log(\bar{K}_{t+1}) &= 0.1334 + 0.9638\log(\bar{K}_t) \qquad \text{if } Z_t = Z_h
\end{aligned}
$$

The R^2 for both regressions are in excess of 0.999999. Note, however, that den Haan (2010) points out that despite having large R^2 values, the accuracy of the solution can still be poor, and suggests simulation the capital stock under the policy rule and comparing it to the capital stock that is calculated by aggregating across the distribution. We do this for 3000 time periods. The average error between the implied law of motion from the forecast equations and the computed law of motion is 0.02%, with a maximum error of 0.10%.

For the benchmark economy the computed law of motion is:

$$\begin{aligned}\log(\bar{K}_{t+1}) &= 0.0924 + 0.9716\log(\bar{K}_t) \qquad \text{if } Z_t = Z_l \\ \log(\bar{K}_{t+1}) &= 0.0929 + 0.9723\log(\bar{K}_t) \qquad \text{if } Z_t = Z_h\end{aligned}$$

The R^2 for both regressions are in excess of 0.99999. Similar to above, we check the accuracy of the law of motion. We find that the average error between the implied law of motion and the actual capital stock computed from the distribution is 0.01%, with a maximum error of 0.07%.

C.3.2 Consumption Externality Economy

In the economy with the aggregate consumption externality, we add contemporaneous consumption as a state variable in our approximation of the true aggregate state, $\hat{S} = (Z, \bar{K}, C)$. We therefore need an additional law of motion for how aggregate consumption evolves. We conjecture the same form of law of motion for the average capital stock, however, we allow the evolution of aggregate consumption to depend on both the average capital stock and aggregate consumption:

$$\begin{aligned}\log(\bar{K}_{t+1}) &= a_0(Z_t) + a_1(Z_t)\log(\bar{K}_t) \\ \log(C_{t+1}) &= b_0(Z_t, Z_{t+1}) + b_1(Z_t, Z_{t+1})\log(\bar{K}_t) + b_2(Z_t, Z_{t+1})\log(C_t)\end{aligned}$$

Note that because capital is predetermined in the current period, the forces rule for capital depends only contemporaneous variables. Because aggregate consumption is an equilibrium outcome in the next period, we allow for the forecast to depend on subsequent period's realization of the Z shock. Thus, there are four sets of coefficients to be estimated for the law of motion for consumption. The computed forecast equations are:

$$\log(\bar{K}_{t+1}) = 0.0872 + 0.9736\log(\bar{K}_t) \qquad \text{if } Z_t = Z_l$$

$$\log(\bar{K}_{t+1}) = 0.0626 + 0.9816 \log(\bar{K}_t) \qquad \text{if } Z_t = Z_h$$

and

$$\begin{aligned}
\log(C_{t+1}) &= -0.0205 + 0.0023 \log(\bar{K}_t) + 0.9675 \log(C_t) && \text{if}(Z, Z') = (Z_l, Z_l) \\
\log(C_{t+1}) &= -0.5061 + 0.2882 \log(\bar{K}_t) + 0.5297 \log(C_t) && \text{if}(Z, Z') = (Z_l, Z_h) \\
\log(C_{t+1}) &= -0.3560 + 0.1893 \log(\bar{K}_t) + 0.6626 \log(C_t) && \text{if}(Z, Z') = (Z_h, Z_l) \\
\log(C_{t+1}) &= -0.0506 + 0.0360 \log(\bar{K}_t) + 0.9295 \log(C_t) && \text{if}(Z, Z') = (Z_h, Z_h)
\end{aligned}$$

with R^2 in excess of 0.9999, 0.9999999, 0.9999, 0.9999, 0.99999, 0.99999, respectively. As before, we check the accuracy of the two laws of motion. We find that the average error between the implied law of motion and the actual capital stock computed from the distribution is 0.02%, with a maximum error of 0.30%, for the path of aggregate consumption the mean error is 0.02% with a maximum error of 0.24%. While the externality economy has slightly larger forecast errors, the fit of the predicted aggregates is still excellent.

C.4 Digression: Why Quasi-Aggregation?

One of the implications of the results in the main text is that the wealth distribution (and especially the fraction of the population with little or no wealth) is quantitatively important for the macroeconomic consumption and investment response to an aggregate technology shock. This, however, does not imply that Krusell and Smith's (1998) original quasi-aggregation result fails.[73] Recall that this result states that only the mean of the current wealth distribution (as well as the current aggregate shock Z) is required to accurately predict the future capital stock and therefore future interest rates and wages.

The previous experiment compared consumption and investment dynamics in *two economies* that differed substantially in their wealth distributions. For a given economy, if the wealth distribution does not move significantly in response to aggregate shocks, then it would be irrelevant for predicting future aggregates and prices. However, in the high wealth-inequality economy the wealth distribution *does* move over the cycle. For example, the share of households at the borrowing constraint displays a coefficient of variation of 7%. However, what is really crucial for quasi-aggregation to occur is whether the movement, over the cycle, in the key features of the wealth distribution is explained well by movements in Z and K, the state variables in the forecast equations of households. We find that it is, even in the high wealth inequality economy.

[73] In fact, our computational method that follows theirs rather closely relies on quasi-aggregation continuing to hold.

For example, if we regress the fraction of people at the borrowing constraint tomorrow on Z in simulated data, we obtain an R^2 of around 0.8. Therefore the vast majority of the variation in households at the borrowing limit is very well predicted by the aggregate state variables (Z, K). This finding is robust to alternative definitions of constrained households (households exactly at wealth 0, households who save less than 1%, less than 10% or less than 25% of the quarterly wage) and alternative moments of the wealth distribution. It is this finding that makes quasi-aggregation to hold, despite the strong impact of the wealth distribution on the aggregate consumption and investment response to an aggregate technology shocks.

C.5 Recovering the Value Function

As we solve the model by exploiting the euler equation, in order to perform the welfare calculations in section 6.4, we need to recover the value functions as a function of the idiosyncratic and aggregate states. To calculate them we use policy function iteration. We make an initial guess for the value function, v^0, then calculate v^1 by solving the recursive household decision problem (we need not perform the maximization, as we have already computed the optimal policy function). We approximate the value function with a cubic spline interpolation in assets, as well as in aggregate capital (and for the demand externality model also aggregate consumption). If v^1 is sufficiently close to v^0 (in the sup-norm sense) we stop, otherwise we proceed to compute v^2 taking v^1 as the given value function. We proceed until convergence. For the economies with retirement, we first recover the value function for retired households, v_R and then proceed to recover the value function for working age households, v_W