Heterogeneity in Consumption Responses to Long-lasting Income Shocks

Yu Zhang∗

March 24, 2015

Abstract

Standard life-cycle incomplete-markets models predict large cross-sectional differences for age and wealth in agents’ consumption responses to long-lasting income shocks. Whether these predictions hold in the data is an open question. Using household microdata for the US, I document economically important age and wealth cross-sectional heterogeneity in the adjustment of nondurable consumption in the face of long-lasting income shocks. In response to labor-income shocks that last at least three years, households with heads younger than 50 adjust consumption by twice as much as others. Households with less liquid wealth than median adjust consumption by twice as much as others. A calibrated standard life-cycle incomplete-markets model predicts heterogeneity in consumption responses that are quantitatively similar to empirical estimates.

∗Zhang: Department of Economics, Princeton University, 001 Fisher Hall, Princeton, New Jersey 08544, United States (email: zhangyu@princeton.edu). First draft: November 25, 2014. I am extremely grateful to Richard Rogerson and Greg Kaplan for their support and detailed advice. I also thank Mark Aguiar, Gabriel Chodorow-Reich, Liang Dai, Emil Verner, Zongbo Huang, Nobuhiro Kiyotaki, Jason Ravit, Christopher Sims, Chang Sun, Yulong Wang, and seminar participants at the Princeton macroeconomics workshop for helpful comments and conversations. All errors are mine.
1 Introduction

Incomplete markets models in the spirit of Aiyagari (1994), Bewley (1986) and Huggett (1993) are one of the benchmark models in macroeconomics and predict that households are not able to perfectly smooth consumption in the face of income shocks. Building on this prediction, incomplete markets models are used widely to study income inequality and redistribution. In this regard, the ability to match empirical evidence on the joint behavior of income and consumption is a key requisite if incomplete markets models are to deliver reliable predictions.

Existing results suggest that incomplete markets models do well in producing the average level of consumption responses to income shocks. In two recent studies, Blundell et al. (2008) (hereafter referred to as BPP) use data from a panel of US households to estimate the degree to which households adjust consumption in the face of long-lasting income shocks, and Kaplan and Violante (2010) (hereafter KV2010) report that a calibrated life-cycle version of the incomplete markets model can rationalize the average estimate of consumption responses found in BPP.

But there are other issues for which the heterogeneity in consumption responses matters. First, proposed policies can be non-uniform in nature. For example, there is recent interest in age-dependent taxation of income and joint taxation of capital and labor, put forward by Erosa and Gervais (2002), Albanesi and Sleet (2006), Weinzierl (2011), and Farhi and Werning (2013). Second, the predictions of incomplete markets models may be important beyond the aggregate level. For example, Heathcote et al. (2010) use a calibrated incomplete markets model to show that households are affected differently by changes in the wage distribution. In both cases, the reliability of the model’s welfare implications depends on whether the model produces the correct distribution of consumption responses in the face of the proposed policies or structural changes.

Little is known, however, about the ability of incomplete markets models to predict the right distribution of consumption responses. The KV2010 model generates two sharp predictions about cross-sectional differences in consumption responses to long-lasting income shocks: the response is smaller for older households and for wealthier households. Yet BPP find no significant statistical evidence to support these predictions. This tension leads KV2010 to the preliminary conclusion that the model might be misaligned with the data in the cross-sections of consumption responses. The object of this paper is to revisit this tension.

To do this, I use an extended PSID sample augmented with CEX expenditure information, a refined estimator, and parameter stability tests. I find statistically significant and economically important age and wealth heterogeneity in how households adjust consumption in face of long-lasting shocks to labor income. While younger households adjust log nondurable consumption by more than 50% of the size of a long-lasting shock to log after-tax labor income, households with heads older than 50 change log nondurable consumption by 25% or less of the shock size. Along the wealth dimension, households with

---

1See, for example, Storesletten et al. (2004), Heathcote et al. (2010), and Kaplan (2012a).
3In addition to the negative results in BPP, Etheridge (2014), using British data, also finds insignificant age differences in consumption responses.
a larger fraction of total income from liquid assets (which I interpret as a proxy for a larger liquid-wealth to labor-income ratio), are more than twice as capable of smoothing consumption over long-lasting shocks to log after-tax labor income than households with little to no income from liquid assets. My empirical results thus suggest the presence of large cross-sectional differences in the access of US households to private insurance against long-lasting income shocks. Such differences can be important for studying welfare and for designing public insurance and taxation, as discussed, for example, in Heathcote and Tsuijyama (2014).

Comparing model predictions with my empirical results, I find that incomplete markets models are in fact able to predict the correct cross-sections in consumption responses. When calibrated similarly to that of KV2010, the life-cycle incomplete markets model produces age and wealth heterogeneity in consumption responses to long-lasting income shocks that are quantitatively similar to those I find in the data. Additional features of the calibration to that of KV2010 are an empirical value of income persistence and a finer targeting of the wealth distribution. The model embeds two separate channels that allow for heterogeneous consumption responses to long-lasting income shocks. First, there is a horizon effect, in the sense that labor-income risks are truncated at retirement. Second, there is a buffer-stock effect, in the sense that households with high financial wealth relative to labor income are better buffered against income shocks of a given size. I demonstrate that the importance of both channels is supported by the data. On the one hand, in a subsample that excludes high-wealth households, there is still a decreasing age profile of consumption responses, providing evidence for the horizon effect. On the other hand, in a subsample that excludes agents near retirement, the proxy for the liquid-wealth to labor-income ratio remains negatively correlated with consumption responses, supporting the buffer-stock effect.

The identification of consumption responses to long-lasting income shocks in this paper follows and refines the pioneering work of BPP, who construct a panel of household income and consumption from the PSID and the CEX, and document partial insurance. This study has three key differences from BPP, which fails to find significant heterogeneity in consumption responses to long-lasting shocks to labor income.4

First, I extend the PSID sample used in BPP (1980–92) to cover all years during which the PSID was carried out annually (1967–96), which nearly doubles the effective sample size. Following Guvenen and Smith (2013), I employ information from early waves of the CEX in 1972 and 1973 to impute consumption in the PSID before 1980.

Second, I use an alternative estimator, a refined version of the IV estimator proposed in KV2010. Unlike the minimum distance estimator adopted in BPP, this estimator treats year conditional variances of shocks to income as nuisance parameters. This makes the estimator more efficient in estimating consumption responses, especially in panels with a large number of years but not necessarily a large number of observations per year.

Third, to establish heterogeneity in consumption responses I apply parameter stability tests in the spirit of Andrews (1993) and Andrews and Ploberger (1994), instead of stratifying the sample using arbitrary age/wealth cutoffs. These parameter stability tests search over a variety of break points in age and wealth

---

4BPP does document that low-wealth households have significant consumption responses to transitory shocks to labor income, while high-wealth households do not.
and take into account uncertainty in the estimates caused by the arbitrary nature of break points, thus preventing spurious rejection of homogeneity by using discretionary cutoffs.

This paper also expands on the empirical literature on the cross-section of consumption responses to shocks. Notable examples include Jappelli and Pistaferri (2013), who document rich heterogeneity in consumption responses to transitory government transfers, and Mian et al. (2013), who focus on wealth heterogeneity in consumption responses to shocks to household net worth.5

In the next Section, I define the transmission coefficient of long-lasting labor income shocks to consumption—the measure of consumption responses used in this paper. In Section 3, I describe the data and the empirical method used in estimating age and wealth patterns in the transmission coefficient. I report my empirical results in Section 4. In Section 5, I present a calibrated life-cycle incomplete markets model and compare the heterogeneity in consumption responses in the model and in the data. Section 6 concludes.

2 Theory

Following BPP and KV2010, I assume that real (log) net labor income, log $Y$, can be decomposed into an anticipated component $\kappa$, a persistent component $z$, and a transitory component $\nu$. Hence, the income process of each household $i$ is (subscript $t$ denotes age):

$$\log Y_{it} = \kappa_{it} + z_{it} + \nu_{it}$$ (1)

I further assume that the persistent component of labor income, $z$, follows an AR(1) process with a persistence parameter $\rho$ that is close to one:

$$z_{it} = \rho z_{i,t-1} + \eta_{it}$$ (2)

Here, $\eta_{it}$ denotes long-lasting shocks to labor income. The shock $\eta_{it}$ is fully permanent if $\rho$ is one; it is slowly mean-reverting if $\rho$ is less than but still close to one. The transitory component $\nu$ is an MA(1) process. This form of the income process provides a reasonably flexible but parsimonious way to capture income shocks with different persistences. Let $c_{it}$ denote log consumption of household $i$.

The transmission coefficient of long-lasting labor income shocks to consumption (hereafter the transmission coefficient), $\phi$, is defined as the covariance between changes in log consumption and long-lasting shocks to labor income, divided by the variance of such shocks:

$$\phi \equiv \frac{\text{cov}(\Delta c_{it}, \eta_{it})}{\text{var}(\eta_{it})}$$ (3)

That is, the transmission coefficient is the elasticity of consumption with respect to the persistent component of labor income. The transmission coefficient measures the degree to which consumption responds

---

5 Along the wealth dimension, the paper with findings most similar to mine is Casado (2011). He finds that households in wealthier regions of Spain have significantly lower consumption responses to long-lasting shocks to labor income than those in poorer regions. See also Carroll et al. (2014) for a review of wealth heterogeneity in the consumption response to transitory income shocks.
to long-lasting shocks to labor income. An alternative measure would be the marginal propensity to con-
sume (MPC) in terms of levels. However, given that I use self-reported consumption data that do not
cover all consumption categories, estimating an MPC in levels is likely to produce downward-biased re-
results, whereas \( \phi \), an elasticity-based measure, should be less susceptible to this problem. Thus, I use \( \phi \) as
the preferred measure of consumption responses.

3 Data and empirical method

Perhaps surprisingly, the literature contains no conclusive empirical evidence of age effects on \( \phi \) and very
limited empirical evidence of wealth effects on \( \phi \). When BPP estimate a specification where \( \phi \) is allowed to
vary linearly with age, they find a negative coefficient, but it is not statistically significant (they also report
a smaller \( \phi \) for the cohorts born in the 1930s compared with those born in the 1940s, but the difference is
again statistically insignificant). When BPP allow consumption responses to vary with wealth, they report
a higher \( \phi \) for the bottom quintile of the wealth-to-income ratio distribution compared with the other
four quintiles but, once again, the difference is statistically insignificant.

This inconclusiveness can be attributed to two key factors. First, given their focus on the increase in income
and consumption inequality over the ‘80s, BPP restrict their sample to a period between 1980 and 1992,
which limits their sample size. Second, while BPP’s methodology, designed to infer the year-conditional
variances of permanent/transitory income shocks jointly with the transmission coefficients, is well-suited
for their investigation of the composition-shift of income shocks, it also increases the degrees of freedom
in the estimation and thus lowers efficiency in estimating the transmission coefficient.

I tailor BPP’s methodology to the estimation of age and wealth cross-sections of \( \phi \) by making two changes
in the estimation. First, I extend the income-consumption panel to the period between 1968 and 1996 to
increase the sample size. Following Guvenen and Smith (2013), I exploit information in early waves of
the CEX—1972 and 1973—to impute consumption in the PSID before 1980. Second, I adopt an alternative
estimator, a refined version of the instrumental variable estimator described in KV2010, which improves
efficiency in estimating the transmission coefficient by treating the year-conditional variances of income
shocks as nuisance parameters.

I also introduce a third methodological difference by applying parameter stability tests with unknown
break points, in the spirit of Andrews (1993) and Andrews and Ploberger (1994). This type of parameter
stability test searches over various break points in age and wealth, but also takes into account uncertainty
in the estimates resulting from the arbitrary nature of the break points. Such tests do not necessarily make
it easier to reject homogeneity in \( \phi \), compared with using arbitrary cutoffs in age and/or wealth, but they
allow me to test for heterogeneity in consumption responses more thoroughly. These tests are explained
in detail in Section 4.

3.1 Description of data: extending BPP’s sample

Estimating the transmission coefficient \( \phi \) in the age and wealth cross-sections requires a panel with mea-
sures of income, consumption, age and wealth. I use data from the Panel Study of Income Dynamics
The PSID is a panel of US households with detailed and continuously available information on demographics and income. I focus on the years during which the PSID is carried out annually, which results in data that cover age and net labor income for the period 1968–96. I use the age of the head of the household as the measure for age. The PSID reports total family income and its components: wages and salaries, transfer income, and asset income. I compute net family income by adding wages and salaries to transfer income, and subtracting imputed federal income tax on non-asset income. I use the measure of asset income in the PSID as a proxy for wealth. A detailed measure of asset income is regularly available since 1974, which includes interest, dividends, rent on non-owner-occupied housing, and the asset portion of income from unincorporated businesses/farms, but excludes service flows from owner-occupied housing. It is reasonable to assume that households with higher asset income also have a higher level of wealth. Accordingly, I use the ratio of asset income to labor income as a proxy for the liquid-wealth-to-income ratio. I recognize that this could be a noisy measure, but it is the best measure regularly available in the PSID during the sample period.

Consumption measures in the PSID are limited to food expenditures (available annually for 1968–96 except in 1972, 1987 and 1988) and occasionally some other items such as utilities. The CEX, on the other hand, has detailed information on household expenditures, but lacks the panel dimension of the PSID. I create imputed panel measures of nondurable consumption by combining the information in both the PSID and the CEX. Following BPP, I estimate a demand system in the CEX, in which log food expenditures (food at home and food in restaurants) is a linear function of log nondurable expenditures, demographics and calendar time. I then invert the estimated demand system for food to map food expenditures, demographics and calendar time in the PSID into imputed nondurable expenditure. BPP use information in only the continuous CEX, starting in 1980. I follow Guvenen and Smith (2013) and include information in the early wave of the CEX, 1972 and 1973, to impute nondurable expenditure in the PSID prior to 1980. I also incorporate recent suggestions in Campos and Reggio (2014) that improve on the original choice of instrumental variables in BPP to avoid bias in the estimated demand system. Table 6 reports detailed demand system estimates used in the imputation of nondurable expenditure in the PSID sample.

My sample selection procedure is standard. I restrict the focus to households in the PSID Core sample, headed by a male, aged 25 to 65, and not reported to be retired. I drop households with missing information on race, education, or state of residence, and I drop income growth outliers, or those who have top-coded income, food expenditure or taxes. Conditional on these criteria, I consider both a sample of continuously married couples and an unrestricted sample of households.

---

6I impute federal income tax on non-asset income using TAXSIM for the years 1977–96, and I assume that couples file jointly. For years prior to 1977, when TAXSIM is not applicable, I use the imputed federal income tax variable provided in the PSID.

7Values of owner-occupied housing are reported separately, but the value of housing equity is not typically available. Moreover, owner-occupied housing is not a very liquid form of wealth.

8Wealth supplements offered in survey years 1984, 1989 and 1994 offer potentially better wealth measures, but since I need contemporaneous values the resulting sample size would be prohibitively small.

9I use the same definition of nondurable expenditure as Attanasio and Weber (1995) and BPP, which includes food, alcohol, tobacco, and expenditures on other nondurable goods, such as services, heating fuel, public and private transport (including gasoline), personal care, and semidurables (defined as clothing and footwear), and excludes expenditures on health and education, which are more durable and more dependent on family composition.

10Defined as those whose income grows more than 500%, falls by more than 80%, or with an income level below $100.
Extending the sample from 1968 to 1996 (relative to 1980–92) nearly doubles the effective sample size available for estimation. This paper’s estimator (described below) requires observing income in the past three, the current and the next two years, and consumption in the current and the past year. For statistical power, I want a large sample size. With BPP’s age restriction of 30 to 65, their sample would yield 5120 valid observations, while the extended sample yields 9887. With my age restriction of 25 to 65, BPP’s sample would yield 6464 valid observations, while the extended sample yields 12273. This increase in sample size is crucial, given that I want to estimate heterogeneity in \( \phi \) in the cross-section.

### 3.2 Estimating the transmission coefficient

I broadly follow the instrumental variable procedure described in KV2010 to estimate the transmission coefficient of long-lasting income shocks to consumption (\( \phi \)). The estimator assumes unit persistence of long-lasting shocks to income, but performs reasonably well when misspecified (for \( \rho \) close to 1), as shown in the Appendix. Other identifying assumptions are: (1) income shocks are serially uncorrelated, (2) log consumption changes are orthogonal to non-contemporaneous shocks to log-income, and (3) transitory income shocks are MA(1).

To identify \( \phi \), define 
\[
g_{it}^g(y_i) \equiv \Delta y_{it} + \Delta y_{i,t+1} + \Delta y_{i,t+2} \text{ and } s_{it}^g(y_i) \equiv \Delta y_{i,t-2} + \Delta y_{i,t-1} + \Delta y_{i,t} + \Delta y_{i,t+1} + \Delta y_{i,t+2}. 
\]

It follows that
\[
\phi \equiv \frac{\text{cov}(\Delta c_{it}, \eta_{it})}{\text{var}(\eta_{it})} = \frac{\text{cov}(\Delta c_{it}, g_{it}^g(y_i))}{\text{cov}(\Delta y_{it}, s_{it}^g(y_i))}. \tag{4}
\]

Identification in the data is achieved by substituting sample covariances in (4). Standard errors are obtained by a bootstrap with 250 repetitions, clustering over households.

Intuitively, the estimator captures the degree of comovement between consumption and income shocks that last through at least three years. To compare with the original IV estimator in KV2010, note that if \( g_{it}^g(y_i) \) is the same as \( s_{it}^y(y_i) \), the estimator is precisely that described in KV2010; the covariance between \( \Delta c_{it} \) and \( \eta_{it} \), however, is better estimated using the sample covariance between \( \Delta c_{it} \) and \( g_{it}^g(y_i) \) rather than using the sample covariance between \( \Delta c_{it} \) and \( s_{it}^y(y_i) \).

The difference between \( g_{it}^g(y_i) \) and \( s_{it}^y(y_i) \) represents a small refinement to the IV estimator in KV2010. The major difference between my estimation (or any other estimation of \( \phi \) akin to the IV estimation in

---

11The years between 1993 and 1996 warrant some comments. Prior to 1993, an annualized measure of food expenditure is available from the PSID, whereas since 1993 food expenditure is recorded in disaggregated time periods, so that no official annualized measure is available, and the user is responsible for time aggregation. I check the quality of my time-aggregation procedure by exploiting the fact that both the time-disaggregated measure and the official annualized measure are available for 1992. I confirm that my annualized measure matches the official annualized measures provided by the PSID almost perfectly. As a side note, BPP do not use data between 1993 and 1996 in their main estimation, but they do consider these years in an unpublished sensitivity analysis that they report to maintain their overall results.

I also note that the variance of transitory changes in reported food expenditure in the PSID is larger during 1993–96 than before 1993, indicating a larger measurement error in consumption, possibly due to the switch to a computer-assisted interview system. As will become clear later, I ensure that the estimate of consumption responses to long-lasting income shocks used in this paper is robust to classical measurement error in consumption.

12I use the same identification assumptions as BPP. The third assumption is motivated by the fact that log income growth is significantly negatively correlated at the first and the second lag in my sample; the same pattern that prompts BPP to assume an MA(1) transitory income component.

13In this particular case, clustered standard errors are preferable and often smaller than non-clustered ones because there is negative intra-cluster correlation in independent variables. Most of the results remain unchanged, however, when using non-clustered standard errors.
KV2010) and BPP’s minimum distance estimation, however, lies in the fact that I treat the year-conditional variances of $\eta_{it}$ as nuisance parameters (which need not be explicitly estimated along with $\phi$), whereas in BPP’s minimum distance estimation the year-conditional variances of $\eta_{it}$ are always explicitly estimated.\footnote{The year-conditional variances of $\eta_{it}$ are allowed to vary under both methods.} In panels with relatively large $T$ and relatively small $N$, BPP’s minimum distance estimation does not benefit as much from the increase in $T$, since degrees of freedom also increase with $T$, whereas estimations analogous to the IV in KV2010 take full advantage of increases in the length of the sample in the estimation of $\phi$.\footnote{I note that the property of BPP’s minimum distance procedure is not well understood under small $N$, which warrants further investigation in its own right. I encounter convergence problems in several cases when I use BPP’s minimum distance procedure to estimate $\phi$ for high/low age groups, using different age thresholds in their 1980–92 sample.}

Given that I focus exclusively on estimating $\phi$, the estimator based on (4) is a natural choice over BPP’s original minimum distance procedure. I report a Monte Carlo exercise in the Appendix which further shows that the estimator used in this paper is robust to measurement errors in income and consumption and performs reasonably well when misspecified (for $\rho$ close to 1).

When implementing the estimator, I use the measures of nondurable expenditure and net labor income described in Section 3.1. I control for the anticipated component in log net labor income by removing age effects and year effects, interacted with level of education. I also control for potential year-specific imputation bias in log nondurable expenditure (log consumption) by removing year effects. The residual log income and log consumption are then used in the estimation of $\phi$.\footnote{I follow the semiparametric approach in Fernandez-Villaverde and Krueger (2007) and simultaneously estimate year effects using year dummies and age effects as a smooth function of age (non-Caucasian observations—around 5% of the sample—have to be dropped in this step due to lack of information in estimating education-specific age effects). Any fixed effects in income or consumption, for example cohort effects, would be controlled for when I take the first difference in income and consumption. I experiment with various controls that might lead to anticipated or spurious changes in income and/or consumption (such as family composition), and I also experiment with different specifications of age effects in income/consumption. The heterogeneity results that I report are scarcely affected by changes along these dimensions.}

4 Empirical Results

Before looking at the age and wealth cross-sections of $\phi$, I first confirm that my sample and estimation method produce a reasonable average estimate of $\phi$ relative to existing results in the literature. When I do not restrict family type in the sample, the average estimate of $\phi$ is 0.47, with a standard error of 0.05. When I look at only continuously married couples, the average estimate of $\phi$ is 0.37, with a standard error of 0.07. While my average estimate of $\phi$ might seem low, it is actually similar to results reported in other studies. For example, Etheridge (2014) finds an average $\phi$ of 0.42 using British data, with a standard error of 0.14. Although BPP’s preferred number for $\phi$ is 0.64, they remove the effects of employment changes from the stochastic components of income and consumption. If the effects of employment changes on income and consumption were not removed, as here and in Etheridge (2014), their estimate of $\phi$ would be 0.43, with a standard error of 0.06.

The transmission coefficient seems to be constant over the sample period: in both cases, I cannot reject a
constant $\phi$ before and after 1980 or in 1985 (the target year of a time-break test in BPP). In what follows, I restrict the transmission coefficient as constant over the whole period.

4.1 Age profile of consumption responses in the PSID

Figure 1 displays rolling estimates of $\phi$ by age. Specifically, the rolling estimate of $\phi$ at each age $i$ is the estimate of $\phi$ in a subsample with ages between $i - 2$ and $i + 2$, including endpoints. Confidence intervals at the 95% level are plotted as well.

Two observations arise from Figure 1. First, there is some evidence that $\phi$ is lower for households with heads older than 51. In the unrestricted sample, rolling estimates of $\phi$ for those aged above 51 are low (below 0.3) and not statistically distinguishable from zero, whereas for those aged below 49 the rolling estimates of $\phi$ are high (between 0.4 and 0.8) and statistically different from zero. The evidence is weaker in the continuously married sample, as rolling estimates of $\phi$ are lower for those above 51 as well as for those in their mid-30s. Second, the standard errors for the rolling estimates are large, ruling out conclusions of an age effect.  

Table 1 provides another way to examine the age pattern in consumption responses to long-lasting income shocks. When the sample is grouped by age bins, there is again some evidence of a decreasing pattern with age, although standard errors are large.  

---

18 Figure A.3 plots rolling estimates in the original BPP sample with continuously married couples aged from 30 to 65. There, the point estimates share similar features with the right-hand panel of Figure 1. Standard errors are even larger, as expected with a smaller sample size.
Table 1: Transmission coefficient for age bins

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Unrestricted $\phi$</th>
<th>s.e.</th>
<th>Married $\phi$</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>28 ≤ age ≤ 30</td>
<td>0.41</td>
<td>0.14</td>
<td>0.65</td>
<td>0.21</td>
</tr>
<tr>
<td>30 ≤ age &lt; 35</td>
<td>0.50</td>
<td>0.13</td>
<td>0.38</td>
<td>0.15</td>
</tr>
<tr>
<td>35 ≤ age &lt; 40</td>
<td>0.67</td>
<td>0.15</td>
<td>0.32</td>
<td>0.20</td>
</tr>
<tr>
<td>40 ≤ age &lt; 45</td>
<td>0.63</td>
<td>0.14</td>
<td>0.66</td>
<td>0.19</td>
</tr>
<tr>
<td>45 ≤ age &lt; 50</td>
<td>0.76</td>
<td>0.26</td>
<td>0.51</td>
<td>0.31</td>
</tr>
<tr>
<td>50 ≤ age &lt; 55</td>
<td>0.04</td>
<td>0.26</td>
<td>0.14</td>
<td>0.21</td>
</tr>
<tr>
<td>55 ≤ age &lt; 60</td>
<td>0.21</td>
<td>0.12</td>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>60 ≤ age ≤ 62</td>
<td>0.15</td>
<td>0.19</td>
<td>0.20</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Given that large standard errors preclude statistical conclusions when estimating $\phi$ for narrower age groups, a logical next step is to combine age groups to reduce standard errors. The data allow me to follow Chow (1960), Andrews (1993), and Andrews and Ploberger (1994) and perform parameter stability tests of $\phi$ over age, assuming either known break points in the case of Chow (1960) or unknown break points in the case of Andrews (1993) and Andrews and Ploberger (1994). Although the parameter break tests in Andrews (1993) and Andrews and Ploberger (1994) are designed for a one-time change in the parameter, they are shown to have power against more general forms of parameter instability and can therefore be applied to cross-sectional heterogeneity in $\phi$. I start with the simple case of known break points; i.e., the Chow test. For different ages, I estimate $\phi$ separately for the sample with household heads strictly older than the threshold and the sample with household heads weakly younger than the threshold, and test for equality in $\phi$.

Figure 2: Chow tests of consumption responses (at various age thresholds)

Figure 2 displays the results from this exercise. The horizontal axis displays different age thresholds.
potential breaks”). Each point marked by a solid circle is the estimated transmission coefficient for households with heads whose ages fall below the corresponding age threshold ($\phi_L$). Each point marked by a cross is the estimated transmission coefficient for households with heads whose ages are greater than or equal to the corresponding age threshold ($\phi_G$). The bands around each pair of circles and crosses illustrate a Chow test; i.e., a two-sided test of age homogeneity in consumption responses with a known break point ($H_0 : \phi_L = \phi_G$), with width of the bands set to $k \sqrt{(\hat{se}_G)^2 + (\hat{se}_L)^2}$, where $1 - \Phi(k) = 0.025$ and $\Phi(\cdot)$ denotes the cdf of a standard normal distribution.

Figure 2 is designed so that the null hypothesis, $H_0$, of the Chow test is rejected at a certain age threshold if, and only if, the two bands do not overlap. Notice that for age thresholds 45 to 58, the null hypothesis is rejected (in both the unrestricted sample and the continuously married sample): the consumption response of households with heads aged 45 or older (or 46, ..., or 58) is significantly smaller than that of households with heads younger than 45 (or 46, ..., or 58).

Differences in the point estimates between the younger groups and the older groups are large, further highlighting the economic significance. Taking the midpoint age threshold of 45 as an example, the average $\phi$ of the younger group is 0.59 in the unrestricted sample (0.52 in the continuously married sample), while the average $\phi$ of the older group is 0.25 (0.16 in the continuously married sample), a difference of 0.34 (0.36 in the continuously married sample).

To ensure that the age-heterogeneity results are robust to alternative choices of the age break, I apply a class of parameter stability tests with unknown break points, described in Andrews (1993) and Andrews and Ploberger (1994). In essence, if the Chow test for each known age break is:

$$F(age) = \frac{(\hat{\phi}_L - \hat{\phi}_G)^2}{(\hat{se}_G)^2 + (\hat{se}_L)^2}$$

then under the null hypothesis that $H_0 : \phi_L(age) = \phi_G(age)$, $\forall$age $\in [a, \bar{a}]$, I can find the distribution for the following tests and reject the null hypothesis of age homogeneity in $\phi$ if the sample value of SupF (ExpF, AveF) is higher than a critical value:

$$\text{SupF} = \sup_{a_1 \leq \text{age} \leq a_2} F(age)$$

$$\text{ExpF} = \ln \left( \frac{1}{a_2 - a_1 + 1} \sum_{\text{age} = a_1}^{a_2} \exp \left( \frac{1}{2} F(age) \right) \right)$$

$$\text{AveF} = \frac{1}{a_2 - a_1 + 1} \sum_{\text{age} = a_1}^{a_2} F(age)$$

The tests take into account the uncertainty in the estimated transmission coefficients for the younger group and the older group due to arbitrariness in the age thresholds. Applying these tests to the sample, I confirm that the earlier conclusion holds: In both the unrestricted sample and the continuously married sample, the ExpF and AveF tests reject age homogeneity at the 5% significance level. The SupF test rejects...
age homogeneity at the 5% significance level in the unrestricted sample, and at the 10% significance level in the continuously married sample.\footnote{Following Andrews (1993), I set $a_1$ and $a_2$ to exclude the highest and lowest 15% of the age thresholds when calculating the test statistics. I use Monte Carlo methods to compute the critical value of the tests: I simulate the value of SupF, ExpF, AveF under the null hypothesis (with $\phi$ set at the average estimate), the income process (1) and a log-linear consumption rule many times, and find approximate distributions of the test statistics. The approximate test statistic distribution that I obtain turns out to be insensitive to the parameterizations used in the simulations.}

In summary, when applying parameter stability tests, I find that there is both statistically significant and economically significant age heterogeneity in consumption responses to long-lasting income shocks (or equivalently, age difference in $\phi$). This conclusion holds independently of whether I assume that the econometrician knows the exact age points at which $\phi$ changes.

I also consider a parametric approach of the sort used in BPP. When I allow for a linear effect in age in the estimation of $\phi$, I obtain a significant negative coefficient, as reported in Table 2. The parametric estimates imply a drop of .31 in $\phi$ from .61 at age 30 to .31 at age 55 in the unrestricted sample, and a corresponding drop of .37 in $\phi$ (from .57 at age 30 to .20 at age 55) in the continuously married sample.

| Table 2: Age heterogeneity in consumption responses: linear specification $\phi(age_{it}) = \phi_0 + \phi_1 (age_{it} - \text{ave}(age_{it}))$ |
|-------------------------------|-----------------|-----------------|-----------------|-----------------|
| Unrestricted                  | Continuously-Married |
| point estimate | standard error | point estimate | standard error |
| $\phi_0$      | 0.487*** | 0.068 | 0.406*** | 0.079 |
| $\phi_1$      | -0.012** | 0.006 | -0.015** | 0.006 |

4.2 Wealth heterogeneity in consumption responses

Next, I consider the effect of wealth on $\phi$. I first report empirical estimates of $\phi$ for different wealth groups. I sort households by their position in the year-specific, asset-income-to-labor-income ratio (AY/LY) distribution. I then ask whether a household’s wealth (as measured by AY/LY) in the past year influences its ability to insure against long-lasting income shocks in the current year. Table 3 shows that quintiles higher up in the AY/LY distribution do have lower point estimates of $\phi$. The point estimates suggest that households in the highest quintile of the AY/LY distribution are more than twice as capable of smoothing consumption over long-lasting labor income shocks than households in the lowest two quintiles, who have little to no income from liquid assets.
Table 3: Transmission coefficient for quintiles in the AY/LY ratio distribution

<table>
<thead>
<tr>
<th>Group</th>
<th>Unrestricted</th>
<th>Married</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>φ</td>
<td>median(AY/LY)</td>
</tr>
<tr>
<td>bottom quintile</td>
<td>0.71</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(.11)</td>
<td>(.12)</td>
</tr>
<tr>
<td>2nd quintile</td>
<td>0.71</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(.10)</td>
<td>(.11)</td>
</tr>
<tr>
<td>3rd quintile</td>
<td>0.70</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(.12)</td>
<td>(.14)</td>
</tr>
<tr>
<td>4th quintile</td>
<td>0.43</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>(.13)</td>
<td>(.15)</td>
</tr>
<tr>
<td>top quintile</td>
<td>0.31</td>
<td>0.250</td>
</tr>
<tr>
<td></td>
<td>(.11)</td>
<td>(.12)</td>
</tr>
</tbody>
</table>

Figure 3 summarizes the results from a parameter stability test in φ over wealth, conducted in the same way as that in Section 4.2, with deciles in the year-specific AY/LY ratio distribution as potential cut-offs. The parameter stability tests (assuming either known or unknown break points) all reject the null hypothesis of wealth homogeneity in φ at the 5% significance level. The test and the point estimates suggest that higher wealth does reduce the consumption response to long-lasting income shocks, and that the effect of wealth on φ is both statistically and economically significant.

Figure 3: Chow tests of consumption responses (at various wealth thresholds)

The results for wealth are not without issues. First, unlike age, which is nearly perfectly measured,
the wealth/labor-income ratio in the current results is proxied by the asset-income/labor-income ratio.\textsuperscript{21} Second, asset income in the PSID includes income from business, farm and rents. I examine what portion of the wealth results is driven by business-owners, farm-owners, or second-home (for rental) owners, who might be systematically different from workers. When I remove these groups, which make up less than 10% of the sample, and consider only interest/dividend/trust income relative to labor income, it is reassuring that I still find differences in the point estimates for $\phi$ across quintiles of the AY/LY ratio distribution, but the differences are now smaller and not statistically significant.

Notwithstanding these drawbacks, the results do confirm a wealth effect on consumption insurance against long-lasting labor-income shocks. The cautionary notes call for further work to refine the estimation of this wealth effect.

4.3 Discussion

Because age is correlated with wealth in the data, it is of interest to ask whether there is an age effect independent of wealth, and vice versa. Intuitively, age exerts an independent force, since shocks late in life are effectively less permanent. I try to separately identify the conditional effects of age and wealth with the following two exercises.

For the conditional effect of age on $\phi$, I re-estimate the age profile of consumption responses, excluding high-wealth households. When I exclude households in the highest quintile of the AY/LY distribution, the age-heterogeneity results are almost unchanged: while the standard errors become somewhat larger, parameter stability tests reject age homogeneity in $\phi$ at the 5% significance level. This suggests that the conditional effect of age does contribute in generating the empirical decreasing age profile of $\phi$. As a stricter test, I also experiment with excluding households in both the highest and the second-highest quintiles in the AY/LY distribution when re-estimating the age profile of consumption responses. Not surprisingly, a further increase in the standard errors precludes statistical conclusions in this stricter test, but the point estimates still display patterns similar to those found in Section 4.1.

For the conditional effect of wealth on $\phi$, I re-estimate the effect of wealth on consumption responses excluding households with heads older than 55 (or, as a stricter test, excluding households with heads older than 50). Table 4 displays the results. In both the unrestricted and the continuously married samples, the results suggest similar heterogeneity across wealth quintiles when households with older heads are excluded. Parameter stability tests reject wealth homogeneity in $\phi$ at the 10% significance level in all cases, and reject wealth homogeneity in $\phi$ at the 5% significance level if I exclude only households with heads older than 55. This provides evidence that the conditional effect of wealth is also at play in generating heterogeneity in $\phi$.

\textsuperscript{21}I examined the quality of this proxy using the direct wealth measures in the PSID (available in survey years 1984, 1989, 1994 and biennially after survey year 1999) and found little correlation. It is unclear whether this is due to the quality of the PSID direct wealth measures (which are imputed using answers to bracketed questions), or due to the nature of the proxy.
Table 4: Conditional effect of wealth

<table>
<thead>
<tr>
<th>AY/LY Group</th>
<th>Wealth Quintile</th>
<th>Unrestricted</th>
<th>Married</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>≤ 55</td>
<td>≤ 50</td>
</tr>
<tr>
<td></td>
<td>φ</td>
<td>φ</td>
<td>φ</td>
</tr>
<tr>
<td>bottom quintile</td>
<td>0.71(0.11)</td>
<td>0.71(0.11)</td>
<td>0.74(0.10)</td>
</tr>
<tr>
<td>2nd quintile</td>
<td>0.71(0.10)</td>
<td>0.71(0.10)</td>
<td>0.74(0.11)</td>
</tr>
<tr>
<td>3rd quintile</td>
<td>0.70(0.12)</td>
<td>0.73(0.13)</td>
<td>0.79(0.13)</td>
</tr>
<tr>
<td>4th quintile</td>
<td>0.43(0.13)</td>
<td>0.51(0.15)</td>
<td>0.62(0.15)</td>
</tr>
<tr>
<td>top quintile</td>
<td>0.31(0.11)</td>
<td>0.40(0.13)</td>
<td>0.45(0.15)</td>
</tr>
</tbody>
</table>

5 Accounting for heterogeneity in φ

As noted in the introduction, life-cycle versions of incomplete markets models predict that consumption responses to persistent income shocks decrease with both age and wealth. Having presented empirical evidence in support of these qualitative predictions, in this section I examine whether these estimated responses are quantitatively consistent with a reasonably calibrated version of this model.

5.1 The model and the calibration

The model and its calibration follow KV2010, except that: 1) I use the empirically estimated persistence of the income process; and 2) I estimate relative risk-aversion, the discount factor, and looseness of the borrowing limit jointly, using wealth moments.

The economy consists of a continuum of households with time-separable CRRA preferences over consumption; there are no bequest motives. There is no aggregate uncertainty. Until retiring at $T_{ret}$, each household receives net labor income $Y_{it}$ governed by the process characterized in equations (1) and (2). After $T_{ret}$, households receive social security transfers as a function of the average of gross pre-retirement labor income $\tilde{Y}_i = \frac{1}{T_{ret}} \sum_{t=1}^{T_{ret}} Y_{it}$. At each age $t > T_{ret}$, households survive with probability $\xi_t$. At a final age $T$, households terminate with certainty. Households trade one-period bonds at a fixed interest rate $R$ (taken as given), with a borrowing limit $A_{it} \geq \tilde{A}_{it}$, defined as:

$$\tilde{A}_{it+1} \equiv \sum_{\tau=1}^{T_{ret}} \left( \frac{\psi}{R} \right)^{\tau} \min(Y_{it}) + \left( \frac{\psi}{R} \right)^{T_{ret} - t} \sum_{\tau=T_{ret}+1}^{T} \left( \frac{1}{R} \right)^{\tau} \min(\tilde{Y}_i)$$

This form follows Guvenen and Smith (2013), where $\psi \in [0, 1]$ measures the looseness of the borrowing limit. Lowest future income realizations are discounted at rate $R/\psi$ until $T_{ret}$ and at rate $R$ afterwards. The natural borrowing limit corresponds to $\psi = 1$. The zero borrowing limit effectively corresponds to $\psi = 0$. Intermediate cases of the borrowing limit are captured by $\psi \in (0, 1)$. Initial conditions $A_{i0}$, $z_{i0}$ are drawn from a joint distribution $H(A_{i0}, z_{i0})$. Households have access to a perfect annuity market, which

---

\(^{22}\) A tax function $\tau_{it} = \tau(\tilde{Y}_{it})$ implies a one-to-one mapping between net labor income and gross labor income at each age ($Y_{it} = \tilde{Y}_{it} - \tau(\tilde{Y}_{it})$).
allows me to formulate household $i$’s problem:

$$
V(A_{i0}, z_{i0}, \epsilon_{i0}) = \max_{(C_{it}, A_{it})} \sum_{t=0}^{T} \beta^t \frac{C_{it}^{1-\gamma}}{1-\gamma}
$$

subject to:

$$
C_{it} + A_{it+1} \leq RA_{it} + Y_{it}, \quad A_{it} \geq A_{it+1} \quad \text{for } t \leq T^{ret}
$$

and

$$
C_{it} + A_{it+1} \leq RA_{it} + P(\tilde{Y}_t), \quad A_{it+1} \geq A_{it+1} \quad \text{for } t > T^{ret}
$$

where $Y_{it}$ follows (1), (2).

The model is calibrated as follows. Each household enters the model at age 25 and retires at age 65.23 Terminal age is 90. The interest rate is taken to be exogenous and set to 3%. The mapping between gross and net labor income follows Gouveia and Strauss (1994). The social security transfer function $P(\cdot)$ mimics the actual system in 1990; benefits are a piecewise linear function of average gross labor income during working life. The initial conditions for income and wealth are chosen to match data from the PSID and the Survey of Consumer Finances (SCF). The income profile is estimated from the PSID. Variance of the long-lasting income shocks is taken to be age-invariant and chosen to match the empirical life-cycle increase in income mean and dispersion. Following the method in Guvenen (2009), I estimate the persistence parameter $\rho$ for the income process to be 0.964 in my sample. I choose the rest of the parameters to match key features of the wealth distribution. I first estimate the distribution of the working-life wealth-to-income ratio in the SCF from 1983 to 1995, excluding the top 5% in wealth, given that the PSID undersamples the top of the wealth distribution.24 I then take three parameters $\beta$, $\gamma$ and $\psi$ as free parameters, and estimate the model using six moments from the SCF: the fraction of households with negative net worth, and the 10th, 25th, 50th, 75th and 90th percentile of the working-life wealth-to-income distribution. I estimate a discount factor $\beta$ of 0.93, a relative risk aversion $\gamma$ of 3.03, and a looseness parameter of the borrowing limit $\psi$ of 0.26.

Table 5 lists both the data and the model moments for the wealth-to-income ratio distribution. The results suggest that the model fits the target distribution of the working-life, wealth-to-income ratio well.

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Fraction with negative NW</th>
</tr>
</thead>
<tbody>
<tr>
<td>10th</td>
<td>25th</td>
</tr>
<tr>
<td>Data</td>
<td>0.09</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Model</td>
<td>0.12</td>
</tr>
</tbody>
</table>

23 This is a departure from KV2010, which uses 60 as the retirement age. I use 65 to be consistent with the estimates in Section 4.

24 I measure wealth in the SCF as net worth and income as gross labor income. To conform with sample selection in the PSID, I exclude households in the SCF with heads younger than 25 or older than 65, or households with female heads. To ensure that the wealth-to-income ratio is not biased upward due to low income observations, I look at households with at least one member working full-time and making an annual labor income no lower than the minimum hourly wage times 2000 annual hours.
5.2 Model results

Figure 4 plots the age profile of transmission coefficients in the model against the age profile of transmission coefficients in the data. I simulate an artificial panel of 50000 households to calculate the model’s transmission coefficients.

![Figure 4: Age profile of consumption responses in the data and in the model](image)

The calibrated model predicts an age profile of consumption responses that is quantitatively similar to that found in the data. Consumption responses to long-lasting shocks to labor income decrease with age in the calibrated model. The left-hand panel of Figure 4 shows that the model-generated age profile of \( \phi \) is within the 95% confidence interval at most age points for the unrestricted sample. In terms of the point estimates, the model accounts for around 75% of the drop in \( \phi \) from age 40 to retirement. The model overpredicts \( \phi \) for those younger than 32, which is due mainly to the fact that \( \psi \) is estimated to be low, leading to a tight borrowing limit. The literature offers two solutions to this problem. First, the young might have access to insurance within the extended family through non-monetary channels. For example, Kaplan (2012b) shows that the option of moving back with parents is quantitatively relevant in providing insurance for young workers. Second, the persistence of income shocks might be lower for young workers, as shown in Karahan and Ozkan (2013).

The right-hand panel of Figure 4 compares model predictions with estimates from the continuously married sample. The model matches the age decline in \( \phi \) in the continuously married sample, but overpredicts the average consumption response.\(^{25}\) The fit for ages around 35 is somewhat worse for this sample.

To ensure conformity in comparison with the empirical results, the model’s transmission coefficients above and in what follows are calculated using the same IV estimator as in Sections 3 and 4. A close alternative would be to calculate the model’s transmission coefficients directly from the realizations of the individual

\(^{25}\)The model average \( \phi \) is 0.57. The empirical average \( \phi \), as reported in Section 4, is 0.47 (standard error 0.05) in the unrestricted sample, and 0.37 (standard error 0.07) in the continuously married sample.
shocks. This is feasible in the model and produces the true value of $\phi$, but is not feasible for the actual data. In the model, the value of $\phi$ calculated by the IV estimator is close to the true value but can contain an upward bias, the size of which depends on the fraction of constrained households, as pointed out by KV2010. However, in Figure A.2 of the Appendix, I show that this upward bias is quantitatively small when the model’s age profile of $\phi$ is calculated. This is because the fraction of constrained households is small at each age in the calibrated model.

From Table 6 it can be seen that the model also produces wealth effects on consumption responses that are similar to those found in the data. Given that the ratio of asset income to labor income in the model is just a scaled version of the wealth-to-labor-income ratio (the interest rate is age invariant), I sort observations in the simulated data and in the PSID data by the ratio of asset income to labor income. The values of $\phi$ in the model decrease with wealth within the confidence intervals of the empirical estimates in all but the bottom wealth quintile. Moreover, the bias of the IV in the model is also small in those quintiles. The bottom wealth quintile has a high model transmission coefficient. But that is partly caused by the upward bias of the IV estimator associated with a large number of constrained households; the true model transmission coefficient in the bottom wealth quintile is 0.78. Therefore, one potential resolution of the discrepancy in $\phi$ between the model and the data in the bottom wealth quintile, is that households in the bottom wealth quintile of the data are, in fact, not as constrained as those in the calibrated model, such that the the IV bias is present in the model but not in the data. In that case, one should compare empirical values in the bottom quintile with the true model transmission coefficient, which would suggest a reasonable fit.

<table>
<thead>
<tr>
<th>AY/LY</th>
<th>Unrestricted</th>
<th>Married</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>$\hat{\phi}$</td>
<td>$\hat{\phi}$</td>
<td>$\hat{\phi}$ (IV)</td>
</tr>
<tr>
<td>bottom quintile</td>
<td>0.71(0.11)</td>
<td>0.60(0.12)</td>
<td>0.86</td>
</tr>
<tr>
<td>2nd quintile</td>
<td>0.71(0.10)</td>
<td>0.61(0.11)</td>
<td>0.65</td>
</tr>
<tr>
<td>3rd quintile</td>
<td>0.70(0.12)</td>
<td>0.42(0.14)</td>
<td>0.57</td>
</tr>
<tr>
<td>4th quintile</td>
<td>0.43(0.13)</td>
<td>0.37(0.15)</td>
<td>0.45</td>
</tr>
<tr>
<td>top quintile</td>
<td>0.31(0.11)</td>
<td>0.14(0.12)</td>
<td>0.34</td>
</tr>
</tbody>
</table>

In the model, two forces drive the age and wealth heterogeneity in consumption responses. First, there is a horizon effect. As agents get closer to the end of their working lives, any given persistent income shock has less time to play out, reducing its impact on consumption. Corresponding to the earlier discussion in Section 4.3, this is the conditional effect of age in the model. Second, there is a buffer-stock effect. As households accumulate more financial wealth relative to their labor income, they can better insure against income shocks. This is the conditional effect of wealth in the model.

Figure 5 displays the conditional effects of age and wealth on $\phi$ in the model. Household-age observations in the simulated data are first sorted by the position in quintiles of the overall working-life distribution of the wealth-to-income ratio. For each quintile of the wealth-to-income ratio distribution, I then plot $\phi$ at

---

26 The origin of the bias is as follows: In the model, borrowing-constrained households who receive transitory negative income shocks will expect consumption to increase in the near future. This creates a negative correlation between consumption growth and lagged transitory income shocks, which violates one of the orthogonality conditions of the IV estimator.

27 The idea that some households in the data might be less constrained than those in the pure self-insurance model is consistent with the findings in Heathcote et al. (2009) and KV2010.
each age where there are more than 50 observations. To reveal what really occurs in the model, the true values of $\phi$ are plotted. Thus, the results illustrated in Figure 5 are not blurred by the statistical nuance of the IV.

Evidently, both conditional effects are important in the model. The slopes of the lines represent the conditional effect of age; i.e., the horizon effect. All five lines slope downward, indicating an age-related decline in $\phi$ near retirement that does not come from the effect of wealth. The horizon effect is particularly strong after age 55 in the model. At the very end of working life, $\phi$ for all quintiles drops towards zero. The gaps between the lines for different wealth quintiles represent the conditional effect of wealth; i.e., the buffer-stock effect. The monotonicity of $\phi$ in wealth is clear, except at the very end of working life. The buffer-stock effect is quantitatively large: the transmission coefficient for households in the top wealth quintile is 30% to 50% smaller than the coefficient for households in the bottom quintile, conditional on age before age 55.

Figure 5: Joint age-wealth distribution of $\phi$ in the model

6 Concluding remarks

From the data I find that there is substantial heterogeneity across households in the response of non-durable consumption to long-lasting income shocks. In the face of labor-income shocks that last through at least three years, households with older heads adjust consumption by much less. Households with more liquid wealth also adjust consumption by much less. A calibrated version of the life-cycle incomplete markets model produces consumption responses that are quantitatively similar with regard to age and wealth heterogeneity to those estimated from the data. This is achieved jointly by the buffer-stock
effect, inherent in the incomplete markets assumption, and the horizon effect, generated from the life-cycle assumption.

What do these results mean for future research? First, it could be useful to verify these findings using panel data from countries other than the United States: Italy, Spain and the United Kingdom all provide data suitable for a similar exercise.

Second, the results in this paper imply significant cross-sectional heterogeneity among US households in the ability to self-insure. To the extent that the optimal design of public insurance depends on self-insurance capabilities, these results will be relevant for optimal policy in the cross-section. For example, policies concerned with long-term income maintenance for groups with a high ability to self-insure are not likely to provide significant welfare gain. In this regard, the verification of the life-cycle incomplete markets model that I offer here implies that the model is suitable for analyzing or designing age-dependent or means-tested public insurance policies and taxes.
References


Appendix

A1. Performance of estimator for $\phi$ under measurement error, long-run mean-reversion, and history dependence

The estimator for $\phi$ obtained in Section 3.2 is robust to classical measurement errors in consumption and income. The estimator will be misspecified when $\rho < 1$; i.e. when there is mean-reversion in the long-lasting random component of log income. Nevertheless, a simple Monte-Carlo analysis of the estimator shows that when $\rho$ is close to one, the estimator performs reasonably well despite mean-reversion in the income process. The estimator will also be misspecified when consumption growth correlates with past income shocks (history dependence), as discussed in KV2010. However, the resulting bias is quantitatively small in the age profile of $\phi$ for the calibrated model in this paper.

The reason why the estimator for $\phi$ is robust to classical measurement errors is clear: since classical measurement error in log income is akin to transitory income shocks, it is not correlated with $g_t^y(y_t)$, which picks up only long-lasting changes in observed log income. On the other hand, classical measurement error in the dependent variable (log consumption) does not harm estimation, for the usual reason. By the same logic, the estimator for $\phi$ is also robust to measurement errors that are MA(1). \(^{28, 29}\)

When there is mean-reversion in the long-lasting income shock ($\rho < 1$), the estimator for $\phi$ will be misspecified:

$$\frac{\text{cov}(\Delta c_{it}, g_t^y(y_t))}{\text{cov}(\Delta y_{it}, g_t^y(y_t))} = \frac{\text{cov}(\Delta c_{it}, \rho^2 \eta_{it} + \rho \eta_{it+1} + \eta_{it+2} + (\rho^3 - 1)z_{it-1} + \varepsilon_{it+2} - \varepsilon_{it-1})}{\text{cov}(\Delta y_{it}, \rho^2 \eta_{it} + \eta_{it+2} + \rho \eta_{it+1} + \rho^3 \eta_{it-1} + \rho^4 \varepsilon_{it-2} + (\rho^5 - 1)z_{it-3} + \varepsilon_{it+2} - \varepsilon_{it-3})}$$

$$= \frac{\text{cov}((\rho - 1)(\eta_{it-1} + \rho \eta_{it-2} + \rho^2 z_{it-3}) + \eta_{it} + \Delta \varepsilon_{it}, \eta_{it+2} + \rho \eta_{it+1} + \rho^2 \eta_{it} + \rho^3 \eta_{it-1} + \rho^4 \eta_{it-2} + (\rho^5 - 1)z_{it-3} + \varepsilon_{it+2} - \varepsilon_{it-3})}{\text{var}(\eta_{it}) + (1 - \rho)[(1 - \rho^3)\text{var}(z_{it-3}) - \rho \text{var}(\eta_{it-1}) - \rho^4 \text{var}(\eta_{it-2})]}$$

$$= \frac{\text{cov}(\Delta c_{it}, \eta_{it})}{\text{var}(\eta_{it}) + (1 - \rho)[\text{var}(z_{it-3}) - \rho \text{var}(z_{it-1})]}$$ (10)

This type of misspecification bias will depend on the age of agents. When the variation in the initial condition for $z$ is not too small and $\rho$ is not too close to 1, $\text{var}(z_{it-3})$ will be larger than $\rho \text{var}(z_{it-1})$, and there will be a downward bias in the estimation of $\phi^m$. If, however, $\text{var}(z_{it-3}) < \rho \text{var}(z_{it-1})$, for example when initial variation in $z$ is small and variance of the $z$ component is quickly accumulating from $t - 3$ to $t - 1$, there will be an upward bias in the estimation of $\phi^m$. In both cases, the bias will be small if $\rho$ is reasonably close to 1.

\(^{28}\)If measurement errors are systematic, then the estimator here might not perform well. However, for such systematic measurement errors to affect the heterogeneity result in this paper, it must be the case that certain groups (old, high liquid wealth) over-report income changes and/or under-report consumption changes, while other groups (young, low liquid wealth) under-report income changes and/or over-report consumption changes.

\(^{29}\)Measurement errors, however, will bias down estimates of consumption response to transitory income shocks, since typical measurement errors in income will be observationally equivalent to transitory income shocks. This is why I do not focus on transitory income shocks.
A simple Monte Carlo experiment illustrates the pattern of bias in $\hat{\phi}$ when $\rho < 1$. The income process in (1) is parameterized with $T = 40$, $var(z_{i0}) = 0.15$, $var(\epsilon_{it}) = 0.05$, $\theta = 0.10$. To isolate the effect of income mean-reversion, I assume a log-linear consumption policy $\Delta c_{it} = \phi \eta_{it}$ with $\phi = 0.5$, so that consumption orthogonality conditions are satisfied. The income process is simulated for 50000 households in each experiment. I repeat the experiment 200 times for each of the following values of $\rho$: 1, 0.99, 0.97, 0.93, and set $var(\eta_{it})$ in each case such that $var(z_{iT}) = 0.55$. Figure 1 plots the average estimate of $\phi$ at each time $t$. The lines from top to bottom correspond to different true value of $\rho$. The absolute value of the bias in $\hat{\phi}$ when $\rho < 1$ is smaller than 0.05 in all cases. Age specificity of the bias is concentrated before age 35, and is also small in magnitude.

The estimator will also be misspecified if consumption growth is correlated with past income shocks (history dependence of consumption). KV2010 point out that in simulated data generated by the life-cycle incomplete markets model, the IV estimator overestimates consumption responses for households near the borrowing limit. In the model, constrained households who receive negative transitory shocks are forced to consume much of the shock, but they will expect consumption growth in the future as the shock wears off. This creates a negative correlation between consumption growth and lagged transitory income shocks, which violates the orthogonality conditions of the IV estimator. With potentially non-zero $\text{cov}(\Delta c_{it}, \epsilon_{it-1})$ in the model, the previous expression for the IV estimator becomes

$$\frac{\text{cov}(\Delta c_{it}, \epsilon_{it})}{\text{cov}(\Delta y_{it}, \epsilon_{it})} = \frac{\text{cov}(\eta_{it}, \Delta c_{it}) - \rho^{-2} \text{cov}(\Delta c_{it}, \epsilon_{it-1})}{\text{var}(\eta_{it}) + (1 - \rho) \left[ \text{var}(z_{it-3}) - \rho \text{var}(z_{it-1}) \right]}.$$  \hspace{1cm} (11)

Since in the model $\text{cov}(\Delta c_{it}, \epsilon_{it-1}) < 0$ for constrained agents, the model IV estimates will contain an additional upward bias. Figure A.2 plots the age profile of $\phi$ estimated by the IV estimator against the age.
profile of true consumption responses in the calibrated life-cycle incomplete markets model as in Section 5 in the main text. The differences between the true and the IV estimated transmission coefficients in the model come from both mean-reversion ($\rho = 0.964$) and history dependence of consumption. As can be seen in Figure A.2, the upward bias of the IV estimator in the age profile of $\phi$ is quantitatively small in the calibrated model. The bias in the age profile of $\phi$ shows up mainly before age 40, when more of the households are constrained in the model.

Figure A.2: Performance of IV estimator $\hat{\phi}$ in the calibrated model
Figure A.3: Age profile of consumption response (rolling estimates), in BPP original sample

Figure A.4: “Searching for breaks” in BPP original sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Variable</th>
<th>Estimate</th>
<th>Variable</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln c</td>
<td>0.970</td>
<td>Age$^2$</td>
<td>-0.000349</td>
<td>Born 1955-59</td>
<td>-0.0825</td>
</tr>
<tr>
<td></td>
<td>(0.132)</td>
<td></td>
<td></td>
<td></td>
<td>(0.0581)</td>
</tr>
<tr>
<td>ln c × high school graduate</td>
<td>-0.0294</td>
<td>Northeast</td>
<td>0.0325</td>
<td>Born 1950-54</td>
<td>-0.112</td>
</tr>
<tr>
<td></td>
<td>(0.0260)</td>
<td></td>
<td>(0.00527)</td>
<td></td>
<td>(0.0529)</td>
</tr>
<tr>
<td>ln c × college or above</td>
<td>-0.232</td>
<td>Midwest</td>
<td>0.00522</td>
<td>Born 1945-49</td>
<td>-0.164</td>
</tr>
<tr>
<td></td>
<td>(0.0271)</td>
<td></td>
<td>(0.0111)</td>
<td></td>
<td>(0.0481)</td>
</tr>
<tr>
<td>ln c × (year-1980)</td>
<td>-0.00418</td>
<td>South</td>
<td>0.0149</td>
<td>Born 1940-44</td>
<td>-0.177</td>
</tr>
<tr>
<td></td>
<td>(0.000374)</td>
<td></td>
<td>(0.0114)</td>
<td></td>
<td>(0.0438)</td>
</tr>
<tr>
<td>ln c × one child</td>
<td>0.0964</td>
<td>Family size</td>
<td>0.0700</td>
<td>Born 1935-39</td>
<td>-0.179</td>
</tr>
<tr>
<td></td>
<td>(0.0368)</td>
<td></td>
<td>(0.0176)</td>
<td></td>
<td>(0.0389)</td>
</tr>
<tr>
<td>ln c × two children</td>
<td>0.123</td>
<td>ln $p_{food}$</td>
<td>-0.789</td>
<td>Born 1930-34</td>
<td>-0.179</td>
</tr>
<tr>
<td></td>
<td>(0.0381)</td>
<td></td>
<td>(0.101)</td>
<td></td>
<td>(0.0340)</td>
</tr>
<tr>
<td>ln c × three children+</td>
<td>0.165</td>
<td>ln $p_{transports}$</td>
<td>0.475</td>
<td>Born 1925-29</td>
<td>-0.186</td>
</tr>
<tr>
<td></td>
<td>(0.0315)</td>
<td></td>
<td>(0.113)</td>
<td></td>
<td>(0.0302)</td>
</tr>
<tr>
<td>One child</td>
<td>-0.858</td>
<td>ln $p_{fuel+utils}$</td>
<td>-0.878</td>
<td>Born 1920-24</td>
<td>-0.195</td>
</tr>
<tr>
<td></td>
<td>(0.349)</td>
<td></td>
<td>(0.0918)</td>
<td></td>
<td>(0.0263)</td>
</tr>
<tr>
<td>Two children</td>
<td>-1.072</td>
<td>ln $p_{alcohol+tobacco}$</td>
<td>-0.716</td>
<td>Born 1910-19</td>
<td>-0.113</td>
</tr>
<tr>
<td></td>
<td>(0.372)</td>
<td></td>
<td>(0.121)</td>
<td></td>
<td>(0.0225)</td>
</tr>
<tr>
<td>Three children+</td>
<td>-1.516</td>
<td>ln $p_{cpi}$</td>
<td>2.656</td>
<td>White</td>
<td>0.0899</td>
</tr>
<tr>
<td></td>
<td>(0.310)</td>
<td></td>
<td>(0.287)</td>
<td></td>
<td>(0.0177)</td>
</tr>
<tr>
<td>High school graduate</td>
<td>0.242</td>
<td>Born 1970-74</td>
<td>0.0151</td>
<td>Constant</td>
<td>-2.009</td>
</tr>
<tr>
<td></td>
<td>(0.237)</td>
<td></td>
<td>(0.0758)</td>
<td></td>
<td>(1.036)</td>
</tr>
<tr>
<td>College and above</td>
<td>2.099</td>
<td>Born 1965-69</td>
<td>-0.0180</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.241)</td>
<td></td>
<td>(0.0694)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.0410</td>
<td>Born 1960-64</td>
<td>-0.0489</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00536)</td>
<td></td>
<td>(0.0637)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observations: 50108

The last two years are included as anchor points, because in my experimentation they significantly improve the fit of the imputation procedure (as measured by a CEX in-sample validation exercise, results of which are available upon request). The left-hand side variable is log nominal food expenditure. Right-hand side variables are log real nondurable expenditure plus interactions, demographic controls and price levels. I instrument log real nondurable expenditure (and their interactions) with the cohort-year specific average of the log husband’s and wife’s hourly real wage rates (and their interactions with age, education, and a time trend) as per recommendations in Campos and Reggio (2014). The standard errors are in parentheses. The p-value from both the F test of excluded instruments and the Angrist-Pischke multivariate F test of excluded instruments is smaller than 0.0001 for all endogenous variables.