CENTRO STUDI IN ECONOMIA E FINANZA

CENTRE FOR STUDIES IN ECONOMICS AND FINANCE

WORKING PAPER NO. 23

Intertemporal choice and consumption mobility

Tullio Jappelli and Luigi Pistaferri

First draft: October 1999

This version: July 2003

DIPARTIMENTO DI SCIENZE ECONOMICHE - UNIVERSITÀ DEGLI STUDI DI SALERNO
Via Ponte Don Melillo - 84084 FISCIANO (SA)
Tel. 089-96 3167/3168 - Fax 089-96 3169 – e-mail: csef@unisa.it

CSEF WORKING PAPER NO. 23

Intertemporal choice and consumption mobility

Tullio Jappelli* and Luigi Pistaferri**

Abstract

The theory of intertemporal consumption choice makes sharp predictions about the evolution of the entire distribution of household consumption, not just about its conditional mean. In a first step, we study the empirical transition matrix of consumption using a panel drawn from the Bank of Italy Survey of Household Income and Wealth. In a second step, we estimate the parameters that minimize the distance between the empirical and the theoretical transition matrix of the consumption distribution. The transition matrix generated by our estimates matches remarkably well the empirical matrix, both in the aggregate and in samples stratified by education. Our estimates strongly reject the consumption insurance model and suggest that households smooth income shocks to a lesser extent than implied by the permanent income hypothesis.

Keywords: Consumption dynamics, mobility

JEL Classification: D52; D91; I30.

Acknowldegments: We thank Orazio Attanasio, Chris Carroll, Miles Kimball, Ra.aele Miniaci, John Pencavel, and seminar partecipants in Padua, Seattle, and Michigan for comments. The Training and Mobility of Researchers Network Program (TMR) and the "Taube Faculty Research Fund" at the Stanford Institute for Economic and Policy Research provided financial support.

* CSEF, University of Salerno, and CEPR.

** Stanford University and CSEF, University of Salerno.

Table of contents

1. *Introduction*
   2. Tests of consumption mobility
   3. Measuring consumption mobility

4. Intertemporal choice and mobility
   4.1. The Permanent Income Hypothesis
   4.2. The rule-of-thumb model
   4.3. Consumption insurance
   4.4. Nesting the three models
   4.5. Measurement error

5. Estimation method

6. Estimation results
   6.1. Full sample estimates
   6.2. Group estimates
   6.3. Measurement error in income
   6.4. Relations with previous tests
   6.5. Implications for social mobility

7. Conclusions

References

1 Introduction

The theory of intertemporal choice suggests that measures of household welfare should be based on consumption, not income. For instance, the permanent income hypothesis implies that households set consumption equal to permanent income, smoothing out transitory income fluctuations, so that people who are currently "income-poor" are not necessarily "permanent income-poor". Therefore in this model the cross-sectional variance of consumption equals the cross-sectional variance of permanent income. On the other hand, the theory of full consumption insurance implies that the cross-sectional variance of consumption is constant over time. Departing from these insights, several recent studies have examined trends in consumption inequality in the US and elsewhere. Some of these studies are primarily descriptive; others examine the validity of theoretical predictions by contrasting them with the data.[1]

Our point of departure from this literature is that measures of consumption inequality do not always provide an accurate measure of household behavior and welfare (and changes thereof). Consumption inequality is a static concept, and as such it cannot handle violations of the life cycle-permanent income hypothesis (such as borrowing constraints or myopic behavior), which would imply a role for transitory income fluctuations over and above permanent fluctuations, or buffer stock behavior, which would imply more smoothing of income shocks than predicted by the standard model. The handling of these issues, we argue, calls for an analysis of consumption mobility.

The distinction between consumption inequality and consumption mobility is, effectively, a distinction between static and dynamic features of a distribution. Inequality refers to the dispersion of consumption at a point in time. Mobility describes movements within the consumption distribution as time goes by. Studies of consumption inequality may record no change in the dispersion of the underlying distribution even in the presence of intra-

[1] See Cutler and Katz (1992), Deaton and Paxson (1994), Johnson and Shipp (1997), Blundell and Preston (1998), Blundell, Pistaferri and Preston (2003).

distributional movements, with direct implications for welfare analysis. Despite the importance of these issues, to the best of our knowledge the present paper represents the first attempt to analyze consumption mobility, both theoretically and empirically.[2] As we shall see, the analysis of consumption mobility delivers new implications of various theoretical models of intertemporal choice and generates new empirical tests and insights of those models.

The paper attempts to understand which model of intertemporal consumption choice is capable of explaining the amount of consumption mobility we observe in the data. We focus on several consumption theories, among which the theory of consumption insurance, the rule-of-thumb model, and the PIH model have received the widest attention. We nest these popular consumption models and estimate the parameters that minimize the distance between the empirical and the theoretical transition matrix of the consumption distribution. The exercise is performed constructing a transition matrix for consumption and testing different hypotheses concerning consumption dynamics. Since to measure consumption mobility one needs to follow households over time, the empirical analysis is conducted on a panel drawn from the Bank of Italy Survey of Household Income and Wealth for the years 1987 to 1995. The survey we use is representative of the Italian population, spans nine years of data, contains a measure of total non durable consumption and has good quality income data. Since there are virtually no panel datasets with broad consumption measures, a by-product of this paper is to bring the data set to the attention of empirical macroeconomists.

To see how the theory of intertemporal choice delivers implications for consumption mobility, consider first the extreme case of full consumption insurance. According to this theory, the cross-sectional distribution of consumption of any group of households is constant over time. Of course aggregate consumption can increase or decrease, so that consumption growth for any household can be positive or negative, but the relative position of each household in the cross-sectional distribution does not change over time. Consumption insurance makes

[2]In contrast, there is a long tradition of studies of earnings and income mobility. Existing contributions can be divided into two broad groups. A first group analyzes transition probabilities across quantiles of the earnings distribution by Markov-chain models (e.g., Shorrocks, 1978). A second approach is to specify and estimate a process for the conditional mean of earnings (e.g., Lillard and Willis, 1978).

therefore strong predictions about the entire consumption distribution, not just its mean or variance. In particular, consumption insurance implies absence of consumption mobility between any two time periods, regardless of the nature of the individual income shocks and the time frame considered. If one observes people moving up and down in the consumption distribution one must therefore conclude that some people are not insulated from idiosyncratic shocks, a contradiction of the consumption insurance hypothesis.[3]

A second case we consider is the rule-of-thumb model which predicts that households set consumption equal to income in each period. Given that any change in current income translates into an equivalent change in consumption, one should expect a relatively high degree of consumption mobility if shocks are not correlated with the rank position in the initial distribution of consumption.

In more realistic models with incomplete markets and insurance opportunities, individuals use saving as a self-insurance device and are able to smooth away at least some of the income variability. Within this class of models, the best known is the PIH, in which income shifts over time because of transitory (e.g., mean reverting) and permanent (e.g., persistent or non-mean reverting) shocks. If people behave according to the PIH, consumption reacts mostly to permanent unanticipated income shocks but is almost insensitive to transitory ones. Households will therefore move up and down in the consumption distribution only in response to permanent shocks. Thus one should expect a degree of mobility that is intermediate between the level predicted by the consumption insurance hypothesis and the rule-of-thumb model, a proposition that is formally proved in the Appendix.

The rest of the paper is organized as follows. Section 2 presents the mobility index and the test of consumption mobility. The data and the empirical results are presented in Section 3. In Section 4 we review the implications for consumption dynamics of the theories of intertemporal consumption choice and consider how to account for measurement error in consumption. In Section 5 we estimate the parameters of the consumption rule and the amount of measurement error in consumption by minimizing the distance between

[3]Although this implication of consumption insurance was mentioned in a theoretical paper by Banerjee and Newman (1991), to our knowledge it has never been explored empirically.

the empirical and the simulated transition matrix of the consumption distribution. The results, presented in Section 6, reject statistically each of the simple representations of the consumption decision rule, and reveal that households smooth income shock to a lesser extent than implied by the PIH. The estimated parameters are also able to reproduce remarkably well the difference in consumption mobility that we observe in samples stratified according to education and have interesting implications for analyzing the determinants of social mobility. Section 7 summarizes our results.

2 Tests of consumption mobility

To summarize the transition matrix for consumption through an appropriate index of mobility, we build on Shorrocks' approach (1978). Assume that $\mathbf{P}$ is an unobservable $q \times q$ stochastic transition matrix of household consumption, q being the number of consumption classes in the distribution. These classes could be determined exogenously or estimated from the quantiles of the empirical distribution. For notational simplicity we consider transition probabilities from period t to period $t+1$; extending the argument to transition probabilities in periods $t+2$, $t+3$, and so on, is straightforward. The generic element of $\mathbf{P}$ is p_{ij}, the probability of moving from class i in period t to class j in period $t+1$ conditioning on being in class i in period t. Define n_{ij} as the number of households that move from class i in period t to class j in period $t+1$, $n_i = \sum_{i=1}^{q} n_{ij}$ as the total number of observations in each row i of $\mathbf{P}$, and $n = \sum_{i=1}^{q} n_i$ the total number of observations. The maximum likelihood estimator of the first-order Markov transition probabilities is $\widehat{p}_{ij} = \frac{n_{ij}}{n_i}$ (Anderson and Goodman, 1957). The Shorrocks index of mobility is then defined as:

$$S(\mathbf{P}) = \frac{q - trace(\mathbf{P})}{q-1} \qquad (1)$$

Shorrocks (1978) proves that this mobility index satisfies a series of desirable properties, such as that of normalization ($0 \leq S(\mathbf{P}) \leq 1$), monotonicity ($S(\mathbf{P})$ increases with mobility), strong immobility, and strong perfect mobility. In particular, if the probability of being in class i in period t equals the probability of being in class j in period $t+1$, the typical entry of the transition matrix is $p_{ij} = 1$ for all $i = j$ and 0 otherwise. In this case $trace(\mathbf{P}) = q$, and $S(\mathbf{P}) = S(\mathbf{I}) = 0$, where $\mathbf{I}$ is the identity matrix. This is a case of strong immobility. If the probability of being in class i in period t is independent of that of being in class j in period $t + 1$, the typical entry of the transition matrix is $p_{ij} = q^{-1}$ for all i and j, $trace(\mathbf{P}) = 1$, and $S(\mathbf{P}) = 1$. This is a case of strong perfect mobility. Therefore $S(\mathbf{P})$ can be interpreted as the average probability across all classes that an individual will leave her initial class in the next period.

In the empirical analysis we will be interested in comparing statistically actual consumption mobility with that implied by theoretical models of consumption and in assessing whether consumption mobility differs statistically over time or between population groups. One way of making inference about an empirical transition matrix relies on the Shorrocks index itself.

Let's assume that class boundaries are exogenously fixed, and consider the theory of statistical inference on independent sample proportions.[4] The central limit theorem implies that as $n \to \infty$, $trace(\widehat{\mathbf{P}}) \overset{a}{\sim} N\left(\sum_{i=1}^{q} p_{ii}; \sum_{i=1}^{q} \frac{p_{ii}(1-p_{ii})}{n_i}\right)$, so that $S(\widehat{\mathbf{P}})$, the maximum likelihood estimator of $S(\mathbf{P})$, is asymptotically normally distributed:

$$S\left(\widehat{\mathbf{P}}\right) \overset{a}{\sim} N\left(\frac{q - \sum_{i=1}^{q} p_{ii}}{q-1}; \frac{1}{(q-1)^2} \sum_{i=1}^{q} \frac{p_{ii}\left(1 - p_{ii}\right)}{n_i}\right),$$

and one can test the null hypothesis that $S(\mathbf{P})$ assumes a given value $S\left(\mathbf{P}^0\right)$ using the statistic:

$$\frac{\frac{q-\sum_i \widehat{p}_{ii}}{q-1} - S\left(\mathbf{P}^0\right)}{\sqrt{\frac{1}{(q-1)^2} \sum_i \frac{\widehat{p}_{ii}\left(1-\widehat{p}_{ii}\right)}{n_i}}} \sim N(0,1) \tag{2}$$

where $\mathbf{P}^0$ is the transition matrix under the null hypothesis.

To assess if consumption mobility differs statistically over time or between population groups one can extend the test of the difference between two sample proportions and construct a test of differential mobility between groups, based on the statistic:

[4] Inference when class boundaries are estimated rather than exogenously fixed is discussed in Formby, Smith and Zheng (2001). In the empirical application we neglect this source of extra randomness.

$$\frac{S(\widehat{\mathbf{P}}^{g_0}) - S(\widehat{\mathbf{P}}^{g_1})}{\sqrt{s.e.(S(\widehat{\mathbf{P}}^{g_0}))^2 + s.e.(S(\widehat{\mathbf{P}}^{g_1}))^2}} \sim N(0,1) \tag{3}$$

where g_0 and g_1 are appropriately defined to allow comparisons over time or between population groups. Under the null hypothesis of no differential mobility between the two groups $(\mathbf{P}^{g_0} = \mathbf{P}^{g_1})$, the statistic (3) is also asymptotically distributed as a standard normal.

The main disadvantage of these tests is that the Shorrocks index of mobility is based on the trace of a matrix, and therefore the same index can be produced by very different underlying transition matrices. The modified χ^2 goodness-of-fit statistic proposed by Anderson and Goodman (1957) takes into account not only differences in the trace but also differences in the off-diagonal elements of the matrix. The statistic is defined as:

$$\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{\left(\widehat{p}_{ij} - p_{ij}^0\right)^2}{\widehat{p}_{ij}} \sim \chi^2_{q(q-1)} \tag{4}$$

and can be used to test the null hypothesis that $p_{ij} = p_{ij}^0$ for all i, j. As with the Shorrocks index, the χ^2 statistic allows also to test if the transition matrix differs statistically over time or between population groups:

$$\sum_{i=1}^{q}\sum_{j=1}^{q} N_i \frac{\left(\widehat{p}_{ij}^{g_0} - \widehat{p}_{ij}^{g_1}\right)^2}{\widehat{p}_{ij}} \sim \chi^2_{q(q-1)}$$

where $\widehat{p}_{ij}$ is the estimate of p_{ij} obtained pooling data for the two groups or time periods g_0 and g_1, and $N_i^{-1} = \frac{1}{n_i^{g_0}} + \frac{1}{n_i^{g_1}}$.

In the empirical application we will also be interested in matching the empirical transition matrix with a simulated matrix that depends on a vector of unknown parameters θ. This estimation problem can be addressed by implementing a minimum χ^2 method, i.e. minimizing the function:

$$\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{\left(\widehat{p}_{ij} - p_{ij}(\theta)\right)^2}{p_{ij}(\theta)} \tag{5}$$

The properties of this estimator are discussed in Neyman (1949). In the Appendix we show that (5) can be rewritten as:

$$\left(\widehat{\mathbf{p}}-\mathbf{p}\left(\theta\right)\right)\Omega\left(\theta\right)^{-1}\left(\widehat{\mathbf{p}}-\mathbf{p}\left(\theta\right)\right)' \tag{6}$$

where −after deleting a column from the theoretical and empirical transition matrices to avoid singularity− $\widehat{\mathbf{p}}$ is the vector of estimated transition probabilities, $\mathbf{p}\left(\theta\right)$ the vector of theoretical transition probabilities, and $\Omega\left(\theta\right)$ the covariance matrix of the distance vector $\left(\widehat{\mathbf{p}}-\mathbf{p}\left(\theta\right)\right)$. The function (6) has therefore the optimal minimum distance form of Chamberlain (1982) that econometricians are familiar with.

Neyman (1949) also proposed a modified minimum χ^2 method, where the function to minimize is

$$\sum_{i=1}^{q}\sum_{j=1}^{q}n_i\frac{\left(\widehat{p}_{ij}-p_{ij}\left(\theta\right)\right)^2}{\widehat{p}_{ij}}$$

The Appendix proves that this function can be rewritten as $\left(\widehat{\mathbf{p}}-\mathbf{p}\left(\theta\right)\right)\widehat{\mathbf{\Omega}}^{-1}\left(\widehat{\mathbf{p}}-\mathbf{p}\left(\theta\right)\right)'$, where $\widehat{\mathbf{\Omega}}$ uses the estimated $\widehat{p}_{ij}$ to construct an estimate of the covariance matrix of $\left(\widehat{\mathbf{p}}-\mathbf{p}\left(\theta\right)\right)$.

When the expression for $p_{ij}\left(\theta\right)$ is available in closed form, implementation of the minimum χ^2 criterion is straightforward. When it is not, as in our case, one must rely on simulations to generate the transition probability conditional on θ, and then apply the minimum χ^2 method to the simulated $p_{ij}\left(\theta\right)$. Details are provided in Section 5 and in the Appendix.

3 Measuring consumption mobility

From the previous section it is clear that mobility can only be computed with longitudinal data on consumption. For this purpose we use the 1987-1995 panel of the Italian Survey of Household Income and Wealth (SHIW). This data set contains measures of consumption, income, and demographic characteristics of households. The SHIW provides a measure of total non-durable consumption, not just food, thus overcoming one of the main limitations of other panels, such as the PSID, that have been used to test for intertemporal consumption choice.

The SHIW is conducted by the Bank of Italy which surveys a representative sample of the Italian resident population. Sampling is in two stages, first municipalities and then households. Municipalities are divided into 51 strata defined by 17 regions and 3 classes of population size (more than 40,000, 20,000 to 40,000, less than 20,000). Households are randomly selected from registry office records. From 1987 through 1995 the survey was conducted every other year and covered about 8,000 households, defined as groups of individuals related by blood, marriage or adoption and sharing the same dwelling. Starting in 1989, each SHIW has re-interviewed some households from the previous surveys. The panel component has increased over time: 15 percent of the sample was re-interviewed in 1989, 27 percent in 1991, 43 percent in 1993, and 45 percent in 1995.[5] The response rate (ratio of responses to contacted households net of ineligible units) was 25 percent in 1989, 54 percent in 1991, 71 percent in 1993, and 78 percent in 1995.[6] While these figures uncover considerable sample attrition especially in the early years of the survey, they are comparable to those obtained in other microeconomic data sets. For instance, in 1994 the net response rate in the US Consumer Expenditure Survey was 83 percent for the Interview sample and 81 percent for the Diary sample. Given the rotating sample structure, the number of repeated observations on households in our sample ranges from a minimum of two to a maximum of five. Ample details on sampling, response rates, processing of results and comparison of survey data with macroeconomic data are provided by Brandolini and Cannari (1994).

The total number of consumption transitions is 10,508. To minimize measurement error we exclude cases in which the head changes over the sample period or gives inconsistent age figures. In most cases, the excluded households are those facing breaking-out events (widowhood, divorce, separation, etc.), leading to changes in household head. Inconsistent age figures can reflect unrecorded change in household head or measurement error. After these exclusions, the sample has 9,214 consumption and income transitions. Consumption is the sum of all expenditure categories except durables. Income is defined as the sum of labor income and transfers of all household members, excluding income from assets. These are the

[5] In the panel component, the sampling procedure is also determined in two stages: (i) selection of municipalities (among those sampled in the previous survey); (ii) selection of households reinterviewed. This implies that there is a fixed component in the panel (for instance, households interviewed 5 times between 1987 to 1995, or 4 times from 1991 to 1995) and a new component every survey (for instance, households reinterviewed only in 1989).

[6] Response rates increase in 1991 because in that year households included in the panel were chosen among those that had previously expressed their willingness to being re-interviewed (Brandolini, 1999).

standard consumption and income concepts used in studies that test the implications of the permanent income hypothesis.[7]

Table 1 reports sample statistics of log consumption, income and other household characteristics. All statistics are computed using sample weights. The panel is relatively stable over the sample period. Consumption grows considerably between 1987 and 1989 and is roughly constant afterwards. Over time, family size declines while the number of income recipients increases. Other demographic characteristics remain roughly unchanged. Self-employment slightly falls over time. Income strongly declines in 1993, a recession year, and consequently dispersion increases. In all years, household disposable income is more variable than consumption. Note also the stability of the cross-sectional variance of log consumption as opposed to the wide fluctuations in the cross-sectional variance of log income. The pattern of the Gini coefficients for consumption and income confirms that the income distribution is less equal than the consumption distribution (34 vs. 28 percent). Interestingly, the 1993 recession boosts income inequality while leaving consumption inequality unaffected. As pointed out by Deaton and Paxson (1994), these descriptive statistics are consistent with models in which households are able to smooth away at least some of the income shocks. The focus of the present analysis, however, is not consumption inequality but consumption mobility. For this purpose, we need to construct a consumption transition matrix.

There are two methods for constructing such matrix. One is to keep the width of the consumption interval constant and let the number of observations within each interval vary. The alternative, more standard method, is to keep constant the marginal probabilities and let the interval width change, for instance dividing the distribution into discrete quantiles. We proceed using quartiles throughout; results with deciles are qualitatively similar and are not reported for brevity.

An important advantage of studying transition probabilities is that they are not affected by any specific form for the utility function. As the ordering of household consumption is invariant to monotonic transformation of the utility function, so are quantile probabilities.

[7] Adding back asset income or asset income net of imputed rents does not change the main results of the paper.

We focus on the transition matrix of the logarithm of non-durable per capita consumption, rather than the level of per capita consumption, because this allows us to nest and simulate different models of consumption behavior. We check the sensitivity of the results using different consumption equivalent scales and consider the interactions between consumption and labor supply.

Table 2 reports the transition matrix of log per capita consumption from 1987-89 to 1993-95. Recall that the SHIW is conducted every two years, so we observe transitions from period $t-2$ to period t. The elements of the main diagonal report the proportion of households that did not change quartile. For instance, the entry in the top left cell of the 1993-95 panel indicates that 68 percent of the households in the first quartile in 1993 were still in that quartile two years later. Off-diagonal elements signal consumption mobility. For instance, the second entry in the first row indicates that 25 percent of households moved from the first quartile in 1993 to the second quartile in 1995. The transition matrices for other years are similar, displaying substantial amount of consumption mobility.

In the simulation analysis we will make the assumption that consumption mobility is generated by a symmetric distribution of income shocks, which implies that our simulated transition matrix is also symmetric. It is therefore of interest to check if the transition matrix is symmetric using the maximum likelihood test suggested by Bishop, Fienberg and Holland (1975). The statistic is of the form $\Psi = \sum_{i>j} \frac{(p_{ij}-p_{ji})^2}{p_{ij}+p_{ij}} \sim \chi^2_{q(q-1)/2}$. The p-value of the test is close to 1 for all years, and does not reject the null hypothesis of symmetry.

The mobility index $S(\widehat{\mathbf{P}})$ corresponding to each of the transition matrices in Table 2 is reported in Table 3 together with the associated standard error and the number of transitions. On average, there is a 60 per cent probability of moving up or down in the distribution over a period of two years. Consumption mobility ranges from 59 percent in 1993-95 to 67 percent in 1991-93, and is precisely estimated in each year. The swings in mobility that we observe after 1991 are likely to be associated with the deep 1991-93 recession and to the subsequent expansion of 1993-95.

The transition matrix and the associated mobility index based on the distribution of per capita consumption do not take into account the fact that household expenditures are affected by demographic variables and labor supply choice.

Changes in family size, for instance the arrival of children, alter family needs, hence consumption allocations. If household expenditures are characterized by economies of scale, one would observe mobility in consumption per capita even if the distribution of consumption per adult equivalent is constant over time. We thus compute transitions using log consumption per adult equivalent rather than per capita.[8] The pattern of the transition matrix and of the associated mobility index is unaffected. As a further check, we restrict attention to households whose demographic structure did not change over the sample period and find, again, similar consumption transitions.[9] In the remaining of the paper we thus focus on consumption per capita.

If leisure is an argument of the utility function, and if consumption and leisure are non-separable, consumption decisions are affected by predictable changes in households' labor supply (Attanasio, 2000). This implies that the dynamics of consumers' rank in the consumption distribution depends, among other things, on changes in hours of work. Failure to control for changes in labor supply might therefore induce consumption mobility even in the absence of income and other idiosyncratic shocks. The interaction between consumption and labor supply is unlikely to affect our results, however. First of all, in our sample hours worked by individual household members and the proportion of spouses working do not change appreciably over the period considered. Second, if we exclude households reporting changes in labor force participation, the consumption transition matrix is almost identical to the full sample matrix.

As it stands, the mobility index in Table 3 summarizes the transition matrix. In the next section we derive from theory meaningful null hypotheses against which data can be confronted. We explore the implications for consumption mobility implied by popular models

[8]There is a large literature on the cost of children and on the economies of scale in consumption, see Deaton (1997) for a survey. Any particular choice of an equivalence scale is therefore to a certain extent arbitrary, depending on the estimation method and assumptions about the utility function. We rely on a plausible equivalence scale that is consistent with current literature, assigning a weight of 1 to the first adult, 0.8 to any additional adult and 0.25 to each household member less than 18 years old. We obtain similar results changing the parameters of the equivalence scale within a range of realistic estimates (0.1 to 0.5 for children, 0.6 to 1 for adults).

[9]For instance, excluding households with changes in family composition results in a mobility index of 0.576 in 1993-95.

(PIH, rule-of-thumb, and consumption insurance). In Section 5 we discriminate among these models by estimating the parameters that minimize the distance between the empirical and theoretical transition matrix of the consumption distribution.

4 Intertemporal choice and mobility

To explore the relation between the consumption and the income distributions, it is useful to start by presenting a fairly general characterization of the income process. Starting with Hall and Mishkin (1982), it has become quite standard in panel data studies of income and consumption dynamics to express log income of household h in period t as:

$$\ln y_{h,t} = \beta X_{h,t} + p_{h,t} + e_{h,t} \tag{7}$$

where $X_{h,t}$ is a set of deterministic variables such as age and region of residence, $p_{h,t}$ and $e_{h,t}$ permanent and transitory components, respectively.[10] The latter is the sum of an idiosyncratic ($\varepsilon_{h,t}$) and an aggregate component (ε_t); both are assumed to be serially uncorrelated. Since the permanent component of income changes very slowly, the standard assumption is to model it as a random walk process of the form:

$$p_{h,t} = p_{h,t-1} + z_{h,t} \tag{8}$$

where $z_{h,t}$ is the permanent innovation, which is again the sum of an idiosyncratic ($\zeta_{h,t}$) and an aggregate shock (ζ_t); both components are serially uncorrelated. We also assume that $\varepsilon_{h,t}$ and $\zeta_{h,t}$ are mutually uncorrelated disturbances with variances σ^2_ε and σ^2_ζ, respectively. Since we operate with a short panel, transitory and permanent aggregate shocks will be estimated by a vector of time dummies, d_t, rather than used as random components.

The decomposition of income shocks into transitory and permanent components dates back to Friedman (1957). Some of the income shocks are transitory (mean reverting) and their effect does not last long. Examples include fluctuations in overtime labor supply, bonuses, lottery prizes, and bequests. On the other hand, some of the innovations to earnings

[10]The logarithmic transformation eliminates heteroskedasticity in the distribution of income in levels.

are highly persistent (non-mean reverting) and their effect cumulates over time. Examples of permanent innovations are generally associated with job mobility, promotions, lay-off, and severe health shocks.

Given our assumptions income growth can be written as:

$$\Delta \ln y_{h,t} = \Delta d_t + \beta \Delta X_{h,t} + \zeta_{h,t} + \Delta \varepsilon_{h,t} \tag{9}$$

As we shall see, this income process delivers different implications for consumption mobility for different models of intertemporal choice. We also consider how these implications change in the presence of measurement error in consumption.

4.1 The Permanent Income Hypothesis

We consider a version of the PIH with CRRA preferences, where infinitely lived households maximize expected utility under perfect credit markets, subject to an intertemporal budget constraint. We assume that income income follows the process (7)-(8) and that it is the only source of uncertainty of the model. As in Blundell and Preston (1998), we approximate the Euler equation for consumption with a second-order Taylor expansion and assume that $r = \delta$, that consumption equals permanent income, and that the conditional variance of income shocks varies only in the aggregate. One can show that under such assumptions, individual consumption growth depends on an aggregate component and unanticipated idiosyncratic income shocks:

$$\Delta \ln c_{h,t} = m_t^{PIH} + \frac{r}{1+r} \varepsilon_{h,t} + \zeta_{h,t} \tag{10}$$

Equation (10) indicates that the optimal rule is to respond one-to-one to permanent shocks and to revise consumption only by the annuity value of the income innovation in case of transitory shocks. This is in fact the basic insight of the PIH, where people self-insure against high-frequency income shocks but adjust their consumption fully in response to low-frequency shocks. As we shall see, a convenient feature of equation (10) is that it readily lends itself to be nested with consumption rules derived from different models.

Suppose now that we observe a given cross-sectional distribution of consumption at time $t-1$ and that the income shocks are not perfectly correlated with the consumption rank of each household in the cross-section. Since aggregate shocks are by definition identical for all households, they do not change each consumer's rank in the consumption distribution and therefore they will not induce any consumption mobility: if they were the sole source of consumption fluctuations the mobility index would be zero.[11] However, other shocks are idiosyncratic, and will move people up and down in the consumption distribution, to an extent that depends on the variance of the two shocks. But since the impact of transitory shocks is scaled down by the factor $\frac{r}{1+r}$, we expect the variance of the permanent shocks to have the greatest impact on mobility. The purpose of the simulations in the next section will be precisely to assess the amount of mobility that one should expect in the permanent income model for given parameters of the income process.

Recent simulation results produced by Carroll (2001) show that with constant relative risk aversion, impatient consumers and an income process similar to the one we use, the implication of the PIH that transitory income shocks have a negligible impact on consumption still holds true. Permanent shocks, however, have a somewhat lower impact in buffer stock models. In fact, in such models permanent income shocks reduce the ratio of wealth to permanent income, thus increasing also precautionary saving. Under a wide range of parameter values, Carroll shows that in this class of models the marginal propensity to consume of a permanent income shock is about 0.9, not far from that of the approximation in (10). Therefore, empirically it is difficult to distinguish the PIH from buffer stock models on the basis of the marginal propensity to consume out of permanent income shocks. But the main intuition is still valid: if individuals smooth consumption and understand the income generating process, transitory income shocks should have a negligible impact on consumption. To account for the effect remarked by Carroll and others, in the empirical analysis we take into account the degree of consumption smoothing arising from precautionary savings estimating:

[11] Suppose that income shocks were instead perfectly and positively correlated with the rank of household consumption in the cross-section. Then, the poorest households receive the largest negative shocks and the richest the largest positive shocks, implying no mobility as in the consumption insurance case.

$$\Delta \ln c_{h,t} = m_t^{PIH} + \phi \left(\frac{r}{1+r} \varepsilon_{h,t} + \zeta_{h,t} \right)$$

In buffer stock models ($\phi < 1$), assets accumulated for precautionary purposes allow people to smooth income shocks to a larger extent than in the PIH model ($\phi = 1$).

4.2 The rule-of-thumb model

Let's assume that consumption equals income in each period, i.e.:

$$\ln c_{h,t} = \ln y_{h,t}$$

This model has been often proposed as a simple, yet extreme alternative to the PIH to describe the behavior of households that do not use savings to buffer income shocks but spend all they receive. Some authors rationalize this model by appealing to the presence of binding liquidity constraints in each period. We term it rule-of-thumb model because liquidity constrained consumers cannot borrow but can save, and react differently to positive and negative income shocks. The rule-of-thumb model is an interesting case to study because it represents an upper bound for the sensitivity of consumption to income shocks and may therefore approximate the behavior of consumers with short horizons or limited resources.

Using the income process above the dynamic of consumption is given by:

$$\Delta \ln c_{h,t} = m_t^K + \varepsilon_{h,t} - \varepsilon_{h,t-1} + \zeta_{h,t} \tag{11}$$

where m_t^K is the effect of the aggregate shocks on consumption in the rule-of-thumb model. According to the rule-of-thumb model the growth rate of consumption is therefore equally affected by current and lagged transitory shocks and by permanent shocks. The main difference with the PIH is that in the rule-of-thumb model transitory shocks impact one-to-one on consumption. It is precisely for this reason that in the rule-of-thumb model one should expect more consumption mobility than under the permanent income rule: there is another channel through which households can move to a different quartile from one period to the next.

4.3 Consumption insurance

To illustrate the implications of the theory of intertemporal choice with complete insurance markets, let us **keep** the assumption that households have preferences of the CRRA type, $u(c) = (1-\gamma)^{-1}c^{1-\gamma}$, where γ^{-1} is the elasticity of intertemporal substitution. The implications of the model are identical for any power utility function. As shown, among others, by Mas-Colell, Whinston and Greene (1995), the optimal transition law for consumption with complete markets can be obtained by assuming that there is a social planner who maximizes a weighted sum of individual households' utilities. The Lagrangian of this problem can be written as:

$$L = \sum_h \lambda_h \sum_s \sum_t \pi_{s,t} u(c_{h,s,t}) + \sum_s \sum_t \mu_{s,t} \left(C_{s,t} - \sum_h c_{h,s,t} \right)$$

where h, s and t are subscripts for household h in the state of nature s in period t, λ_h is the social weight for household h, $\mu_{s,t}$ is the Lagrange multiplier associated with the resource constraint, $\pi_{s,t}$ the probability of the realization of state s in period t, and $C_{s,t}$ aggregate consumption in state s and period t.

The first order condition can be written in logarithms as:

$$-\gamma \ln c_{h,s,t} = \ln \mu_{s,t} - \ln \lambda_h - \ln \pi_{s,t}$$

To obtain the growth rate of consumption, subtract side-by-side from the same expression at time $t-1$:

$$\Delta \ln c_{h,t} = -\gamma^{-1} \Delta \ln \mu_t + \gamma^{-1} \Delta \ln \pi_t \equiv m_t^{CI} \tag{12}$$

where we drop the subscript s because only one state is realized in each period. The two terms on the right-hand-side of equation (12) represent genuine aggregate effects. The first term is the growth rate of the Lagrange multiplier, the second is the growth rate of the state probabilities. Note that first-differencing has eliminated all household fixed effects (μ and π in equation 12 are not indexed by h).

Equation (12) states that the growth rate of consumption of each household is the same. This implies that the initial cross-sectional distribution of consumption is a sufficient statistic to describe all future distributions. Since all households experience the same consumption growth rate, their rank in the consumption distribution is stationary. Note that the stationarity of the cross-sectional distribution is directly implied by the assumption that insurance markets fully insulate households from idiosyncratic shocks. The statistical counterpart of consumption insurance is that the transition matrix for household consumption is an identity matrix. The extreme assumptions of this model are clearly unrealistic. However, the model provides a lower-bound for the impact of income shocks on consumption and is therefore a useful theoretical benchmark.

The discussion in Sections 4.1-4.3 can be summarized by the following:

Proposition 1 *Under CRRA preferences, consumption mobility is zero in the consumption insurance model, intermediate in the permanent income model, and highest in the rule-of-thumb model.*

Proof: See Appendix.

4.4 Nesting the three models

The distinction between the three models is useful but too stylized for empirical applications. Consumption insurance is no less unrealistic than assuming that all income is consumed in each period, or that all households follow exactly the PIH. In the empirical application we therefore nest the three models and estimate the parameters of the following flexible consumption rule:

$$\ln c_{h,t} = \ln c_{h,t-1} + \phi \left(\frac{\lambda + r}{1+r} \varepsilon_{h,t} - \lambda \varepsilon_{h,t-1} + \zeta_{h,t} \right) \tag{13}$$

Since aggregate shocks do not affect consumption mobility, for notational simplicity the equation above omits the aggregate component m_t^j. However, when we estimate the income process we control for aggregate shocks by introducing time dummies in the regression (d_t in equation 9).

The two parameters λ and ϕ allow to distinguish various forms of departure from the stylized models of intertemporal choice. Consider first the case in which $\phi = 1$. The parameter λ represents the extent to which consumption responds to income over and above the amount warranted by the PIH, i.e., the excess sensitivity of consumption to current and past income shocks. One way to interpret this parameter is that each household sets consumption equal to income with probability λ (perhaps because of binding liquidity constraints) and follows the PIH with probability $(1-\lambda)$. Note that with $\lambda = 0$ the expression (16) reduces to the PIH, while with $\lambda = 1$ one obtains the rule-of-thumb model where consumption equals income each period.

Consider now the situation in which $\phi = 0$. Income shocks play no role in the consumption insurance model. But intermediate cases in which $0 < \phi < 1$ are interesting and potentially informative, as discussed in Section 4.1. Some consumers have assets accumulated for precautionary reasons which allow them to smooth income shocks to a larger extent than in the PIH model (where $\phi = 1$).

4.5 Measurement error

The consumption transition law is derived assuming that there is no measurement error in consumption. In practice the index could potentially be upward biased by reporting errors. If respondents report their consumption with errors, one will find units moving up and down even if their true rank in the consumption distribution is unchanged; hence, measurement errors affect consumption dynamics and the mobility index in Section 2 will tend to report higher mobility. In the estimation it is therefore important to account explicitly for measurement error.

Suppose that true consumption is measured with a multiplicative error:

$$\ln c^*_{h,t} = \ln c_{h,t} + \widetilde{v}_{h,t} \tag{14}$$

where $\ln c^*$ is *measured* consumption and $\widetilde{v}$ is an independently and identically distributed measurement error. Without loss of generality, we assume that in each period the standard deviation of measurement error ($\sigma_{\widetilde{v}}$) is a fraction α of the standard deviation of measured consumption, $\sigma_{\widetilde{v}} = \alpha \sigma_{\ln c^*}$. Since the variables are expressed in logs, α can be interpreted as the percentage variability in observed consumption due to reporting error.

The consumption dynamics in equation (13) changes in the following way:

$$\ln c^*_{h,t} = \ln c^*_{h,t-1} + \phi \left(\frac{\lambda + r}{1 + r} \varepsilon_{h,t} - \lambda \varepsilon_{h,t-1} + \zeta_{h,t} \right) + \alpha \left(v_{h,t} - v_{h,t-1} \right) \quad (15)$$

where we have adopted the linear transformation $v_{h,\tau} = \alpha^{-1} \widetilde{v}_{h,\tau}$, so that $v \sim i.i.d.\,(0, \sigma_{\ln c^*})$, to make clear that α is an unknown parameter to estimate. Equation (15) shows that measurement error induces a further reason for consumption to vary. Clearly, not only consumption dynamics changes, but the implied consumption mobility as well.[12]

5 Estimation method

We now discuss estimation of the parameters of interest. One complication with the panel we use is that while income and consumption refer to calendar years, data are collected every other year from 1987 to 1995. The simulated transition laws for consumption must therefore be slightly modified to tackle this problem. Equation (15) rewrites as:

$$\begin{aligned} \ln c^*_{h,t} &= \ln c^*_{h,t-2} + \phi \left(\frac{\lambda + r}{1 + r} \varepsilon_{h,t} + \frac{(1-\lambda) r}{1 + r} \varepsilon_{h,t-1} - \lambda \varepsilon_{h,t-2} + \zeta_{h,t} + \zeta_{h,t-1} \right) \\ &\quad + \alpha \left(v_{h,t} - v_{h,t-2} \right) \end{aligned} \quad (16)$$

The parameters to be estimated are the variances of the permanent and transitory income shocks, the fraction of measurement error in consumption, the degree of excess sensitivity, the degree of income smoothing and the real interest rate. As for the interest rate, we assume a value of 2 percent throughout. We estimate the income variances σ^2_ε and σ^2_ζ in a first step. In a second step we use the estimated income variances to generate income shocks ε and ζ that appear in the consumption rule, and estimate the parameters ϕ, λ, and α by simulated minimum χ^2 method.

As explained in Section 4, we specify the income process as $\ln y_{h,t} = d_t + \beta X_{h,t} + p_{h,t} + \varepsilon_{h,t}$, where $y_{h,t}$ is per capita household disposable income and d_t a set of time dummies. Using the 1987-95 panel, we regress $\ln y_{h,t}$ on a set of demographic variables (North, South, a dummy for gender, a fourth-order age polynomial, and education dummies) and time dummies, so

[12] The clearest case in which this happens is in the model with consumption insurance (where $\phi = 0$): in the absence of measurement error there is absolutely no mobility in the consumption distribution.

to remove the deterministic component of income. We save the residuals $u_{h,t} = p_{h,t} + \varepsilon_{h,t}$ and carefully examine their covariance properties. We estimate covariances using equally weighted minimum distance methods, as suggested by Altonji and Segal (1997).[13]

We find that the estimated covariances are consistent with the income process in equations (7) and (8), i.e. that there is a random-walk permanent component and a serially uncorrelated transitory shock. Recall that because of the sample design of the SHIW we can only construct the covariance matrix for two years apart income residuals, $u_{h,t} - u_{h,t-2} = \zeta_{h,t} + \zeta_{h,t-1} + \varepsilon_{h,t} - \varepsilon_{h,t-2}$. To check the consistency of the estimated income process with the model in equations (7) and (8), note that the income process implies the following testable restrictions on the covariance matrix of the first difference of the income residuals:

$$\begin{aligned} E\left[(u_{h,\tau} - u_{h,\tau-2})^2\right] &= 2\sigma_\zeta^2 + 2\sigma_\varepsilon^2 \\ E\left[(u_{h,\tau} - u_{h,\tau-2})(u_{h,\tau-2} - u_{h,\tau-4})\right] &= -\sigma_\varepsilon^2 \\ E\left[(u_{h,\tau} - u_{h,\tau-2})(u_{h,\tau-j} - u_{h,\tau-j-2})\right] &= 0 \text{ for all } j \geq 4 \end{aligned}$$

Provided that the restrictions are met in the data, one can estimate the variance of the transitory shock σ_ε^2 from the first order autocovariance of income residuals and the variance of the permanent shock σ_ζ^2 combining information on the variance and the first-order autocovariance of the residuals. We find that the estimated autocovariance at the second order is very small (-0.0056) and not statistically different from zero (a t-statistic of -1.1); the autocovariance at the third order is again small (-0.0178) and not statistically different from zero (a t-statistic of -1.1). In contrast, the first order autocovariance (which provides an estimate of $-\sigma_\varepsilon^2$) is precisely estimated (a t-statistics of 6.4) at -0.0794. The estimate of the overall variance $(2\sigma_\zeta^2 + 2\sigma_\varepsilon^2)$ is 0.2122 (with a t-statistics of 19.4), so we infer that $\sigma_\zeta^2 = 0.0267$ and $\sigma_\varepsilon^2 = 0.0794$.[14] These parameter estimates are broadly consistent with the evidence available for the US, where researchers have found variances of similar

[13]Covariances can be estimated by equally weighted minimum distance or optimal minimum distance. As shown by Altonji and Segal (1997), the latter can produce inconsistent estimates in small samples, so we adopt the former.

[14]Unfortunately, with data collected every two years we cannot distinguish between this income process and one where the transitory component is an MA(1) process.

magnitude.[15]

The remaining unknown parameters are ϕ, the degree to which consumers are unable to insure income shocks through precautionary savings, λ, the degree of excess sensitivity of consumption, and α, the fraction of the cross-sectional variance of measured log consumption that is due to measurement error. Each of the three parameters ranges from 0 to 1.[16] We therefore estimate ϕ, λ and α minimizing the distance between the empirical and the theoretical transition matrix using the modified χ^2 criterion presented in Section 2.

Since theoretical transition probabilities do not have a closed form expression, we use a simulated minimum χ^2 estimation method.[17] A sketch of the estimation method is the following. We start by generating, for each household, draws for the transitory and the permanent income shocks and for the measurement error in consumption.[18] The income shocks are drawn from a normal distribution with mean zero and variances equal to the estimated variance from the income process ($\sigma_\varepsilon^2 = 0.0794$ and $\sigma_\zeta^2 = 0.0267$, respectively). The measurement errors $v_{h,t}$ and $v_{h,t-2}$ are drawn from a normal distribution with mean zero and variance equal to the variance of measured log consumption at t and $t-2$, respectively. The number of draws is set to $S = 100$ for each household, for a total of HS simulated observations (H being the number of households). We then choose a starting value for the parameter vector and, for each household, compute next period consumption, $\ln c_{h,t}^*$. We finally compute the theoretical transition probabilities (averaging across the S simulations) and obtain the parameter estimates as those that minimize the (optimal) distance between empirical and theoretical transition probabilities. The Appendix reports technical details about the properties of this estimator and the minimization algorithm.

[15] For instance, Carroll and Samwick (1997) using the PSID, estimate $\sigma_\zeta^2 = 0.0217$ and $\sigma_\varepsilon^2 = 0.0440$.

[16] Recall that $\alpha = 0$ implies no measurement error, while $\alpha = 1$ signals that the variability in measured consumption is entirely explained by measurement error.

[17] Alan and Browning (2003) estimate the parameters of the Euler equation (the elasticity of intertemporal substitution and the amount of measurement error) using simulated Euler equation residuals. Gourinchas and Parker (2002) estimate the coefficient of relative risk aversion and the intertemporal discount rate using a method of simulated moments conditional on the assumption that the PIH is the true consumption model.

[18] In each year we choose a sample size identical to the number of actual sample transitions (for instance, 2,982 in 1991-93 and 3,211 in 1993-95).

6 Estimation results

In this section we report full sample estimates of the parameters of the consumption rule and of the transition matrix for consumption. We also split the sample by educational attainment of the head, estimate a separate income process for each education group and evaluate patterns of consumption mobility of households with different levels of educational attainment. Finally, we check the sensitivity of the estimates to the presence of measurement error in income.

6.1 Full sample estimates

The results of the full sample estimates are similar across periods, so we focus on the most recent one (1993-95), that also features the largest number of transitions. The stability of the results across different sample periods suggests that the simulations are only marginally affected by the initial distribution of consumption (the income process and the associated variances of the shocks are in fact assumed to be the same across the different samples).

As a preliminary analysis, we constrain the parameter space in the simulated minimum χ^2 estimation method and compare the empirical and theoretical transition matrices in three benchmark models: PIH ($\phi = 1,\ \lambda = 0$), rule-of-thumb ($\phi = 1,\ \lambda = 1$) and consumption insurance ($\phi = 0$). These benchmark models illustrate our estimation strategy and provide a gateway to the results that follow.

We first evaluate the three benchmark models in the absence of measurement error ($\alpha = 0$). The Shorrocks mobility index is highest in the rule-of-thumb model (65 percent), intermediate in the case of the permanent income hypothesis (44 percent), and zero under consumption insurance. The ranking of mobility agrees with Proposition 1 because idiosyncratic income shocks translate into consumption changes entirely in the rule-of-thumb model, partially in the PIH via intertemporal smoothing, and are fully insured in the risk sharing model. However, from a statistical point of view, none of these models is able to match the amount of empirical mobility. Recall from table 3 that the empirical Shorrocks index is 59 percent. The hypothesis that the simulated Shorrocks index equals the empirical one is rejected in each of the models considered. Cell-by-cell comparison of the theoretical and empirical transition matrices reveals that each of the three models is rejected also according to the χ^2 goodness-of-fit statistics.[19]

To bridge the gap between simulated and empirical mobility we therefore consider the effect of measurement error in consumption and allow for a more flexible response of consumption to income shocks than predicted by either full insurance, rule-of-thumb model or PIH. As we know from Section 4.5, measurement error always increases consumption mobility, regardless of the model considered. We also know from equation (16), nesting the three baseline models, that raising the excess sensitivity parameter λ or the insurance parameter ϕ also increases consumption mobility, regardless of the size of the measurement error.

We therefore implement the simulated minimum χ^2 estimation method freeing the parameter space. The parameter estimates of ϕ, λ and α are reported in column (1) of Table 4. Since the restriction $\phi = 1$ is not rejected on economic and statistical grounds, we impose the restriction in column (2). The results indicate that the variability of consumption due to measurement error is 38 percent and that the excess sensitivity coefficient is 16 percent. Both estimates are precisely estimated and statistically different from zero at the 1 percent level. For values of $\phi = 1$, $\lambda = 0.16$ and $\alpha = 0.38$, the simulated mobility index is almost identical to the empirical one (60.13 against 59.37 percent). The χ^2 goodness of fit statistic is 15 with a p-value of 9 percent, indicating that the model fits well the transition probabilities: we cannot reject the hypothesis that the empirical transition probabilities are jointly equal to the simulated ones.[20]

The simulation predicts almost perfectly the empirical transition matrix cell-by-cell, not just the aggregate mobility index. In Table 5 we report the simulated transition probabilities and (in parenthesis) the empirical transition probabilities, the same reported for 1993-95 in

[19] In the rule-of-thumb case ($\alpha = 0$, $\lambda = \phi = 1$), the χ^2 value is 58, in the PIH ($\alpha = \lambda = 0$, $\phi = 1$) 250, and in the consumption insurance case ($\alpha = \lambda = \phi = 0$) 2856. Each of these values exceeds the critical value of $\chi^2_{12;0.05} = 21$.

[20] Results for other years are similar with the exception of 1991-93. In that period actual mobility increases to 67 percent, a fact that is not captured by our simulations. One possible explanation is that the variance of the permanent shock, which is assumed to be time stationary, changed in 1993 due to the unprecedented strong recession. However, we cannot rule out that in 1993 the amount of measurement error is greater than in the other two years. Another possibility is that the 1993 recession impacted unevenly on households, a particular form of non-stationarity that we neglect in our simulation exercise.

Table 2. The comparison between the two sets of numbers is striking: regardless of cell, the difference between the empirical and simulated values is at most 2 percentage points.

The estimated value of the excess sensitivity parameter ($\lambda = 0.16$) is broadly consistent with previous evidence on the effect of transitory income shocks on consumption expenditure. Using CEX quarterly panel data, Souleles (1999) and Parker (1999) examine, respectively, the response of household consumption to income tax refunds and to predictable changes in Social Security with-holdings. Souleles finds evidence that the marginal propensity to consume is at least 35 percent of refunds within a quarter, and Parker that consumption reacts significantly to changes in tax rates. In both studies, the impact of transitory income shocks is too high to be consistent with the PIH model, but in the range of estimates produced by our hybrid model. Browning and Crossley (2001) survey several other studies reporting evidence that consumption overreacts to anticipated income innovations.

6.2 Group estimates

Except for the extreme case of consumption insurance, models with incomplete insurance suggest that if different population groups are systematically exposed to different idiosyncratic shocks (and therefore face different income processes), consumption mobility should differ across groups in a predictable way. Therefore, comparison of different population groups with different income generating process is potentially quite interesting. Indeed, even more compelling evidence for the ability of our simulations in explaining consumption transitions comes from comparing consumption mobility in two education groups: compulsory schooling or less and high school or college degree.

Focus on education is warranted for at least three reasons: (1) education is an exogenous characteristic by which one can partition the sample; (2) there is wide evidence that different education groups face different earnings opportunities and uncertainties; (3) education is likely to be correlated with variables affecting preferences and therefore with different consumption behavior. We run the income regressions separately for households headed by individuals with high and low education. We then estimate the autocovariance matrix as explained in Section 5, and find $\sigma_{\zeta}^2 = 0.0296$ and $\sigma_{\varepsilon}^2 = 0.0754$ for the less well educated, and

$\sigma^2_\zeta = 0.0198$ and $\sigma^2_\varepsilon = 0.0895$ for those with at least a high school degree. The estimated variances signal that the less well educated face a higher variance of permanent income shocks, a pattern also uncovered by Carroll and Samwick (1997) with US data. Since in our sample the income process varies considerably by education groups, we have an ideal setting to test the validity of models of intertemporal choice and of the robustness of our procedure.

The summary statistics reported in Table 6 indicate that also consumption mobility differs between the two groups in a systematic way. The Shorrocks index is higher among the less well educated than among those with higher educated (0.62 against 0.55). Applying the test on difference of means outlined in Section 2, we reject the hypothesis that the two indexes are equal at the 1 percent significance level.

Quite clearly, the consumption insurance model is unable to explain differences in consumption mobility emerging from income shocks, transitory or permanent. In that model all shocks are insured, so consumption mobility between two groups exposed to different shocks should be identical. Therefore the fact that mobility is higher in the group with lower education provides further evidence against the consumption insurance model. For quite different reasons, the rule-of-thumb model with $\lambda = 1$ (or any model where excess sensitivity to transitory income shocks plays a prominent role) predicts little or no difference between education groups. In the simulations the lower variance of the transitory shock for the less well educated is offset by a higher variance of the permanent shock, resulting in approximately the same mobility rates in the two groups.

We therefore estimate equation (21) and the associated consumption transitions allowing for a flexible specification of the parameters of the consumption rule and differential response between the two education groups. The parameter estimates of the simulated minimum χ^2 method and the associated χ^2 statistic are reported in Table 6. Also in this case we cannot reject the hypothesis that $\phi = 1$ in each of the two groups. We find a value of $\alpha = 0.38$ (s.e. 0.01) and $\lambda = 0.4$ (0.05) in the group with low education and $\alpha = 0.28$ (0.01) and $\lambda = 0.09$ (0.05) in the group with high school or college degree. The model replicates quite well also the difference in empirical and simulated mobility between the two groups: the simulated mobility index (0.64 and 0.55 for low and high education, respectively) is quite close to empirical mobility in each group. And in each of the two cases the χ^2 statistic does not reject the null hypothesis that the simulated probabilities are equal the empirical ones at the 1 percent significance level.

As a final check of the validity of the estimates, we test whether the parameters are the same in the two groups. Call $\widetilde{\theta}_h$ and $\widetilde{\theta}_l$ the $k \times 1$ vectors of simulated minimum χ^2 estimates of θ for high- and low-educated individuals. Given the asymptotic normal distribution of the estimator and the fact that the two samples are independent, the null hypothesis of no group difference can be tested using the statistic:

$$\left(\widetilde{\theta}_h - \widetilde{\theta}_l\right)' \left(\text{var}\left(\widetilde{\theta}_h\right) + \text{var}\left(\widetilde{\theta}_l\right)\right)^{-1} \left(\widetilde{\theta}_h - \widetilde{\theta}_l\right)$$

which is distributed χ^2_k under the null. The test statistic, reported in the last row of the first panel of Table 6, rejects overwhelmingly the null hypothesis of parameter equality. The other two panels of Table 6 report the simulated transition probabilities and (in parenthesis) the empirical transition probabilities for the two education groups. Once more, each of the simulated probabilities is remarkably close to the empirical transitions irrespective of the group considered.

From an economic point of view, the result that the less well educated individuals are more responsive to transitory income shocks than the high income group is of particular interest. To the extent that these households are less likely to have access to credit and insurance markets than households with higher education, our findings support the hypothesis that excess sensitivity stems from the effect of borrowing constraints, rather than from other sources. The results that the less well educated report noisier consumption data is in line with intuition and expectations.

6.3 Measurement error in income

In this section we consider the robustness of our conclusions in the presence of measurement error in income. This error inflates the variance of the transitory shock but does not affect the variance of the permanent shock. To see this point, assume that true income is measured with a multiplicative error: $\ln y^*_{h,t} = \ln y_{h,t} + \omega_{h,t}$, where $\omega_{h,t}$ is an independently and identically normally distributed measurement error with mean zero and variance σ_ω^2. Using the income process (4)-(5): $\ln y^*_{h,t} = \beta X_{h,t} + p_{h,t} + \varepsilon_{h,t} + \omega_{h,t}$, the two years apart income residual is now: $u_{h,t} - u_{h,t-2} = \zeta_{h,t} + \zeta_{h,t-1} + \varepsilon_{h,t} - \varepsilon_{h,t-2} + \omega_{h,t} - \omega_{h,t-2}$. The covariance matrix of the first difference of the income residuals depends now on the variance of the measurement error:

$$
\begin{aligned}
E\left[\left(u_{h,\tau} - u_{h,\tau-2}\right)^2\right] &= 2\sigma_\zeta^2 + 2\sigma_\varepsilon^2 + 2\sigma_\omega^2 \\
E\left[\left(u_{h,\tau} - u_{h,\tau-2}\right)\left(u_{h,\tau-2} - u_{h,\tau-4}\right)\right] &= -\sigma_\varepsilon^2 - \sigma_\omega^2 \\
E\left[\left(u_{h,\tau} - u_{h,\tau-2}\right)\left(u_{h,\tau-j} - u_{h,\tau-j-2}\right)\right] &= 0 \text{ for all } j \geq 4
\end{aligned}
$$

However, it can be checked that measurement error inflates the estimated variance of the transitory shock by σ_ω^2, but not the variance of the permanent shock σ_ζ^2, which is still identified by the difference between the variance and (minus twice) the first-order autocovariance. The conclusion is that even though the estimate of the variance of the permanent shock is unaffected by serially uncorrelated measurement error, the estimate of the variance of the transitory shock is not.

This implies that in the model with full consumption insurance, idiosyncratic income shocks play no role regardless of measurement error in income. In the permanent income model, the impact of measurement error in income is bound to be small, because transitory shocks play a very limited role. In contrast, measurement error may have a large impact in the rule-of-thumb model. Since we cannot identify σ_ω^2 from the data, we repeat our simulation: (a) dropping the self-employed from the sample on which we estimate the income process,[21] and (b) downsizing the variance of the transitory shock, i.e., assuming that one third or one half of the estimated first-order autocovariance reflects measurement error. The results of these experiments are very similar to the simulations reported in Tables 4, 5 and 6 and are not reported for brevity.

[21] Brandolini and Cannari (1994) note that in the SHIW income from self-employment is less well estimated than wages or salaries.

6.4 Relation with previous tests

It is useful to contrast our approach with previous tests of models of intertemporal choice. First of all, our simulation method produces structural estimates of the propensity to consume out of transitory and permanent income shocks. These parameters are of great policy interest, for instance to evaluate the effect of a tax cut or other changes in the household budget constraint. Excess sensitivity of consumption has sometimes been inferred from the income growth coefficient in Euler equations estimates. However, there is much disagreement concerning the interpretation of the excess sensitivity parameter due to various identification problems in the estimation of the Euler equation (Attanasio, 2000). While the Euler equation literature is concerned with estimation of preference parameters derived from the first order conditions of the consumers' optimization problem, we attempt at estimating the structural parameters of the consumption rule. This does not come without costs, however. We make specific assumptions about preferences and the income generating process, and our estimates are therefore conditional on the validity of the theoretical framework and on the stability of the income process. This paper is therefore part of a growing literature in macroeconomics that attempts to estimate structural models by means of simulated estimation methods.

Second, and for quite different reasons, our approach to test for consumption insurance differs from previous tests based on regression analysis. Cochrane (1991), Mace (1991) and Townsend (1994) regress household consumption growth on aggregate variables and idiosyncratic shocks (such as change in disposable income, unemployment hours, and days of illness). The implication of the theory is that none of these shocks should impact household consumption growth, as in equation (12). Focussing instead on the prediction that consumption insurance implies absence of consumption mobility has the advantages that we need not identify any of these shocks, and that we need not assume that they are uncorrelated with unobservable or omitted preference shocks, including household fixed effects. Moreover, measurement error in the shock variables biases tests based on regression analysis towards the null hypothesis of full consumption insurance; our testing strategy is instead robust to such problem

Deaton and Paxson (1994) and Attanasio and Jappelli (2001) test another implication of the theory of consumption insurance, i.e., that the cross-sectional variance of consumption is constant over time. However, the distribution of consumption at time t might have the same variance of the distribution at time $t-1$ even if there is mobility in the underlying distributions.[22] Tests based on the dynamics of the cross-sectional variance of consumption are therefore biased towards the null. Our test instead still signals rejection of the consumption insurance model even in situations in which the cross-sectional variance is constant over time but there is mobility in the underlying distribution.

6.5 Implications for social mobility

Our estimates allow us to single out the separate contributions of incomplete markets, excess sensitivity, and measurement error in generating the actual mobility we observe in the data. Furthermore, we can characterize consumption mobility both in the short and in the long run by using recursively the transition law for consumption and the realizations of the income shocks (equation 16). We define as "poverty" being in the bottom consumption quartile in the initial period, so the concept we use is one of escaping relative poverty $(1-p_{11})$, not one defined in terms of absolute standards of living.

The theory of full consumption insurance delivers the most striking implications for poverty persistence. In that model, absent measurement error, anyone who happens to be poor in the initial period will be poor in relative terms in all subsequent periods. A society that insures all possible shocks is one where relative poverty is a permanent individual characteristic, and where the probability of moving up (or down, for that matter) the social ladder is exactly zero, both in the short and in the long-run. Of course, if there is aggregate growth the proportion of households that are poor in absolute terms falls over time, and eventually everybody crosses an absolute poverty line: but the individual rank in society remains unchanged.

Figure 1 shows that moving from this extreme representation of reality to a world with incomplete markets generates social mobility. In a world in which households change con-

[22] For instance, suppose that a poor and a rich household switch ranks in the consumption distribution. This will not change the cross-sectional variance of consumption but represents a violation of consumption insurance.

sumption one-for-one in response to permanent income shocks, and smooth transitory shocks by saving and dissaving (the PIH model, obtained by setting $\phi = 1$, $\alpha = 0$, and $\lambda = 0$ in equation 16), we find a 24 percent probability of moving from the first consumption quartile in period t to a higher quartile in period $t + 2$. Since each period the household receives new income shocks, we can generate a consumption distribution also in year $t + 4$, $t + 6$, and so on until $t + 20$ (recall that our panel and transition law for consumption span two years of data). From each distribution we then create consumption quartiles and compute the probability of moving to higher quartiles in period $t + 4$, $t + 6$ and so on conditional on being in the first quartile in period t. This set of calculations traces the lowest line in Figure 1. Since the income process is non-stationary, income shocks compound and the chance of escaping poverty increases over time, up to a long-run value of 43 percent.

A second source of consumption mobility is due to the amount of excess sensitivity of consumption to transitory income shocks that we find in the data. The intermediate line in Figure 1 is obtained using a transition law for consumption with $\phi = 1$, $\lambda = 0.16$ and $\alpha = 0$. Although the line lies above the one estimated for the PIH model, the distance between the two is rather small, reflecting the fact that the estimated λ is well below unity.

Measurement error represents a third source of consumption mobility. The upper line in Figure 1 plots the estimated probability of escaping poverty for the full model ($\phi = 1$, $\lambda = 0.16$ and $\alpha = 0.38$, as in Table 4). The distance between this line and the intermediate line (with $\alpha = 0$) indicates that measurement error adds about 10 percentage points to the probability of leaving the first consumption quartile. Notice also that the probability in $t+2$ is 33 percent, matching the actual value ($0.33 = 1 - p_{11}$ in Table 5) and that measurement error impacts equally short and long-run mobility.

One interesting implication of Figure 1 is that there is an inverse relation between market completeness and social mobility. Societies where individuals insure or smooth a great deal of their idiosyncratic income shocks have the least social mobility. On the other hand, societies with imperfect insurance and credit markets will experience higher social mobility. In principle, these implications of the theories of intertemporal choice could be confronted with cross-country data, but one should bear in mind that the ordering of social mobility depends not only on the availability of insurance and credit markets, but also on the characteristics and persistence of the income shocks.

From a theoretical point of view, it is difficult to compare societies with different degrees of social mobility. Atkinson and Bourguignon (1982) show that mobility structures are irrelevant when the social welfare function is a weighted sum of households' time-separable expected utilities (in Section 3 we indeed rely on such time-separable utilities). In order to assess the welfare implications of mobility one should model explicitly preferences for the fundamentals that affect the social value of mobility, such as social aversion to inequality and the value that the society assigns to the equality of opportunities in attenuating disparities in initial endowments and origins (Benabou and Ok, 2001; Gottschalk and Spolaore, 2002).[23] Although our results have no immediate implications for the welfare analysis of social mobility, it is clear that even a moderate amount of preference for a society that values social mobility would make models where income shocks do not impact consumption undesirable or unsustainable.

7 Conclusions

The implications of the theories of intertemporal consumption choice for consumption mobility are as yet unexplored. In this paper we study transition probabilities for total non-durable consumption using the 1987-95 panel contained in the Bank of Italy Survey of Household Income and Wealth. We summarize the transition matrix of consumption by appropriate mobility indexes and find that there is substantial consumption mobility: in any year, about 60 percent of the households moves up or down in the consumption distribution.

In the remainder of the paper we attempt to understand which model of intertemporal consumption choice is capable of explaining the amount of consumption mobility we observe in the data. From the theoretical point of view, the consumption insurance model provides the clearest implications for consumption mobility. In a model where all idiosyncratic income shocks are insured, the initial cross-sectional distribution of consumption is a sufficient statis-

[23] This leads to different functional forms of social welfare, such as the concave aggregating function of individual time-separable utilities analyzed by Gottschalk and Spolaore (2002).

tic for all future distributions, and therefore, apart from measurement error in consumption, the model predicts zero consumption mobility. On the other hand, the rule-of-thumb model is one where income shocks have the greatest impact on consumption; it therefore generates substantial consumption mobility. Finally, in models with optimizing agents and incomplete markets (such as the permanent income model or models with precautionary saving) households react mainly to permanent income shocks. Thus, the degree of mobility predicted by the model is intermediate between the two other models.

We carefully parametrize an income process to distinguish between transitory and permanent shocks and use the estimated parameters to simulate theoretically the degree of mobility stemming from each of the consumption models examined. We then compare them statistically with the actual amount of mobility estimated in the data. The results reject statistically each of the simple representations of the consumption decision rule, and reveal that households smooth income shocks to a lesser extent than implied by the PIH.

Several criteria suggest that our estimates describe the dynamics of the consumption distribution remarkably well. First, the aggregate mobility index generated by the estimates is almost identical to the empirical mobility index. Second, the estimates are able to match the empirical transition matrix cell by cell. Third, and most importantly, the group-specific estimates by education match the different patterns of consumption mobility we find in the data. Finally, the results are robust with respect to different definitions of consumption (in per capita or per adult equivalent terms), to the presence of measurement error in income and to various other sensitivity checks on sample exclusions and definitions.

There are three important by-products of our analysis. First of all, we produce structural estimates of the sensitivity of consumption to permanent and transitory income shocks that are potentially useful to evaluate fiscal policy experiments that affect the timing of income receipts and, more generally, households' budget constraints. In this respect, we find considerable asymmetric response to transitory income shocks by education groups: a low response in the group with higher education and a relatively high response for households with lower education.

Second, we provide a powerful test of the consumption insurance model. So far these tests have focused on mean and variance restrictions of the distribution of consumption growth. Mean restrictions require consumption growth to be orthogonal, on average, to idiosyncratic income shocks. If shocks are measured with error, however, these tests are biased towards the null hypothesis of full consumption insurance. Variance restrictions require the cross-sectional variance of consumption growth to be constant over time. But the variance might be stationary even if the underlying consumption distribution is shifting. Thus, variance restriction tests too are biased towards the null. Our test is free from these problems, because we look at the entire consumption distribution, not just its mean or variance. On the other hand, the implementation of this test and, more generally, the evaluation of consumption mobility requires genuine panel data, while mean and variance restriction tests can be performed with repeated cross-sectional data.

Finally, the estimates allow us to single out the separate contributions of incomplete markets, excess sensitivity, and measurement error in generating the actual mobility we observe in the data for both the short and the long run. Interestingly, we show that social mobility is inversely related to market completeness, so that societies where individuals insure or smooth a great deal of their idiosyncratic income shocks have the least social mobility. This suggests that one reason why complete markets fail is precisely because societies assign a positive value to social mobility and equality of opportunities, an implication that we plan to explore in future research.

References

[1] Alan, S., and M. Browning (2003), "Estimating Euler equations using simulated residual estimation", CAM Working Paper 3, University of Copenhagen.

[2] Altonji, J. G., and L. M. Segal (1996), "Small-sample bias in GMM estimation of covariance structures", *Journal of Business and Economic Statistics* 14, 353-66.

[3] Anderson, T. W., and L. A. Goodman (1957), "Statistical inference about Markov chains", *Annals of Mathematical Statistics* 28, 89-110.

[4] Atkinson, A. B., and F. Bourguignon (1982), "The comparison of multidimensional distributions of economic status", *Review of Economic Studies* 49, 182-201.

[5] Attanasio, O. P. (2000), "Consumption", in *Handbook of Macroeconomics*, John B. Taylor and Michael Woodford. Amsterdam: Elsevier Science.

[6] Attanasio, O. P., and T. Jappelli (2001),"Intertemporal choice and the cross-sectional variance of marginal utility," *Review of Economics and Statistics* 83, 13-27.

[7] Banerjee, A. V. and A. F. Newman (1991), "Risk-bearing and the theory of income distribution", *Review of Economic Studies* 58, 211-35.

[8] Benabou, R., and E. A. Ok (2001), "Mobility as progressivity: ranking income processes according to equality of opportunity", NBER Working Paper 8431.

[9] Bishop Y.M., S.E. Fienberg, and P.W. Holland (1975), Discrete multivariate analysis: Theory and Practice, Cambridge: MIT Press.

[10] Blundell, R. W., and I. Preston (1998), "Consumption inequality and income uncertainty", *Quarterly Journal of Economics* 113, 603-640.

[11] Blundell, R. W., L. Pistaferri, and I. Preston (2003), "Partial insurance, information, and consumption dynamics", CEPR Discussion Paper 3666.

[12] Brandolini, A. (1999), "The distribution of personal income in post-war Italy: source description, data quality, and the time pattern of income inequality", *Giornale degli Economisti e Annali di Economia* 112, No. 2, 183-240.

[13] Brandolini, A., and L. Cannari (1994), "Methodological Appendix: The Bank of Italy Survey of Household Income and Wealth", in *Saving and the accumulation of wealth. Essays on Italian households and government behavior*, A. Ando, L. Guiso and I. Visco eds. Cambridge: Cambridge University Press.

[14] Browning, M., and T. F. Crossley (2001), "The lifecycle model of consumption and saving," *Journal of Economic Perspectives* 15, No. 3, 3-22.

[15] Campbell, J. Y., and G. N. Mankiw (1989), "Consumption, income, and interest rates: reinterpeting the time-series evidence," in *NBER Macroeconomics Annual 1989*, Olivier J. Blanchard and Stanley Fischer eds., 185-216. Cambridge: MIT Press.

[16] Carroll, C. D. (2001), "Precautionary saving and the marginal propensity to consume out of permanent income," mimeo, John Hopkins University.

[17] Carroll, C. D., and A. A. Samwick (1997), "The nature of precautionary wealth", *Journal of Monetary Economics* 40, 41-72.

[18] Cochrane, J. (1991), "A simple test of consumption insurance", *Journal of Political Economy* 99, 957-976.

[19] Cutler, D. M., and L. F. Katz (1992), "Rising inequality? Changes in the distribution of income and consumption in the 1980s", *American Economic Review* 82, 546-51.

[20] Deaton, A. (1997), *The Analysis of Household Surveys.* Baltimore: The Johns Hopkins University Press.

[21] Deaton, A., and C. Paxson (1994), "Intertemporal choice and inequality", *Journal of Political Economy* 102, 384-94.

[22] Duffie, D. and K. Singleton (1993), "Simulated moments estimation of Markov models of asset prices", *Econometrica* 61, 929-52.

[23] Formby, J. P., W. J. Smith, and B. Zheng (2001), "Mobility measurement, transition matrices, and statistical inference", University of Alabama and University of Colorado at Denver, mimeo.

[24] Friedman, M. (1957), A theory of the consumption function, Princeton: Princeton University Press.

[25] Goffe, W., G. Ferrier and J. Rogers (1994), "Global optimization of statistical functions with simulated annealing", *Journal of Econometrics* 60, 65-99.

[26] Gottschalk, P., and E. Spolaore (2002), "On the evaluation of economic mobility, *Review of Economic Studies* 69, 191-208.

[27] Gourinchas, P.O. and J. Parker (2002), "Consumption over the life cycle", *Econometrica* 70, 47-91.

[28] Hall, R. E., and F. Mishkin (1982), "The sensitivity of consumption to transitory income: evidence from PSID households," *Econometrica* 50, 461-81.

[29] Johnson, D., and S. Shipp (1997), "Trends in inequality using Consumption Expenditures: the US from 1960 to 1993", *Review of Income and Wealth* 43, 133-151.

[30] Kimball, M. S. (1990), "Precautionary saving in the small and in the large", *Econometrica* 58, 53-73.

[31] Lee, B. and B. Ingram(1991), "Simulation estimation of time series model", *Journal of Econometrics* 47, 197-205.

[32] Lillard, L., and R. Willis (1978), "Dynamic aspects of earning mobility", *Econometrica* 46, 985-1012.

[33] Mace, B. (1991), "Full insurance in the presence of aggregate uncertainty", *Journal of Political Economy* 99, 928-56.

[34] Mas-Colell, A., M. Whinston, and J. Greene (1995), *Microeconomic theory.* Oxford: Oxford University Press.

[35] McFadden D. (1989), "A method of simulated moments for estimation of discrete response models without numerical integration", *Econometrica* 57, 995-1026.

[36] Neyman, J. (1949), "Contributions to the theory of the χ^2 test", in *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability.* Berkeley: University of California Press.

[37] Parker, J. A. (1999), "The reaction of household consumption to predictable changes in social security taxes", *American Economic Review* 89, 959-73.

[38] Shorrocks, A. F. (1978), "The measurement of mobility", *Econometrica* 46, 1013-24.

[39] Silverman, B.W. (1986), *Density estimation for statistics and data analysis*, Chapman and Hall.

[40] Souleles, N. S. (1999), "The response of household consumption to income tax refunds", *American Economic Review* 89, 947-58.

[41] Townsend, R. (1994), "Risk and insurance in village India", *Econometrica* 62, 539-91.

A Proof of proposition 1

Recall the three distribution laws for log consumption:

$$\begin{aligned} \ln c_{h,t+1} &= \ln c_{h,t} + \zeta_{h,t+1} + \frac{r}{1+r}\varepsilon_{h,t+1} \\ \ln c_{h,t+1} &= \ln c_{h,t} + \zeta_{h,t+1} + \varepsilon_{h,t+1} - \varepsilon_{h,t} \\ \ln c_{h,t+1} &= \ln c_{h,t} \end{aligned}$$

respectively in three cases of the PIH, the rule-of-thumb model, and consumption insurance. Recall also that ζ is the permanent shock and ε the transitory shock. Without loss of generality, we set aggregate consumption growth to zero.

Divide the distribution in quantiles. Denote with q_{j-1} and q_j two successive quantiles of the distribution ($q_{j-1} < q_j$), and assume that:

$$\begin{pmatrix} \zeta_{h,t+1} \\ \varepsilon_{h,t+1} \\ \varepsilon_{h,t} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\zeta^2 & 0 & 0 \\ 0 & \sigma_\varepsilon^2 & 0 \\ 0 & 0 & \sigma_\varepsilon^2 \end{pmatrix}\right]$$

Start from the consumption insurance model and note that:

$$\begin{aligned} & \Pr\left(q_{j-1} < \ln c_{h,t+1} < q_j \,|q_{j-1} < \ln c_{h,t} < q_j, \ln c_{h,t+1} = \ln c_{h,t}\right) \\ = \; & \Pr\left(q_{j-1} < \ln c_{h,t} < q_j \,|q_{j-1} < \ln c_{h,t} < q_j\right) = 1 \end{aligned}$$

In the rule-of-thumb model:

$$\begin{aligned} & \Pr\left(q_{j-1} < \ln c_{h,t+1} < q_j \left|q_{j-1} < \ln c_{h,t} < q_j, \ln c_{h,t+1} = \ln c_{h,t} + \zeta_{h,t+1} + \varepsilon_{h,t+1} - \varepsilon_{h,t}\right.\right) \\ & = \Pr\left(q_{j-1} < \ln c_{h,t} + \zeta_{h,t+1} + \varepsilon_{h,t+1} - \varepsilon_{h,t} < q_j \,|q_{j-1} < \ln c_{h,t} < q_j\right) \\ & = \Pr\left(q_{j-1} - \ln c_{h,t} < \zeta_{h,t+1} + \varepsilon_{h,t+1} - \varepsilon_{h,t} < q_j - \ln c_{h,t} \,|q_{j-1} - \ln c_{h,t} < 0 < q_j - \ln c_{h,t}\right) \\ & = \Pr\left(\tfrac{q_{j-1} - \ln c_{h,t}}{\sigma_K} < \tfrac{\zeta_{h,t+1} + \varepsilon_{h,t+1} - \varepsilon_{h,t}}{\sigma_K} < \tfrac{q_j - \ln c_{h,t}}{\sigma_K} \,|q_{j-1} - \ln c_{h,t} < 0 < q_j - \ln c_{h,t}\right) \\ & = \Phi\left(\tfrac{q_j - \ln c_{h,t}}{\sigma_K}\right) - \Phi\left(\tfrac{q_{j-1} - \ln c_{h,t}}{\sigma_K}\right) > 0 \end{aligned}$$

where $\sigma_K = \sqrt{\sigma_\zeta^2 + 2\sigma_\varepsilon^2}$, and $\Phi(.)$ is the c.d.f. of the $N(0,1)$ distribution. The last inequality holds because $q_j - \ln c_{h,t} > 0$, $q_{j-1} - \ln c_{h,t} < 0$, and $\sigma_K > 0$.

In the PIH:

$$\begin{aligned} & \Pr\left(q_{j-1} < \ln c_{h,t+1} < q_j \left|q_{j-1} < \ln c_{h,t} < q_j, \ln c_{h,t+1} = \ln c_{h,t} + \zeta_{h,t+1} + \tfrac{r}{1+r}\varepsilon_{h,t+1}\right.\right) \\ & = \Pr\left(q_{j-1} < \ln c_{h,t} + \zeta_{h,t+1} + \tfrac{r}{1+r}\varepsilon_{h,t+1} < q_j \,|q_{j-1} < \ln c_{h,t} < q_j\right) \\ & = \Pr\left(q_{j-1} - \ln c_{h,t} < \zeta_{h,t+1} + \tfrac{r}{1+r}\varepsilon_{h,t+1} < q_j - \ln c_{h,t} \,|q_{j-1} - \ln c_{h,t} < 0 < q_j - \ln c_{h,t}\right) \\ & = \Pr\left(\tfrac{q_{j-1} - \ln c_{h,t}}{\sigma_{PIH}} < \tfrac{\zeta_{h,t+1} + \frac{r}{1+r}\varepsilon_{h,t+1}}{\sigma_{PIH}} < \tfrac{q_j - \ln c_{h,t}}{\sigma_{PIH}} \,|q_{j-1} - \ln c_{h,t} < 0 < q_j - \ln c_{h,t}\right) \\ & = \Pr\left(\tfrac{q_{j-1} - \ln c_{h,t}}{\sigma_{PIH}} < \tfrac{\zeta_{h,t+1} + \frac{r}{1+r}\varepsilon_{h,t+1}}{\sigma_{PIH}} < \tfrac{q_j - \ln c_{h,t}}{\sigma_{PIH}} \,|q_{j-1} - \ln c_{h,t} < 0 < q_j - \ln c_{h,t}\right) \\ & = \Phi\left(\tfrac{q_j - \ln c_{h,t}}{\sigma_{PIH}}\right) - \Phi\left(\tfrac{q_{j-1} - \ln c_{h,t}}{\sigma_{PIH}}\right) > 0 \end{aligned}$$

where $\sigma_{PIH} = \sqrt{\sigma_\zeta^2 + \left(\frac{r}{1+r}\right)^2 \sigma_\varepsilon^2}$. The last inequality holds because $q_j - \ln c_{h,t} > 0$, $q_{j-1} - \ln c_{h,t} < 0$, and $\sigma_{PIH} > 0$. Notice now that $\sigma_K > \sigma_{PIH}$, so that:

$$\frac{q_{j-1} - \ln c_{h,t}}{\sigma_{PIH}} < \frac{q_{j-1} - \ln c_{h,t}}{\sigma_K} < 0 < \frac{q_j - \ln c_{h,t}}{\sigma_K} < \frac{q_j - \ln c_{h,t}}{\sigma_{PIH}}$$

and finally:

$$\Phi\left(\tfrac{q_j - \ln c_{h,t}}{\sigma_{PIH}}\right) - \Phi\left(\tfrac{q_{j-1} - \ln c_{h,t}}{\sigma_{PIH}}\right) > \Phi\left(\tfrac{q_j - \ln c_{h,t}}{\sigma_K}\right) - \Phi\left(\tfrac{q_{j-1} - \ln c_{h,t}}{\sigma_K}\right)$$

The last inequality proves that the probability of remaining in the same quantile of the consumption distribution is greater under the PIH than under the rule-of-thumb model, i.e. that mobility is higher in the latter case. This completes the proof.

B The simulated minimum χ^2 estimator

Let $\mathbf{P}(\theta)$ the $q \times q$ transition matrix with typical element $p_{ij}(\theta)$, where θ is a vector of k unknown parameters:

$$\mathbf{P}(\theta) = \begin{bmatrix} p_{11}(\theta) & p_{12}(\theta) & ... & p_{1q}(\theta) \\ p_{21}(\theta) & p_{22}(\theta) & ... & p_{2q}(\theta) \\ ... & ... & ... & ... \\ p_{q1}(\theta) & p_{q2}(\theta) & ... & p_{qq}(\theta) \end{bmatrix}$$

Conformably with $\mathbf{P}(\theta)$ let $\widehat{\mathbf{P}}$ the $q \times q$ empirical transition matrix with typical element $\widehat{p}_{ij}$:

$$\widehat{\mathbf{P}} = \begin{bmatrix} \widehat{p}_{11} & \widehat{p}_{12} & ... & \widehat{p}_{1q} \\ \widehat{p}_{21} & \widehat{p}_{22} & ... & \widehat{p}_{2q} \\ ... & ... & ... & ... \\ \widehat{p}_{q1} & \widehat{p}_{q2} & ... & \widehat{p}_{qq} \end{bmatrix}$$

The transition matrices $\mathbf{P}(\theta)$ and $\widehat{\mathbf{P}}$ are subject to the restrictions $\sum_{j=1}^{q} p_{ij}(\theta) = 1$ and $\sum_{j=1}^{q} \widehat{p}_{ij} = 1$ $(i = 1...q)$, respectively. This creates a singularity problem similar to the one in the estimation of a full demand system. To avoid this problem, we drop one column (say, the q-th column) from both $\mathbf{P}(\theta)$ and $\widehat{\mathbf{P}}$.

Let $\mathbf{p}(\theta)$ the $q(q-1) \times 1$ vector of true transition probabilities and conformably with $\mathbf{p}(\theta)$ let $\widehat{\mathbf{p}}$ the $q(q-1) \times 1$ vector of estimated transition probabilities. The distance between the empirical and true transition probabilities is $\mathbf{d}(\theta) = \widehat{\mathbf{p}} - \mathbf{p}(\theta)$, whose covariance matrix $\mathbf{\Omega}(\theta)$ is block diagonal with generic block:[24]

$$\mathbf{\Omega}_i(\theta) = \begin{bmatrix} \frac{p_{i1}(\theta)(1-p_{i1}(\theta))}{n_i} & -\frac{p_{i1}(\theta)p_{i2}(\theta)}{n_i} & ... & -\frac{p_{i1}(\theta)p_{iq-1}(\theta)}{n_i} \\ & \frac{p_{i2}(\theta)(1-p_{i2}(\theta))}{n_i} & ... & -\frac{p_{i2}(\theta)p_{iq-1}(\theta)}{n_i} \\ & & ... & ... \\ & & & \frac{p_{iq-1}(\theta)(1-p_{iq-1}(\theta))}{n_i} \end{bmatrix}$$

[24] We neglect the extra randomness induced by the fact that the class boundaries are pre-estimated.

for $i = 1...q$ (we assume $n_i = n/q$ is an integer for simplicity), so that:

$$\mathbf{\Omega}(\theta) = \begin{bmatrix} \mathbf{\Omega}_1(\theta) & \mathbf{0} & ... & \mathbf{0} \\ & \mathbf{\Omega}_2(\theta) & ... & \mathbf{0} \\ & & ... & ... \\ & & & \mathbf{\Omega}_q(\theta) \end{bmatrix}$$

From Chamberlain (1982), the minimum χ^2 method solves the problem:

$$\min_{\theta} \ \mathbf{d}(\theta)' \mathbf{W} \mathbf{d}(\theta)$$

where $\mathbf{W}$ is a weighting matrix. Call $\widehat{\theta}$ the minimum χ^2 estimate of θ. Chamberlain (1982) and others show that $\widehat{\theta}$ is consistent, asymptotically normal with covariance matrix:

$$\text{var}\left(\widehat{\theta}\right) = \left(\mathbf{G}\left(\widehat{\theta}\right)' \mathbf{W} \mathbf{G}\left(\widehat{\theta}\right)\right)^{-1} \mathbf{G}\left(\widehat{\theta}\right)' \mathbf{W} \mathbf{\Omega}(\boldsymbol{\theta}) \mathbf{W} \mathbf{G}\left(\widehat{\theta}\right) \left(\mathbf{G}\left(\widehat{\theta}\right)' \mathbf{W} \mathbf{G}\left(\widehat{\theta}\right)\right)^{-1}$$

where $\mathbf{G}\left(\widehat{\theta}\right) = \frac{\partial \mathbf{d}\left(\widehat{\theta}\right)}{\partial \theta'}$ is the gradient matrix. It is a well known result that the optimal weighting matrix (in the efficiency sense) is $\mathbf{\Omega}(\theta)^{-1}$. In this case:

$$\text{var}\left(\widehat{\theta}\right) = \left(\mathbf{G}\left(\widehat{\theta}\right)' \mathbf{\Omega}(\boldsymbol{\theta})^{-1} \mathbf{G}\left(\widehat{\theta}\right)\right)^{-1}$$

In our case $\mathbf{p}(\theta)$ has no closed form, so we replace it with an approximation based on simulations, as in the simulated method of moments (McFadden, 1989; Duffie and Singleton, 1991). Recall that the generic $p_{ij}(\theta)$ is:

$$p_{ij}(\theta) = \Pr\left(\ln c^*_{h,t} \in i \middle| \ln c^*_{h,t-2} \in j, \theta\right)$$

e.g., the probability of making a transition to class i from class j conditioning on being in class j. The transition law for consumption is determined by (16), reproduced here:

$$\begin{aligned} \ln c^*_{h,t} &= \ln c^*_{h,t-2} + \phi\left(\frac{\lambda + r}{1+r}\varepsilon_{h,t} + \frac{(1-\lambda) r}{1+r}\varepsilon_{h,t-1} - \lambda \varepsilon_{h,t-2} + \zeta_{h,t} + \zeta_{h,t-1}\right) \\ &\quad + \alpha\left(v_{h,t} - v_{h,t-2}\right) \end{aligned}$$

where in our case $\theta = (\phi \ \alpha \ \lambda)'$.[25] For simulation purposes, we assume $\varepsilon_{h,\tau} \sim N\left(0, \sigma^2_\varepsilon\right)$, $\zeta_{h,\tau} \sim N\left(0, \sigma^2_\zeta\right)$, $v_{h,\tau} \sim N\left(0, \sigma^2_{\ln c^*_{h,\tau}}\right)$ $\forall \tau$.

By construction, the normality of the income shocks and of measurement error generates a symmetric transition matrix for consumption. This feature of the simulations is consistent with the symmetry of the empirical matrix documented in Table 2. Our results do not depend on the normality assumption. We choose normality for simplicity, but note that any symmetric distribution would work as well, because it would imply a symmetric transition matrix.

[25] We neglect the problems associated with the fact that r is given, and that $\sigma^2_{\ln c^*_{h,t-2}}$, $\sigma^2_{\ln c^*_{h,t}}$, σ^2_ε and σ^2_ζ are pre-estimated.

Define $\mathbf{u}_h = \left(\varepsilon_{h,t}\ \varepsilon_{h,t-1}\ \varepsilon_{h,t-2}\ \zeta_{h,t}\ \zeta_{h,t-1}\ \upsilon_{h,t}\ \upsilon_{h,t-2}\right)'$ the vector of disturbances. For each household h, we draw S independent realizations of $\mathbf{u}_h$, and store the HS realizations (H being the number of households).[26] It is necessary to keep these basic drawings of $\mathbf{u}_h^s$ fixed when θ changes, in order to have good numerical and statistical properties of the estimators based on the simulations.

Conditioning on the measured (not simulated) $\ln c^*_{h,t-2}$, the simulated $\mathbf{u}_h^s$, and a choice for θ, one obtains

$$\begin{aligned}\ln c^{*s}_{h,t} &= \ln c^*_{h,t-2} + \phi\left(\frac{\lambda+r}{1+r}\varepsilon^s_{h,t} + \frac{(1-\lambda)\,r}{1+r}\varepsilon^s_{h,t-1} - \lambda\varepsilon^s_{h,t-2} + \zeta^s_{h,t} + \zeta^s_{h,t-1}\right)\\ &\quad +\alpha\left(\widetilde{\upsilon}^s_{h,t} - \widetilde{\upsilon}^s_{h,t-2}\right) \qquad \text{(A1)}\end{aligned}$$

This allows computation of $p^s_{ij}(\theta)$. One can then define $\overline{p}_{ij}(\theta) = S^{-1}\sum_{s=1}^{S} p^s_{ij}(\theta)$ as the approximation of $p_{ij}(\theta)$ obtained by means of simulations.

Call the simulated distance $\overline{\mathbf{d}}(\theta) = \widehat{\mathbf{p}} - \overline{\mathbf{p}}(\theta)$ where $\overline{\mathbf{p}}(\theta)$ is the vector of simulated transition probabilities with generic element $\overline{p}_{ij}(\theta)$. Note that the covariance matrix of $\overline{\mathbf{d}}(\theta)$, $\overline{\mathbf{\Omega}}(\theta) \overset{a.s.}{\to} \left(1+\frac{1}{S}\right)\mathbf{\Omega}(\theta)$ where $\left(1+\frac{1}{S}\right)$ is an inflating factor of the variance of the true distance vector induced by the additional randomness of the simulations. With a large enough number of simulations, however, this factor plays little weight in practice.

The choice of θ minimizes the simulated minimum χ^2 criterion

$$\min_{\theta}\ \overline{\mathbf{d}}(\theta)'\,\overline{\mathbf{\Omega}}(\theta)^{-1}\,\overline{\mathbf{d}}(\theta)$$

Call $\widetilde{\theta}$ the resulting solution. Then, the results in Lee and Ingram(1991), McFadden(1989), and Duffie and Singleton (1993) imply that $\widetilde{\theta}$ is consistent, asymptotically normal with covariance matrix:

$$\text{var}\left(\widetilde{\theta}\right) = \left(1+\frac{1}{S}\right)\left[\mathbf{G}\left(\widetilde{\theta}\right)'\,\mathbf{\Omega}\left(\widetilde{\theta}\right)^{-1}\mathbf{G}\left(\widetilde{\theta}\right)\right]^{-1}$$

Goodness of fit can be assessed using:

$$m = \mathbf{d}\left(\widetilde{\theta}\right)'\,\mathbf{\Omega}\left(\widetilde{\theta}\right)^{-1}\mathbf{d}\left(\widetilde{\theta}\right) \sim \chi^2_{q(q-1)-k}$$

The algorithm that we implement is thus the following:

1. Draw $\mathbf{u}_h^s$ $(h = 1...H,\ s = 1...S)$.
2. Choose a starting value for θ, say θ_0.
3. Compute $\ln c^*_{h,t}$ using (A1), $\overline{p}_{ij}(\theta_0)$, and $\overline{d}_{ij}(\theta_0) = \widehat{p}_{ij} - \overline{p}_{ij}(\theta_0)$ $(i = 1...q,\ j = 1...q-1)$.
4. Compute $\overline{\mathbf{d}}(\theta_0)'\,\overline{\mathbf{\Omega}}(\theta_0)^{-1}\,\overline{\mathbf{d}}(\theta_0)$.
5. Update the value of θ.
6. Repeat steps 3-5 until a pre-specified convergence criterion is met. Eventually this provides the required simulated minimum χ^2 estimate $\widetilde{\theta}$ of θ.

We update the value of θ using the simulated annealing method of Goffe, Ferrier and Rogers (1994).[27] This is a derivative-free minimization method that escapes local minima. Starting from

[26] In each year we choose a sample size identical to the number of actual sample transitions (for instance, it is 2,982 in 1991-93 and 3,211 in 1993-95).

[27] We use the Gauss code on simulated annealing written by E.G. Tsionas and available at http://www.american.edu/academic.depts/cas/econ/gaussres/optimize/optimize.htm.

an initial value, the algorithm takes a step and evaluates the function. Downhill steps are always accepted, while uphill steps are accepted probabilistically according to the Metropolis criterion. As the algorithm proceeds, the length of the step declines until the χ^2 reaches the global minimum.

C Test equivalence

Here we prove the statement in Section 2 that the χ^2 goodness of fit criterion (5):

$$\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{\left(\widehat{p}_{ij} - p_{ij}\left(\theta\right)\right)^2}{p_{ij}\left(\theta\right)} \tag{A2}$$

is equivalent to $\left(\widehat{\mathbf{p}} - \mathbf{p}\left(\theta\right)\right) \mathbf{\Omega}\left(\theta\right)^{-1} \left(\widehat{\mathbf{p}} - \mathbf{p}\left(\theta\right)\right)'$.

Note first that (A2) is the sum of q independent χ^2 distributions of the form $m_i = \sum_{j=1}^{q} n_i \frac{\left(\widehat{p}_{ij} - p_{ij}(\theta)\right)^2}{p_{ij}}$. The sum of q independent χ^2 distributions is also a χ^2 distribution with degrees of freedom equal to the sum of the degrees of freedom of the χ^2 distributions that are summed.

Notice that the theoretical and empirical transition probabilities are subject to the restrictions $\sum_{j=1}^{q} p_{ij}\left(\theta\right) = 1$ and $\sum_{j=1}^{q} \widehat{p}_{ij} = 1$ $(i = 1...q)$, respectively. Thus $m_i = \sum_{j=1}^{q} n_i \frac{\left(\widehat{p}_{ij} - p_{ij}(\theta)\right)^2}{p_{ij}(\theta)}$ can be rewritten as:

$$\begin{aligned} m_i &= \sum_{j=1}^{q-1} \frac{n_i}{p_{ij}\left(\theta\right)} \left(\widehat{p}_{ij} - p_{ij}\left(\theta\right)\right)^2 + \frac{n_i}{p_{iq}\left(\theta\right)} \left[\sum_{j=1}^{q-1} \left(\widehat{p}_{ij} - p_{ij}\left(\theta\right)\right)\right]^2 \\ &= \sum_{j=1}^{q-1} \frac{n_i}{p_{ij}\left(\theta\right)} \left(\widehat{p}_{ij} - p_{ij}\left(\theta\right)\right)^2 + \frac{n_i}{p_{iq}\left(\theta\right)} \sum_{j=1}^{q-1} \left(\widehat{p}_{ij} - p_{ij}\left(\theta\right)\right) \sum_{j=1}^{q-1} \left(\widehat{p}_{ij} - p_{ij}\left(\theta\right)\right) \end{aligned}$$

or more compactly:

$$\begin{aligned} m_i &= \mathbf{d}_i\left(\theta\right)' \mathbf{A}_i\left(\theta\right) \mathbf{d}_i\left(\theta\right) + \mathbf{d}_i\left(\theta\right)' \mathbf{B}_i\left(\theta\right) \mathbf{d}_i\left(\theta\right) \\ &= \mathbf{d}_i\left(\theta\right)' \mathbf{\Lambda}_i\left(\theta\right) \mathbf{d}_i\left(\theta\right) \end{aligned}$$

where $\mathbf{d}_i\left(\theta\right) = \widehat{\mathbf{p}}_i - \mathbf{p}_i\left(\theta\right)$ is the distance between empirical and true transition probabilities in row i of the transition matrix (excluding the q-th column), and $\mathbf{\Lambda}_i\left(\theta\right) = \mathbf{A}_i\left(\theta\right) + \mathbf{B}_i\left(\theta\right)$, with:

$$\mathbf{A}_i\left(\theta\right) = \begin{bmatrix} \frac{n_i}{p_{11}(\theta)} & 0 & ... & 0 \\ 0 & \frac{n_i}{p_{12}(\theta)} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \frac{n_i}{p_{iq-1}(\theta)} \end{bmatrix}$$

and $\mathbf{B}_i\left(\theta\right) = \frac{n_i}{p_{iq}(\theta)}\mathbf{ii}'$, where $\mathbf{i}$ is a $(q-1) \times 1$ vector of ones, so that $\mathbf{B}_i\left(\theta\right)$ is a matrix that contains $\frac{n_i}{p_{iq}(\theta)}$ everywhere. It's easy to prove that $\mathbf{\Lambda}_i\left(\theta\right) = \mathbf{\Omega}_i\left(\theta\right)^{-1}$ defined in Appendix B. Since asymptotically $\mathbf{d}_i\left(\theta\right) \sim N\left(\mathbf{0}, \mathbf{\Omega}_i\left(\theta\right)\right)$, it follows that:

$$m_i = \sum_{j=1}^{q} n_i \frac{\left(\widehat{p}_{ij} - p_{ij}\left(\theta\right)\right)^2}{p_{ij}\left(\theta\right)} = \mathbf{d}_i\left(\theta\right)' \mathbf{\Omega}_i\left(\theta\right)^{-1} \mathbf{d}_i\left(\theta\right)$$

is distributed χ^2 with $(q-1)$ degrees of freedom. Moreover:

$$\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{\left(\widehat{p}_{ij} - p_{ij}(\theta)\right)^2}{p_{ij}(\theta)} = \sum_{i=1}^{q} m_i = \sum_{i=1}^{q} \mathbf{d}_i(\theta)' \, \mathbf{\Omega}_i(\theta)^{-1} \, \mathbf{d}_i(\theta) = \mathbf{d}(\theta)' \, \mathbf{\Omega}(\theta)^{-1} \, \mathbf{d}(\theta)$$

is distributed χ^2 with $q(q-1)$ degrees of freedom. This is exactly the function that we minimize in the simulated minimum χ^2 application. This proves the equivalence between $\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{\left(\widehat{p}_{ij}-p_{ij}(\theta)\right)^2}{p_{ij}(\theta)}$ and $\mathbf{d}(\theta)' \, \mathbf{\Omega}(\theta)^{-1} \, \mathbf{d}(\theta)$.

An alternative to the minimum χ^2 criterion $\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{\left(\widehat{p}_{ij}-p_{ij}(\theta)\right)^2}{p_{ij}(\theta)}$ is to use the modified minimum χ^2 criterion $\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{\left(\widehat{p}_{ij}-p_{ij}(\theta)\right)^2}{\widehat{p}_{ij}}$. Following the same steps above, one can show that:

$$\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{\left(\widehat{p}_{ij} - p_{ij}(\theta)\right)^2}{\widehat{p}_{ij}} = \mathbf{d}(\theta)' \, \widehat{\mathbf{\Omega}}^{-1} \mathbf{d}(\theta)$$

where $\widehat{\mathbf{\Omega}}$ is a block-diagonal matrix with generic block:

$$\widehat{\mathbf{\Omega}}_i = \begin{bmatrix} \frac{\widehat{p}_{i1}\left(1-\widehat{p}_{i1}\right)}{n_i} & -\frac{\widehat{p}_{i1}\widehat{p}_{i2}}{n_i} & \cdots & -\frac{\widehat{p}_{i1}\widehat{p}_{iq-1}}{n_i} \\ & \frac{\widehat{p}_{i2}\left(1-\widehat{p}_{i2}\right)}{n_i} & \cdots & -\frac{\widehat{p}_{i2}\widehat{p}_{iq-1}}{n_i} \\ & & \cdots & \cdots \\ & & & \frac{\widehat{p}_{iq-1}\left(1-\widehat{p}_{iq-1}\right)}{n_i} \end{bmatrix}$$

Since $\widehat{p}_{ij}$ is a consistent estimate of $p_{ij}(\theta)$, $\widehat{\mathbf{\Omega}} \overset{a.s.}{\to} \mathbf{\Omega}(\theta)$. In the estimation, we use the modified simulated minimum χ^2 criterion, i.e. use $\widehat{\mathbf{\Omega}}$ (based on the empirical transition probabilities) as an estimate of $\mathbf{\Omega}(\theta)$.

Table 1
Descriptive Statistics

Cross-sectional means and variances are computed using sample weights. The variables c_t and y_t denote household non-durable consumption and disposable income, respectively. Demographic characteristics refer to the household head.

	1987	1989	1991	1993	1995	All years
$\ln c_t$	9.90	10.08	10.02	10.01	10.00	10.02
$var\,(\ln c_t)$	0.26	0.26	0.29	0.29	0.27	0.28
Gini coefficient of c_t	0.28	0.27	0.29	0.29	0.27	0.28
$\ln y_t$	10.25	10.40	10.36	10.27	10.27	10.32
$var\,(\ln y_t)$	0.39	0.37	0.37	0.57	0.47	0.45
Gini coefficient of y_t	0.35	0.32	0.32	0.36	0.35	0.34
South	0.41	0.37	0.34	0.36	0.39	0.37
North	0.43	0.46	0.48	0.47	0.43	0.46
Family size	3.15	3.12	3.04	3.07	3.01	3.07
Self-employed	0.20	0.17	0.17	0.16	0.15	0.16
Years of schooling	7.38	7.97	8.19	8.03	8.10	8.03
Less well educated	0.78	0.72	0.70	0.72	0.70	0.72
More educated	0.22	0.28	0.30	0.28	0.30	0.28
Age	52.00	52.52	52.78	53.05	55.03	53.22
Born $\leq$ 1940	0.60	0.58	0.54	0.49	0.49	0.53
Born $>$1940	0.40	0.42	0.46	0.51	0.51	0.47
Income recipients	1.63	1.72	1.72	1.74	1.78	1.73
Number of obs.	1,097	2,717	4,036	4,006	3,211	15,067

Table 2
The Transition Matrix of Consumption

The table reports consumption transitions from period $t-2$ to period t. The generic element of this table is $\widehat{p}_{ij}$, the estimated probability of moving from quartile i in period $t-2$ to quartile j in period t. Define n_{ij} as the number of households that move from quartile i in period $t-2$ to quartile j in period t and $n_i = \sum_i n_{ij}$ as the total number of observations in each row i of the transition matrix. The maximum likelihood estimator of the first-order Markov transition probabilities is then: $\widehat{p}_{ij} = \frac{n_{ij}}{n_i}$.

1987-89

	1989 quartile			
1987 quartile	1^{st}	2^{nd}	3^{rd}	4^{th}
1^{st}	0.71	0.20	0.07	0.02
2^{nd}	0.23	0.42	0.27	0.08
3^{rd}	0.08	0.29	0.40	0.23
4^{th}	0.03	0.09	0.29	0.60

1989-91

	1991 quartile			
1989 quartile	1^{st}	2^{nd}	3^{rd}	4^{th}
1^{st}	0.66	0.25	0.07	0.01
2^{nd}	0.25	0.41	0.27	0.06
3^{rd}	0.10	0.27	0.41	0.25
4^{th}	0.01	0.07	0.25	0.68

1991-93

	1993 quartile			
1991 quartile	1^{st}	2^{nd}	3^{rd}	4^{th}
1^{st}	0.63	0.26	0.08	0.02
2^{nd}	0.23	0.38	0.29	0.09
3^{rd}	0.11	0.28	0.37	0.25
4^{th}	0.04	0.10	0.26	0.60

1993-95

	1995 quartile			
1993 quartile	1^{st}	2^{nd}	3^{rd}	4^{th}
1^{st}	0.68	0.25	0.07	0.01
2^{nd}	0.24	0.43	0.26	0.07
3^{rd}	0.07	0.27	0.44	0.23
4^{th}	0.02	0.06	0.23	0.69

Table 3
Empirical Mobility Index

The table reports the Shorrocks mobility index, the associated standard error and the number of transitions for separate sample periods. The index is calculated as: $S\left(\widehat{\mathbf{P}}\right) = \frac{q - trace(\widehat{\mathbf{P}})}{q-1}$, where $q = 4$, $trace(\widehat{\mathbf{P}}) = \sum_i \widehat{p}_{ii}$, and $\widehat{p}_{ii}$ is the estimated probability of being in quartile i at both $t-2$ and t. The standard error is computed as: $\sqrt{\frac{1}{(q-1)^2}\sum_i \frac{\widehat{p}_{ii}\left(1-\widehat{p}_{ii}\right)}{n_i}}$.

Sample period	$S\left(\widehat{\mathbf{P}}\right)$	$s.e.\left(S\left(\widehat{\mathbf{P}}\right)\right)$	Number of transitions
1987-1989	0.6269	0.0194	1,097
1989-1991	0.6274	0.0147	1,914
1991-1993	0.6706	0.0118	2,982
1993-1995	0.5905	0.0113	3,211

Table 4
Parameter Estimates, Mobility Index and χ^2 Statistics

The table reports simulated minimum χ^2 estimates of the parameters ϕ, λ and α (asymptotic standard errors in parenthesis), the simulated and empirical Shorrocks mobility index and the χ^2 goodness of fit statistic (p-value of the test in square brackets). The empirical Shorrocks mobility index refers to 1993-95 transitions. In column (2) we imposes the (acceptable) restriction that $\phi = 1$.

	(1)	(2)
ϕ	0.9875	1.0000
	(0.0230)	
λ	0.1586	0.1622
	(0.0377)	(0.0248)
α	0.3875	0.3822
	(0.0186)	(0.0116)
Simulated mobility index	0.6019	0.6013
Empirical mobility index	0.5937	
χ^2 goodness of fit statistic	15.21	15.22
	[0.0853]	[0.1241]

Table 5
Simulated and Empirical Transition Matrix of Consumption

The table reports the simulated consumption transitions between 1993 and 1995 and, in parenthesis, the empirical consumption transitions. The simulated transitions probabilities are obtained from the estimates reported in column (2) of Table 4.

	1995 quartile			
1993 quartile	1^{st}	2^{nd}	3^{rd}	4^{th}
1^{st}	0.6748	0.2515	0.0677	0.0061
	(0.6700)	(0.2528)	(0.0660)	(0.0112)
2^{nd}	0.2513	0.4111	0.2748	0.0628
	(0.2416)	(0.4259)	(0.2665)	(0.0660)
3^{rd}	0.0675	0.2764	0.4175	0.2386
	(0.0660)	(0.2653)	(0.4346)	(0.2341)
4^{th}	0.0061	0.0613	0.2401	0.6926
	(0.0237)	(0.0549)	(0.2332)	(0.6883)

Table 6
Simulated and Empirical Mobility for Different Education Groups

The first panel reports summary statistics for two education groups: the variance of the income shocks, the empirical and simulated Shorrocks index, the estimates of α and λ (asymptotic standard errors in parenthesis), the associated χ^2 goodness of fit statistic (p-value of the test in square brackets), and the χ^2 statistic of the test that the parameters of the two education groups are the same (p-value in square brackets). The education groups are defined as compulsory schooling or less, and high school or college. The other two panels report the simulated transition probabilities and (in parenthesis) the empirical transition probabilities for the two education groups.

	Low education	High education
Variance of permanent shock	0.0296	0.0198
Variance of transitory shock	0.0754	0.0895
Empirical mobility index	0.6166	0.5540
λ	0.3991 (0.0459)	0.0889 (0.0470)
α	0.3814 (0.0120)	0.2846 (0.0125)
Simulated mobility index	0.6386	0.5461
χ^2 goodness of fit statistic	23.03 [0.0106]	20.85 [0.0222]
χ^2 test of parameter equality	55.85 [7.4e−013]	

Low education

	1995 quartile			
1993 quartile	1^{st}	2^{nd}	3^{rd}	4^{th}
1^{st}	0.6432 (0.6449)	0.2610 (0.2633)	0.0845 (0.0830)	0.0113 (0.0088)
2^{nd}	0.2413 (0.2240)	0.3863 (0.4208)	0.2848 (0.2532)	0.0877 (0.1020)
3^{rd}	0.0733 (0.0780)	0.2654 (0.2524)	0.3913 (0.4421)	0.2700 (0.2277)
4^{th}	0.0099 (0.0327)	0.0770 (0.0750)	0.2498 (0.2500)	0.6632 (0.6423)

High education

	1995 quartile			
1993 quartile	1^{st}	2^{nd}	3^{rd}	4^{th}
1^{st}	0.7037 (0.7016)	0.2485 (0.2258)	0.0457 (0.0605)	0.0022 (0.0121)
2^{nd}	0.2472 (0.2308)	0.4463 (0.4656)	0.2686 (0.2389)	0.0379 (0.0648)
3^{rd}	0.0464 (0.0242)	0.2681 (0.2460)	0.4661 (0.4839)	0.2194 (0.2460)
4^{th}	0.0015 (0.0163)	0.0373 (0.0569)	0.2156 (0.2398)	0.7456 (0.6870)

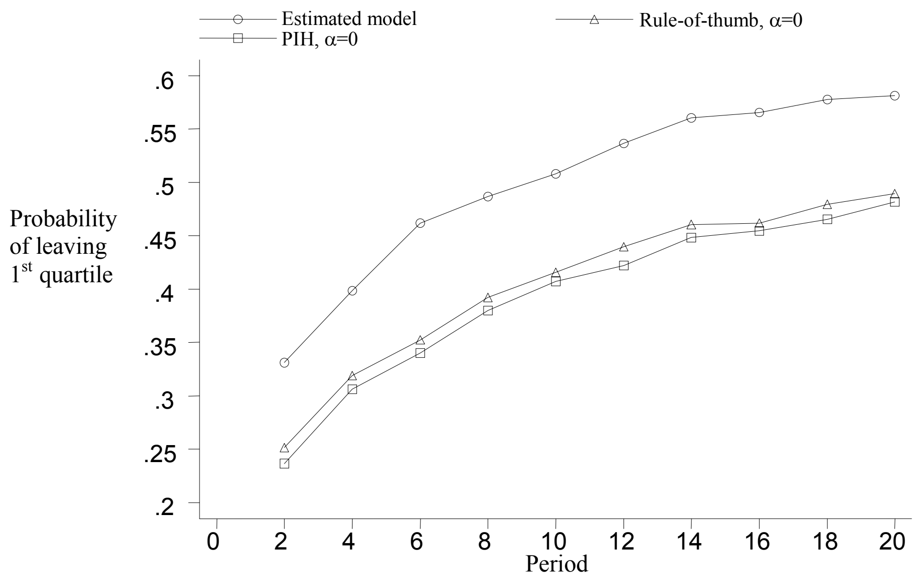

Figure 1: The probability of escaping poverty.