



Estimating Intertemporal Allocation Parameters using Simulated Residual Estimation

Alan, Sule; Browning, Martin

Publication date:
2003

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Alan, S., & Browning, M. (2003). *Estimating Intertemporal Allocation Parameters using Simulated Residual Estimation*. Cph.: Department of Economics, University of Copenhagen.



CAM

**Centre for Applied
Microeconometrics**

**Institute of Economics
University of Copenhagen**

<http://www.econ.ku.dk/CAM/>

**Estimating Intertemporal Allocation Parameters
using Simulated Residual Estimation**

**Sule Alan
Martin Browning**

2003-03

**The activities of CAM are financed by a grant from
The Danish National Research Foundation**

Estimating Intertemporal Allocation Parameters using Simulated Residual Estimation*

Sule Alan
Department of Economics
York University, Canada
salan@econ.yorku.ca

Martin Browning
Institute of Economics
University of Copenhagen
Denmark
Martin.Browning@econ.ku.dk

March 4, 2003

Abstract

There is widespread agreement that given currently available data, we cannot accurately estimate the parameters of intertemporal allocation using GMM on Euler equations, whether they be exact or approximate. Our reading of this literature and our own results is that this is a small sample (strictly, short panel) problem. The alternative seems to be to move to full structural modelling. In the current state of the art this is cumbersome, fragile and unable to deal with significant heterogeneity. We present a novel structural estimation procedure that is based on simulating expectation errors; we refer to it as Simulated Residual Estimation (SRE). We develop variants of the basic procedure that allow us to account for measurement error in consumption,

*We thank Japp Abbring, Orazio Attanasio, Richard Blundell, Hide Ichimura, Hamish Low, Lonnie Magee, Mike Veall and participants in the 2001 NBER Summer workshop and seminars at UCL and CREST for comments on the work reported in this paper. This paper was written during a visit by Sule Alan to the Institute of Economics at the University of Copenhagen. She thanks the Institute and the Centre for Applied Microeconometrics (CAM) for support. We are grateful to the Danish National Research Foundation (Grundforskningsfond) for support through its grant to CAM.

the ‘news’ in interest rate realisations and for heterogeneity in discount factors.

An investigation of the small sample properties of the SRE estimator indicates that it dominates GMM estimation of both exact and approximate Euler equations in the case when we have short panels and noisy consumption data. An empirical application to two panels drawn from the PSID are presented. The results are very encouraging. We find that we can estimate the parameters of intertemporal allocation much more precisely than with a conventional GMM on a log-linearised model. For example, we find that the 95% confidence interval for the EIS is $[0.27, 0.70]$ for the more educated whereas the 1GMM confidence intervals are $[-0.38, 0.90]$ and $[-3.78, 6.22]$ for the linearized and nonlinear models respectively. Moreover, the parameter estimates seem quite reasonable. For example, we find discount factors that are less than, but close to unity. We also find a higher discount factor for the more educated group. We find that the more educated have a higher CRRA which we interpret to indicate that the constant EIS assumption of the iso-elastic form is rejected. Finally we present results for a model that allows for heterogeneity in the discount factor within education groups. We reject strongly the homogeneity assumption and find that discount rates vary significantly within groups.

1 Introduction.

Over the past quarter century many attempts have been made to estimate the parameters governing intertemporal allocation using Euler equation techniques applied to micro data; Browning and Lusardi (1996) discuss the results of 25 studies using micro data and conclude that the results are disappointing. A number of recent Monte Carlo based papers have investigated why we experience this failure (Carroll 2001, Ludvigson and Paxson, 2001, Attanasio and Low, 2002); in section 2 we present some supplementary evidence on this issue. The problems identified are manifold but the most important seem to be the paucity of appropriate data (long panels on consumption) and the problem of dealing with the substantial measurement error in consumption (see Shapiro (1984), Altonji and Siow (1987) and Runkle (1991)). The latter means that we cannot use the exact Euler equation for estimation if the equation is non-linear in parameters (a point first made in the general con-

text of nonlinear GMM by Amemiya (1985)). The use of 'approximate' Euler equations (whether first order or second order) 'solve' the measurement error problem but bring with them new problems in that they introduce latent variables that lead to violations of the orthogonality conditions exploited by GMM methods. Thus Carroll (2001) concludes that "empirical estimation of consumption Euler equations should be abandoned". On the other hand, Attanasio and Low (2002) present results that suggest that the Carroll conclusion is overly pessimistic if we have long panels (40 periods, say) and time series variation in real rates. We do not find this conclusion too comforting for empirical work since we do not have long consumption panels.

Thus the emerging consensus seems to be that we must give up on empirical Euler equations and return to estimating consumption functions ('structural models') based on specifying the environment agents face (see Carroll and Samwick (1997), Gourinchas and Parker (2001) and Attanasio, Banks, Meghir and Weber (1999)). In practice these methods are very similar to calibration (as used in, for example, Hubbard, Skinner and Zeldes (1995)). The problems with this approach are that it is very cumbersome and can only accommodate very limited sources of uncertainty and heterogeneity. Moreover, results may not be robust to small changes in the specification of the structural model (for example, Browning and Ejrnaes (2001) show that the Gourinchas and Parker and Attanasio *et al.* (1999) results are very sensitive to how we account for family composition).

In this paper we focus on estimating the parameters of intertemporal allocation. We propose an alternative approach to GMM estimation of Euler equations that is based on simulating the distribution of expectations errors. We term our new procedure 'simulated residual estimation' (SRE). The key to our approach is that associated with every structural model there is a conditional expectations error distribution. We show that if we know this distribution and observe consumption paths and interest rates, then we can identify utility parameters (the discount factor and the elasticity of intertemporal elasticity) without having to specify the underlying stochastic environment. Without extra information the underlying model is not identified, but this is a strength rather than a weakness if we are only interested in preference parameters, since it gives the method robustness as compared with full-fledged structural estimation.

In section 3 we present an analysis of the distributions of expectations errors associated with models that are widely used in the literature (for example, nearly patient agents with unit root income processes, Deaton's (1991)

buffer stock model with explicit liquidity constraints and models with impatient agents with self-imposed liquidity constraints). This serves to develop intuition and to illustrate many of the points we wish to make. The main conclusion from our investigations is that almost all models that have been suggested in the literature give an expectations error distribution that can be adequately modelled as a mixture of two lognormal distributions.

In section 4 we present our estimator. To estimate, we use a simulation based method that is in the class of Simulated Minimum Distance (SMD) estimators. This involves the specification of ‘auxiliary parameters’ which are then matched to their theoretical predictions to estimate the parameters of interest. We find that the conventional linearized Euler equation provides a very simple and convenient vehicle to do this. The method suggested is many orders of magnitude faster than full structural estimation. Above we stated that we can recover the utility parameters if we know the expectations error distribution. Since we never do know the distribution, we address the problem of testing whether the distribution chosen for the estimation procedure is a good approximation using goodness-of-fit tests applied to the predicted distribution. We also briefly discuss the use of income and asset information in identification and improving precision and in accounting for heterogeneity in preferences and income processes.

In section 5 we present Monte Carlo evidence on our estimator and exact and approximate (GMM based) estimators. We take as designs for these simulations the designs used in the recent papers alluded to above. We find that if consumption is measured with even moderate error, exact Euler equation estimation performs poorly. We also replicate the previous finding that approximate methods do poorly if we have short panels. By contrast, our SMD estimator works well when other estimators do not. In particular, when there is considerable measurement error (for example, half the observed consumption growth variance is due to noise) our estimator works reasonably well even for moderate sample sizes.

Finally, in section 6, we present an empirical application of SRE to two panels drawn from the Panel Study of Income Dynamics (PSID) of the United States. We divide a sample of households into two broadly defined education groups and estimate the coefficient of relative risk aversion (and EIS), discount factor and measurement error variance for each group separately. We find that we can estimate the parameters of intertemporal allocation much more precisely than with a conventional GMM on a log-linearized model. For example, we find that the 95% confidence interval for the EIS is [0.27, 0.70] for

the more educated whereas the GMM confidence intervals are $[-0.38, 0.90]$ and $[-3.78, 6.22]$ for the linearized and nonlinear models respectively. Moreover, the parameter estimates seem quite reasonable. For example, we find discount factors that are less than, but close to unity. We also find a higher discount factor for the more educated group. We find that the more educated have a higher CRRA which we interpret to indicate that the constant EIS assumption of the iso-elastic form is rejected.

The remainder of the paper is organized as follows: The next section provides a detailed analysis of Euler equation for consumption and econometric issues regarding the estimation of such an equation. In Section 3 we present a discussion of the expectational error distributions associated with various models in the literature. Section 4 presents our SRE technique and presents some Monte Carlo results. In Section 5, we discuss the small sample properties of the estimator we propose as well as the properties of the traditional GMM based estimators. Section 6 presents an empirical application of two panels drawn from PSID. Section 7 concludes.

2 Euler Equation Estimation.

2.1 Exact Euler equation estimation.

We consider a standard intertemporal optimization problem for which agent h has expected utility at time t of:

$$E_{h,t} \left[\sum_{j=0}^{T-t} \frac{v(C_{h,t+j})}{(1+\delta)^j} \right] \quad (1)$$

where C is non-durable consumption, $v(\cdot)$ is an increasing, strictly concave sub-utility function, δ is a discount rate and $E_{ht}(\cdot)$ denotes the expectations operator conditional on the information that agent h has at time t . The evolution of assets over time is given by:

$$A_{h,t+j+1} = (1 + r_{h,t+j})A_{h,t+j} + Y_{h,t+j} - C_{h,t+j} \quad (2)$$

where A is assets, Y is stochastic labor income and r is the stochastic real rate of interest. The first order condition for the optimization problem gives the Euler equation for consumption:

$$v'(C_{h,t}) = \frac{1}{(1+\delta)} E_{ht} [(1 + r_{h,t+1})v'(C_{h,t+1})] \quad (3)$$

A widely used functional form for the sub-utility function is the iso-elastic form:

$$v(C_{h,t}) = \frac{(C_{h,t})^{(1-\gamma)}}{(1-\gamma)} \quad (4)$$

where the parameter γ is the coefficient of relative risk aversion (CRRA), which we assume is the same for everyone. Interest usually centres on the inverse of this parameter, the elasticity of intertemporal substitution (EIS):

$$\phi = \frac{1}{\gamma} \quad (5)$$

Low values of the EIS indicate an aversion to fluctuating consumption streams.¹ For the iso-elastic case with exponential discounting the only other preference parameter in this program is the discount rate δ . From the above we have the *exact Euler equation*:

$$\left(\frac{C_{h,t+1}}{C_{h,t}}\right)^{-\gamma} \frac{(1+r_{h,t+1})}{(1+\delta)} = \varepsilon_{h,t+1} \text{ with } E_{h,t}(\varepsilon_{h,t+1}) = 1 \quad (6)$$

This relationship has been the basis of very many estimates of the preference parameters and tests for the validity of the standard orthogonality assumptions in general and for the “excess sensitivity” of consumption to predictable income growth in particular. GMM estimation is based on the assumed orthogonality of the error term $\varepsilon_{h,t+1}$ to all variables dated t or before, such as lagged consumption, interest rate and income variables. As originally emphasized by Hall (1978), this is a very attractive procedure since one can estimate the preference parameters without explicitly parameterizing the stochastic environment that agents face.

Problems for GMM estimation on micro data arise if the consumption data are measured with error. For example, if we allow for a multiplicative measurement error so that observed consumption $C_{h,t}^o$ is given by:

$$C_{h,t}^o = C_{h,t}\eta_{h,t} \text{ with } E(\eta_{h,t}) = 1 \quad (7)$$

then the exact Euler equation for observable consumption becomes

$$\left(\frac{C_{h,t+1}^o}{C_{h,t}^o}\right)^{-\gamma} \frac{(1+r_{h,t+1})}{(1+\delta)} = \left(\frac{\eta_{h,t+1}}{\eta_{h,t}}\right)^{-\gamma} (1+\varepsilon_{h,t+1}) \quad (8)$$

¹We prefer to emphasise the role of this parameter as representing aversion to fluctuations rather than to risk since it is operative even when there is no uncertainty.

The problem this gives is that the composite error term does not have a conditional expectation of unity, even if we assume that $\eta_{h,t+1}$ and $\varepsilon_{h,t+1}$ are independent:

$$\begin{aligned} E_t \left[\left(\frac{\eta_{h,t+1}}{\eta_{h,t}} \right)^{-\gamma} (1 + \varepsilon_{h,t+1}) \right] &= E_t \left(\frac{\eta_{h,t+1}}{\eta_{h,t}} \right)^{-\gamma} E_t (\varepsilon_{h,t+1}) \\ &= (\eta_{h,t})^\gamma E_t (\eta_{h,t+1})^{-\gamma} \neq 1 \end{aligned} \quad (9)$$

It is now widely accepted that household level consumption data information is likely to be very noisy. For example, Runkle (1991) estimates that 76% of the variation in the growth rate of food consumption in the PSID is noise. Dynan (1993) reports that the standard deviation of changes in log consumption in the CEX (American Consumer Expenditure Survey) is 0.2, which seems too large for ‘true’ variations. The other widely used data resource are quasi-panels, constructed from cross-section expenditure survey information by taking within-period means following the same population (e.g. means over all the 25 year olds in one year and all the 26 year olds in the next year). Although this averaging reduces the effect of measurement error, the construction of quasi-panels from samples which change over time induces sampling error which is very much like measurement error.

The presence of measurement error when estimating non-linear equations is problematic. In our context, the basic problem is that measurement error makes it appear as though consumption is less smooth over time than it actually is, which results in too low an estimate for the CRRA (with a consequent bias of the EIS away from zero). Carroll (2001) shows this in simulations with only cross-section variation in interest rates ($r_{h,t} = r_h$ for all t). To show the extent of the problem when we have time varying interest rates, we take a similar environment to Carroll (2001)² with the polar case in which everyone faces the same stochastic interest rate ($r_{h,t} = r_t$ for all h). We construct optimal consumption paths with a CRRA of 4 and then add a multiplicative error on the consumption values. Taking a measurement error variance such that 50% of the time series variation in consumption growth is noise, the average CRRA estimate is 2.78 which indicates substantial bias. Increasing the number of cross section units does not affect the bias.

²Fuller details of our simulation procedures will be given below.

2.2 Approximate Euler equation estimation.

A natural alternative to GMM estimation of the exact Euler equation is GMM estimation of the first or second order approximation to the nonlinear Euler equation (the first derivation is due to Hansen and Singleton(1983); see, for example, Carroll (2001) for the derivations we now present). From equation (6) we have the following (log) quadratic consumption growth equation:

$$\Delta \log C_{h,t+1} - \alpha - \frac{1}{\gamma} r_{h,t+1} - \frac{\gamma}{2} (\Delta \log C_{h,t+1})^2 = e_{h,t+1} \quad (10)$$

where the constant term α contains the discount rate and means of the third and higher order unconditional moments of the error term $\varepsilon_{h,t+1}$. The error term $e_{h,t+1}$ contains the expectational error and also time varying components of the higher conditional moments (conditional on past information).³ The first order log-linear approximation (equation (10) without the squared term) has been used very extensively in the applied micro literature due to the fact that a multiplicative measurement error becomes additive as a result of log linearization. The usual (and uncontroversial.) assumption is that the instruments other than consumption that are used in the estimation are uncorrelated with the measurement error. The *MA*(1) error structure induced in the errors due to the measurement error is easily accounted for in GMM. Most researchers use twice (or more) lagged variables for instruments (but note that we could use first lags of any variable other than consumption since these are assumed uncorrelated with the measurement error). The problem with this approach is that the movements in the higher order moments (for example, the skewness) that are subsumed into the error term will generally cause it to be correlated with lagged variables, which leaves us without any instruments for GMM.

Here we present a brief discussion of the findings of Carroll (2001), Ludvigson and Paxson (2001) and Attanasio and Low (2002); in our simulations we shall replicate many of their results and discuss them in greater detail. Ludvigson and Paxson (2001) solve and simulate a life cycle model with

³This brings out clearly that the one parameter γ controls attitudes to fluctuating consumption paths (through the coefficient on the real rate) and prudence (through the coefficient on the squared term). This close identification of fluctuation aversion and prudence is solely a result of using the iso-elastic form; other forms break the link between aversion to fluctuations and prudence (for example, the quadratic utility function has fluctuation aversion but no prudence).

stochastic income and an additively separable iso-elastic utility function assuming a fixed interest rate of 3% and a discount rate of 5% (so that agents are assumed to be impatient)⁴. They then follow Dynan (1993) and use the simulated data to estimate relative prudence using the second order approximation to the Euler equation (equation (10) with no interest rate)⁵. They find that the estimate of the CRRA is downward biased; that is, it is estimated that agents are less averse to fluctuations than they actually are. Carroll (2001) performs a similar analysis allowing for cross-section variation in the interest rate, but no time series variation. He finds that the estimate of the CRRA is upward biased.

Neither Ludvigson and Paxson nor Carroll allow for time variation in interest rates to identify the EIS. Our own feeling is that trying to estimate the intertemporal price elasticity (the EIS) without some intertemporal variation in price is almost certainly doomed to failure. Attanasio and Low (2002) present results allowing for time series variation in interest rates. They solve and simulate a simple life cycle model with stochastic income and interest rates and then estimate first and second order approximations to the Euler equation. They argue that one can estimate the EIS consistently if the time period of the sample is long enough⁶.

However, for panel lengths of, say, 20 periods there is still considerable bias, so that the Attanasio and Low results are not very encouraging empirically. Attanasio and Low also show in their Monte Carlo study that the precision of the estimates increases considerably with the variance of the interest rate. A potential problem they identify is that even moderately impatient agents will typically hold net wealth stocks that are close to any borrowing limit they face. In this case consumption becomes very sensitive to the income shocks and it is difficult to extract the relatively small variations in consumption growth due to interest rate changes. Note however that this problem is not special to the approximate Euler equations; our simula-

⁴The term ‘impatient’ here and henceforth refers to the condition $\delta - r > 0$. Note that, if income grows overtime, consumers can be impatient even if $\delta = r$. But for all the models considered in this paper zero income growth is assumed.

⁵In the case of iso-elastic utility, the relative prudence parameter is $\frac{\gamma+1}{2}$. Ludvigson and Paxson (2001) assume fixed interest rate and estimate the equation

$$\Delta \log C_{h,t+1} = \alpha + \frac{\gamma+1}{2} (\Delta \log C_{h,t+1})^2 + e_{h,t+1}.$$

⁶They use the term ‘consistent’ as $T \rightarrow \infty$. They experiment with different T and show that the mean estimate of the EIS approaches its true value while the estimated standard errors become smaller as T increases.

tions presented below suggest that the same problem arises for exact Euler equation with no measurement error.

2.3 The implications of these analyses.

We draw the following implications from the analyses of Ludvigson and Paxson (2001), Carroll (2001), Attanasio and Low (2000) and our own supplementary investigations.

1. There is not much point in trying to estimate price elasticities (such as the EIS) without some variation in the price (in this context, variations in the real interest rate).
2. Attempts to gauge the extent of prudence are more successful if agents are impatient since then we will be observing the buffer stock savings due to income uncertainty.
3. Attempts to measure the EIS (reactions to interest rate changes) are more successful if agents are patient and build up wealth. In this case, temporary changes in income do not lead agents to vary consumption (since they have assets to smooth their consumption) so that the extraction of the consumption growth/ interest rate signal is easier.
4. If there is a liquidity constraint and agents are sometimes constrained, then the Euler equation does not hold in periods of constraint and consumption is not affected at all by the interest rate (an effective EIS of zero). This leads to a downward bias in the estimate of the EIS if we do not observe whether or not the agent is constrained and proceed as if they never are.
5. Measurement error introduces considerable bias into GMM exact Euler equation estimates of the EIS, even if there is substantial variation in the interest rate. This problem does not reduce if we have a large number of cross-section units.
6. Generally, approximate Euler equation methods do badly, but the results of Attanasio and Low show clearly that this is a small sample (small T) problem. This requires a small- T solution, for the typical case in which we have short and noisy panels.

The net result of the above is that we agree with Carroll (2001) and Ludvigson and Paxson (2001) that the econometric methods we currently have to hand are not up to estimating the EIS (or the discount rate) on short and noisy panels. What alternatives remain? One is to revert to old style consumption studies that are only loosely linked to conventional life-cycle theory. This is not very attractive to a generation raised on dynamic general equilibrium models and empirical modelling that stays close to the theory. A second alternative is to move to estimation based on structural models. Thus Carroll and Samwick (1997) perform a structural estimation in which they identify the discount rate; all of the other parameters are fixed at ‘reasonable’ values. Gourinchas and Parker (2001) use structural estimation to estimate the EIS and the discount factor. They use CEX information and Method of Simulated Moments estimation which matches the moments generated by the data with that of simulated data. This procedure involves the numerical solution of the dynamic programming problem for every parameter value that the estimation procedure considers. The procedure is extremely slow. An obvious problem regarding this approach is the fact that one needs to specify the underlying stochastic process (income process in their case since they use a fixed interest rate) which is not necessary for Euler equation estimation (whether exact or approximate). It is not clear whether a slight misspecification of the income process will not completely change the results. To examine this would require that the estimation procedure be analyzed under misspecification, which would be extremely time consuming. Although full structural modelling is potentially promising, an alternative is needed that reduces substantially the computational burden without sacrificing the close link to the theory. We present here an alternative that relies on simulating the distribution of expectations errors directly.

3 The distribution of expectation errors.

Below we shall present an alternative approach to GMM which is based on sampling from the conditional distribution of the expectations error. In this section we present an extended discussion of the distributions associated with various models in the literature. In order to illustrate our point, we present a wide range of models with different sets of parameters and different income processes within the time separable iso-elastic utility framework. We consider both fixed and stochastic interest rate models. We assume a finite lifetime

of 70 periods with no bequest motive and we start all agents off with zero wealth. After generating a 70-period consumption path for an individual, we remove the first and the last 20 periods. Further details of the simulation methods are given in the Appendix. Table 1 presents the features of the 14 models we consider. The main differences across models are in the income processes; the degree of impatience; the presence of liquidity constraints and the presence of heterogeneity. We assume that agents face two types of income shocks, permanent and transitory. The assumed income process is as follows

$$Y_{h,t+1} = P_{h,t}u_{h,t+1} \quad (11)$$

where u_{t+1} is an iid lognormal transitory shock with mean 1 and a constant variance $(e^{\sigma_u^2} - 1)$ and $P_{h,t}$ is permanent income which follows the following random walk process

$$P_{h,t+1} = P_{h,t}z_{h,t+1} \quad (12)$$

where $z_{h,t+1}$ is an iid lognormal permanent shock with mean one and a constant variance $(e^{\sigma_z^2} - 1)$. We assume that the innovations to income are independent over time and across individuals so that we assume away aggregate shocks to income⁷.

Model 1 is our benchmark model with a standard CRRA (of 4), the discount rate equal to the constant real rate of interest and no liquidity constraints. Most of our models are simple variants of this benchmark model. Model 2 allows for less aversion to fluctuations (risk), specifically a CRRA of 2. Model 3 allows for impatience. Model 4 is the same as the benchmark model except that we have a higher permanent income variance. In Models 5 agents face iid Normally distributed income shocks with a unit mean and a constant variance (Deaton 1981). Model 6 assumes the same income process as the benchmark model, but in this case the process is given a small probability of zero income in any period (a ‘Carroll’ process). This assumption imposes an implicit liquidity constraint so that agents optimally choose not to borrow in any period over the life cycle. Note that this case differs from a Deaton type buffer stock model where the liquidity constraint is explicit. Since agents optimally choose not to borrow the Euler equation always holds in model 6. Model 7 imposes an explicit liquidity constraint and model 8 imposes the liquidity constraint on impatient agents (a Deaton buffer stock environment). Model 9 is the same as the baseline model with a stochastic

⁷We allow for macro shocks in our empirical work below.

	CRRA	Real	Income	Liquidity
Model	γ	rate, r	process	constraint
1	4	0.05 (0)	$\sigma_z = 0.02, \sigma_\varepsilon = 0.1$	No
2	2	0.05 (0)	$\sigma_z = 0.02, \sigma_\varepsilon = 0.1$	No
3	4	0.03 (0)	$\sigma_z = 0.02, \sigma_\varepsilon = 0.1$	No
4	4	0.05 (0)	$\sigma_z = 0.05, \sigma_\varepsilon = 0.1$	No
5	4	0.05 (0)	IID normal $\sigma_t = 0.1$	No
6	4	0.05 (0)	Carroll Process*	Implicit
7	4	0.05 (0)	$\sigma_z = 0.02, \sigma_\varepsilon = 0.1$	Yes
8	4	0.03 (0)	$\sigma_z = 0.02, \sigma_\varepsilon = 0.1$	Yes
9	4	0.05 (0.025)	$\sigma_z = 0.02, \sigma_\varepsilon = 0.1$	No
10(1&2)	2 or 4	0.05 (0)	$\sigma_z = 0.02, \sigma_\varepsilon = 0.1$	No
11(1&3)	4	0.05/0.03	$\sigma_z = 0.02, \sigma_\varepsilon = 0.1$	No
12 (1&4)	4	0.05 (0)	mixed	No
13	Model 1 with low measurement error (40% noise)			
14	Model 1 with high measurement error (60% noise)			
Note: variance of r in parentheses. Discount rate = 0.05 for all models. σ_z : std of permanent income shocks, σ_ε : std of transitory income shocks				
* Model 1 with 1% probability of zero income.				

Table 1: Models

real interest rate. Models 10, 11 and 12 allow for some heterogeneity. Model 10 allows that there may be heterogeneity in the CRRA; specifically we assume a mixing model in which agents have $\gamma = 4$ or $\gamma = 2$ with probability one half. Model 11 mixes impatient and patient agents (heterogeneity in discount rates). Model 12 allows that agents have different income processes; specifically they have a low or high permanent income variance with probability half. Finally, the last two models experiment with lognormally distributed measurement error. In these cases we allow for measurement error in the benchmark model. In model 13, 40% of consumption growth variation is noise whereas the noise is increased to 60% in model 14.

For models 1 to 9, consumption paths were generated for 500 ex-ante identical individuals. Given time paths $C_{h,t}$, we generate errors for agent h in period $t + 1$ by:

$$\varepsilon_{h,t+1} = \left(\frac{C_{h,t+1}}{C_{h,t}} \right)^{-\gamma} \frac{(1 + r_{t+1})}{(1 + \delta)} \quad (13)$$

Note that the interest rate varies over time only in model 9. In each case we use only the observations from periods 20 to 50 (to give expectations errors for periods 21 to 50) to minimize the impact of starting and end effects. Consumption paths for models 10 to 14 are generated from the consumption paths of models 1 to 9. For the mixing models 10 to 12 we randomly select 250 paths from each of the component models and use these. For the measurement error models 13 and 14, we take the consumption paths from model 1 and generate observed paths by multiplying them with independently distributed lognormal, multiplicative, and unit mean measurement errors. Distributional features of the expectations errors for these 14 models are presented in Table 2.

The first four columns of Table 2 give the first four moments for the distribution. The next column gives the probability for a test that the mean is unity. The final two columns give tests that the distribution is lognormal and a mixture of two lognormals respectively.

The main features of the expectational error distribution for the benchmark model (model 1) are that it has unit mean (as we would expect) and some skewness. The right (positive) skewness is observed for most of our models; it reflects the concavity of consumption function in cash-on-hand (current earnings plus non-labour wealth) and the asymmetry of the income processes we use in all models except for model 5. Although the skewness is

Mod	mean	std	skw	krt	U-test	L-test	M-test
1	1.000	0.081	0.26	3.19	68	99	94
2	1.000	0.040	0.14	3.07	75	54	30
3	1.000	0.090	0.31	3.27	62	84	90
4	1.000	0.173	0.50	3.50	90	89	70
5	1.000	0.024	-0.15	3.08	62	2.3	94
6	1.000	0.189	14.2	439	78	0	0
7	1.000	0.086	0.55	7.16	65	8.4	23
8	0.993	0.161	3.36	50.4	0	0	0
9	1.002	0.131	0.02	3.16	13	0	51
10	1.000	0.064	0.30	4.21	71	0	92
11	1.000	0.086	0.30	3.19	65	92	50
12	1.000	0.135	0.51	4.69	99	0	94
13	1.006	0.136	0.43	3.38	0	98	73
14	1.017	0.206	0.60	3.67	0	63	30
U-test is p-value (in %) for a unit mean; L-test is p-value for lognormality; M-test is p-value for mixture of lognormals. Number of observations = 15,000							

Table 2: Distributions of Expectational Errors for Different Models

small, a formal test of normality rejects decisively. Instead, we find that the expectations errors for this simple model appear to be lognormally distributed (see the column headed L-test). Model 2 differs from the benchmark in having a lower risk aversion. As can be seen, a lower CRRA is associated with a lower standard deviation for the unconditional expectations errors.⁸ The mapping from the CRRA to the expectations error variance seen in this simple comparison forms a partial basis for identification in the estimation scheme we present below. Of course, the variance of the distribution will also depend on environmental factors such as the amount of interest rate variation and the underlying earnings process, so that more is required for identification. Note as well that a lower aversion to fluctuations leads to lower skewness; this reflects the fact that a lower CRRA is closer to linear (risk neutrality which in turn implies no prudence) so that the consumption function is less concave. A comparison of models 1 and 3 reveals that higher impatience is associated with a higher variance and slightly more skewness. Once again we do not reject lognormality. Model 4 indicates that a higher income variance leads to a higher error variance distribution with larger skewness and kurtosis. All of models 1 to 4 are very standard and, as we have seen, generate lognormal expectations error distributions.

In model 5 we vary the income process to being simply an iid normal process with a unit mean.⁹ As can be seen, the expectations error distribution for this process is not right skewed and we reject lognormality but not the mixture model. In model 6 we assume the baseline environment except that agents are sometimes (but rarely) hit by a zero income draw. In this case, current consumption relative to the previous and following period is small (except in the very rare case in which the agent receives two consecutive zero income draws). Since we then take (the inverse of) fourth powers, the associated expectations errors are very large, hence the very pronounced skewness and kurtosis. Although such a process is not very realistic, it serves to illustrate that expectations errors can be very far from lognormal or a mixture of lognormals, as can be seen from the reported probabilities.

Models 7 and 8 introduce a no-borrowing constraint. Even though model 7 agents are not impatient, the constraint sometimes binds and the conse-

⁸The variances of consumption growth for the two models (not shown) are very similar so that the lower value in model 2 is because we are taking the inverse of a square rather than the inverse of the fourth power (see equation (13)).

⁹In practice, we take a discrete approximation to the Normal with many points of support. The support of the discrete approximation is bounded away from zero.

quent error distribution is more skewed and heavier tailed than the benchmark distribution. We do not, however, reject the mixture model. The same is not true for the model 8 in which the agents are impatient. They never accumulate much in the way of assets and often bump up against the borrowing limit which gives them a lower current consumption than they would wish, relative to the future. The effect on the expectations error distribution for the model 8 is interesting. First we see that the mean is less than unity.¹⁰ Second, the other moments are quite different from those of the benchmark model. Finally, we see that a mixture of lognormals does not fit the distribution.

In model 9 we extend the benchmark model by allowing for a stochastic real rate. This is also the model we use to investigate the small sample properties of our competing estimators in Section 5. It is important to note that although we allow for time series variation in interest rates we assume away cross section variation. The distributional properties of model 9 are somewhat different from the benchmark case. In particular, the variance is higher and the skewness is close to zero. We reject lognormality for model 9 but we can fit the distribution well with a mixture model.

Turning to the effect of heterogeneity, we see from the results for model 10 that introducing heterogeneity in the CRRA leads to moments that are similar to those of the models for which it is a mixture (models 1 and 2) but a decisive rejection of lognormality. The mixture of lognormals is not rejected. For discount rate heterogeneity (with no liquidity constraints), model 11, we find very similar results to the component models (1 and 3) with no rejection of lognormality. For income process heterogeneity (model 12) we find fatter tails and a consequent rejection of lognormality but not of the mixture model.

Finally we turn to the effects of measurement error. As can be seen, adding measurement error increases the mean of the expectations errors (see equation (9)) and the error variance. However, the distribution changes in such a way that lognormality is preserved.

In this section we have presented the expectations error distributions associated with a wide range of models. A number of points emerge. First, different underlying environments may give rise to similar distributions. As we shall discuss below, this will impact on the data needs for identification. Second, we do not reject lognormality for many of our models and we do not

¹⁰This is for the error distribution. The ratio of current to lagged consumption also should be below unity since the agent is impatient.

reject a mixture of lognormals with a unit mean for most models. This will be used extensively in our estimation procedure. The two major deviations from the mixture of lognormals occur when we use a Carroll income processes or there are explicit liquidity constraints and agents are impatient (the Deaton buffer stock model). In our empirical work below we select households who are less likely to be in this class of agents.

One important feature to emerge from our analysis is that although we have a non-stationary environment (because of the finite horizon) the unconditional distribution of the expectation errors is very stable.¹¹ Thus for all models the distribution in period 10 is very similar to that in period 60 (10 years from the end of life). In our estimation procedure we shall assume that the expectations errors are stationary.

In addition to examining basic distributional features of the expectational error distributions generated by different models, we checked and confirmed standard orthogonality assumption (no correlation of errors with their lags).¹² Simple regressions of errors on their lags squared, squared errors on lags squared and similar third and fourth moment regressions led us to the conclusion that the assumption of independently distributed expectational errors (allowing correlation across individuals via interest rate shocks) is plausible. This is not to say that we cannot define a model that violates the independence assumption. For this reason we experimented with a model with heteroscedastic income shocks and found that even with considerable heteroscedasticity in income shocks we still do not see serious dependence in the expectational errors.

4 Estimation methods.

4.1 Simulated Minimum Distance.

Our estimation procedure is simulation based. Following Hall and Rust (2002) we refer to the general technique as Simulated Minimum Distance (SMD) since it is based on matching (minimizing the distance between) statistics from the data and from a simulated model. The class of SMD estimators includes the EMM procedure of Gallant and Tauchen (1996) and the Indirect Inference methods of Gouriéroux, Monfort and Renault (1993).

¹¹These results are not shown in the paper but are available upon request.

¹²Again, the results are available upon request.

Here we present a short account of the method as applied generally to panel data; see Hall and Rust (1999) and Alvarez, Browning and Eyrnæs (2001) for details.

Suppose that we observe $h = 1, 2..H$ units over $t = 1, 2..T$ periods recording the values on a set of Y variables that we wish to model and a set of X variables that are to be taken as conditioning variables. Thus we record $\{(Y_1, X_1), \dots (Y_H, X_H)\}$ where Y_h is a $T \times l$ matrix and X_h is a $T \times k$ matrix. For modelling we assume that Y given X is identically and independently distributed over units with the parametric conditional distribution $F(Y_h|X_h; \theta)$, where θ is an m -vector of parameters.¹³ If this distribution is tractable enough we could derive a likelihood function and use either maximum likelihood estimation or simulated maximum likelihood estimation. Alternatively, we might derive some moment implications of this distribution for observables and use GMM to recover estimates of a subset of the parameter vector. Sometimes, however, deriving the likelihood function is extremely onerous; in that case, we can use SMD if we can simulate Y_h given the observed X_h and parameters for the model. Thus we choose an integer S for the number of replications and then generate $S * H$ simulated outcomes $\{(Y_1^1, X_1), \dots (Y_H^1, X_H), (Y_1^2, X_1), \dots (Y_H^S, X_H)\}$; these outcomes, of course, depend on the model chosen ($F(\cdot)$) and the value of θ taken in the model.

Thus we have some data on H units and some simulated data on $S * H$ units that have the same form. The obvious procedure is to choose a value for the parameters which minimizes the distance between some features of the real data and the same features of the simulated data. To do this, define a set of auxiliary parameters that are used for matching. Gallant and Tauchen (1996) suggest first finding a ‘score generator’ (flexible quasi-likelihood function) which nests the true model, and then using the score vector from this as auxiliary parameters. In the Gouriéroux *et al.* (1993) Indirect Inference procedure, the auxiliary parameters are maximizers of a given data dependent criterion which constitutes an approximation to the true DGP. Both of these approaches are motivated by attempts to derive estimators that have efficiency properties that are close to MLE. In Hall and Rust (1999), the auxiliary parameters are simply statistics that describe important aspects of the data; this is very close to calibration. We follow

¹³This could be generalised to allow for correlated heterogeneity by allowing for dependence on the initial values, as in Chamberlain (1980), Blundell and Smith (1991) and Wooldridge (2000).

this approach. Thus we first define a set of J auxiliary parameters (below we shall discuss in detail how to do this for the intertemporal problem):

$$\gamma_j^D = \frac{1}{H} \sum_{h=1}^H g^j(Y_h, X_h), \quad j = 1, 2, \dots, J \quad (14)$$

where $J \geq m$ so that we have at least as many auxiliary parameters as model parameters. Denote the J -vector of auxiliary parameters derived from the data by γ^D . Using the same functions $g^j(\cdot)$ we can also calculate the corresponding values for the simulated data:

$$\gamma_j^S = \frac{1}{S * H} \sum_{s=1}^S \sum_{h=1}^H g^j(Y_h^s, X_h), \quad j = 1, 2, \dots, J \quad (15)$$

and denote the corresponding vector by $\gamma^S(\theta)$ where the notation explicitly shows the dependence on the model parameter values (but not the dependence on the X variables observed). Identification follows if the Jacobian of the mapping from model parameters to auxiliary parameters has full rank:

$$\text{rank}(\nabla_{\theta} \gamma^S(\theta)) = m \text{ with probability } 1 \quad (16)$$

This effectively requires that the model parameters be ‘relevant’ for the auxiliary parameters.

Given sample and simulated auxiliary parameters we take a $J \times J$ positive definite matrix W and define the SMD estimator:

$$\hat{\theta}_{SMD} = \arg \min_{\theta} (\gamma^S(\theta) - \gamma^D)' W (\gamma^S(\theta) - \gamma^D) \quad (17)$$

Alvarez *et al.* (2001) perform a small Monte Carlo study and argue that it is best to work with just identified models ($J = m$). This is largely because the objective function may have many local minima and we can only be sure we have converged to a global minimum (not necessarily unique) if the model is just identified. In the just identified case the choice of W is irrelevant (except for computational reasons) and the minimized criterion should be zero. For just identified models, we would conclude that the model is ‘well-specified’ (relative to a particular choice of m auxiliary parameters) if and only if there is some value of the model parameters such that $\gamma^S(\hat{\theta}_{SMD}) = \gamma^D$. Typically we have $J > m$; in this case we use m of the auxiliary parameters to fit the model and the remaining $J - m$ auxiliary parameters to test for the goodness of fit.

4.2 Simulated Residual Estimation.

We turn now to applying SMD techniques in our specific context. Suppose we have observations on H households followed for T years. We begin by assuming that we only observe household consumption in each period and real rates between periods (which we assume to be time varying but common across agents); thus we observe $\{r, C_h\}_{h=1, \dots, H}$. Below we shall consider the case where we also observe earnings and asset levels. For the moment we assume that we observe consumption with no measurement error; we shall deal with this in the next sub-section. Since we introduce two innovations in modelling (SMD and the use of simulated residuals) we begin by considering how we would use SMD to estimate preference parameters if we used full structural modelling with each agent having the same finite horizon Υ (with Υ chosen to be somewhat larger than T to be able to remove the beginning and end effects). We proceed in a number of steps.

1. First we define (perhaps joint) processes for income, the real rate and anything else that affects intertemporal allocation. Usually these would be estimated using data taken from the population from which we draw our sample.
2. Next we take parameter values for preferences (typically, the EIS and the discount rate).
3. Then we derive Υ period specific consumption (policy) functions conditional on current state variables (typically, the current realization of income and the real rate for dependent processes and current cash on hand). It is rarely possible to do this analytically, so that we need to use numerical policy (or value) iteration methods. The last period consumption function is trivial (consume everything) and the consumption functions for the earlier periods are obtained by backward induction.
4. At this point we are ready to start simulation. To simplify the exposition we assume that we set S (the number of replications of the panel in the SMD estimation procedure) equal to unity and draw H first period income and real rate values (conditional on ‘period 0’ values for income and the real rate) and give each of the synthetic H agents a starting value for assets. From the consumption function for period 1 we calculate first period consumption for each agent. We then draw new values

of income and the real rate and calculate period 2 consumption and so on. The end result of this is a set of Υ real rate, consumption, income and asset realizations for each synthetic unit. We then trim these to remove starting and end effects and to give a time series of length T for each household $\{r^s, C_h^s\}_{h=1, \dots, H}$ (where now the s subscript reminds us that this is simulated data).

5. We now need to choose auxiliary parameters. Since we have two parameters (the EIS and the discount rate) we need two auxiliary parameters. Our choice are the OLS coefficients in the simple regression of consumption growth on the real rate:

$$\Delta \log C_{h,t+1} = \alpha + \phi r_{h,t+1} + e_{h,t+1} \quad (18)$$

These are *not* unbiased estimates of the parameters of interest, but (under weak assumptions) they are unbiased estimates of something and that something is the same for the true data and the simulated data if we have the ‘true’ model. It is this property that makes SMD so useful. Note that we could equally well take the GMM estimates of the two parameters (with the constant and lagged interest rates as instruments) as auxiliary parameters; we prefer the OLS since it is simpler and quicker. We present results below that indicate that the choice of auxiliary parameters is not too important, provided the identification condition is satisfied.

6. The last step gives two sets of estimates: $(\hat{\alpha}_{OLS}^D, \hat{\phi}_{OLS}^D)$ for the data and $(\hat{\alpha}_{OLS}^S, \hat{\phi}_{OLS}^S)$ from the simulated data. We now compare the two sets of estimates. If they are the same, we stop. If they differ, we go back to step 2 and choose new parameter values. In practice, of course, we would embed this in an optimization routine or perform a grid search over the EIS and discount factor parameters.

It will be seen that steps 3 and 4 are very time consuming and estimation will be infeasible. We now present a technique which cuts out these steps. After step 4 we could define simulated expectation errors:

$$\varepsilon_{h,t+1}^S = \left(\frac{C_{h,t+1}^S}{C_{h,t}^S} \right)^{-\gamma} \frac{(1 + r_{t+1})}{(1 + \delta)} \quad (19)$$

Conversely, if we knew the distribution of the expectations errors we could simulate expectations errors, $\varepsilon_{h,t+1}^S$ and then we could construct paths of consumption ratios using:

$$\frac{C_{h,t+1}^S}{C_{h,t}^S} = \left\{ \frac{(1 + \delta)}{(1 + r_{t+1})} \varepsilon_{h,t+1}^S \right\}^{-\frac{1}{\gamma}} \quad (20)$$

This is, of course, very fast (as compared to steps 3 and 4 above). We then use the simulated paths in steps 5 and 6. The error simulation step requires a specification of the distribution of the expectations errors. One part of this easy: it should be serially uncorrelated with an unconditional mean of unity. If we now choose a simple two parameter form such as the lognormal then we have one extra model parameter to estimate. This in turn requires an extra auxiliary parameter; the obvious choice in step 5 above is to use the variance of OLS errors. Using a more flexible distribution such as a mixture of lognormals requires more auxiliary parameters for identification; we return to this below. We refer to our estimation procedure as Simulated Residual Estimation (SRE).

The algorithm for this simple case is:

1. Run OLS on the pooled sample of consumption growth on the real rate and record the estimates of the constant, the slope parameter, the variance of the error term and the correlation coefficient between the error term and interest rate shocks¹⁴.
2. From the standard Normal, draw standardized simulated residuals $\nu_{h,t}^S$ for $t = 2, \dots, T$ and $h = 1, \dots, H$. These standardized errors are kept constant from iteration to iteration.
3. Choose a standard deviation for the expectations error distribution σ_ε and construct simulated expectations errors:

$$\varepsilon_{h,t}^S = \exp \left(-\frac{\ln(1 + \sigma_\varepsilon^2)}{2} + \sqrt{\ln(1 + \sigma_\varepsilon^2)} \nu_{h,t}^S \right) \quad (21)$$

¹⁴We construct the interest rate shocks as $\Delta r_{t+1} = r_{t+1} - r_t$ assuming everybody faces the same shocks (this is to generate a high degree of aggregate shocks). We also tried an AR(1) process instead of a random walk. Our Monte Carlo results indicate that SRE is robust to slight of the interest rate process.

where $\nu_{h,t}^S$ is standard normal variable; see Appendix B for details. Choose values for the intertemporal allocation parameters (γ, δ) and correlation coefficient between the expectations error and interest rate shocks $(\rho_{\varepsilon r})$. Construct consumption ratios using equation (20).

4. Repeat step 1 for the simulated data.
5. If the values from steps 1 and 4 are the same, stop. Otherwise, go to step 3 (so that we keep the same expectations error from iteration to iteration) and revise the choice of $(\gamma, \delta, \sigma_\varepsilon, \rho_{\varepsilon r})$.

In practice we would once again use either an optimization algorithm to revise parameter values or perform a grid search. In either case the computational time is much lower than for full structural estimation.

4.3 Accounting for measurement error.

In the account of SRE given in the last sub-section we ignored the possibility that consumption is measured with error. The log-linearized equation was introduced largely to take account of measurement error since any multiplicative measurement error is incorporated into the error term and as long as it is uncorrelated with the instruments used in the estimation it does not distort the parameter estimates. The only complication arising for GMM estimation of the approximate Euler equation is that the error terms in the consumption growth equation will have an MA(1) structure since we are first differencing the noise. This suggests an auxiliary parameter that will allow us to take account of measurement error in SRE.

If we assume that the measurement error is multiplicative lognormal with unit mean then we need to estimate one extra parameter, the standard deviation of the measurement error, σ_η . If we assume that the only source of auto-correlation in the error term is the measurement error, we can simply use the extent of the first order auto-correlation as an auxiliary parameter. In general, first order auto-correlation in such the Euler equation may also indicate other features such as iid preference shocks. As far as our estimation method is concerned, this can lead to a biased estimate of the magnitude of noise in the data but the structural estimates will not be affected. The steps of the estimation with measurement error are as follows:

1. Run an OLS of consumption growth on the real rate and record the estimates of the constant, the slope parameter, the variance of the error

term, the correlation coefficient between the error term and interest rate shocks and the auto-correlation parameter of the regression errors.

2. From the standard Normal, draw standardized simulated residuals $\nu_{h,t}^S$ and measurement errors $\tilde{\eta}_{h,t}^S$ for $t = 2, \dots, T$ and $h = 1, \dots, H$.
3. Choose standard deviations for the expectations error distribution, σ_ε , and the measurement error, σ_η , correlation coefficient between the error term and interest rate shocks ($\rho_{\varepsilon r}$) and construct simulated expectations errors, $\varepsilon_{h,t}^S$, and measurement errors, $\eta_{h,t}^S$, as above. Choose values for the intertemporal allocation parameters (γ, δ). Construct consumption ratios using equation (20). Introduce measurement error by multiplying the consumption ratio by measurement error ratios to define ‘observed’ simulated consumption ratios:

$$\frac{C_{h,t+1}^S \eta_{h,t+1}^S}{C_{h,t}^S \eta_{h,t}^S} \quad (22)$$

4. Repeat step 1 for the simulated data with the ‘observed’ simulated consumption ratios.
5. If the values from steps 1 and 4 are the same, stop. Otherwise, go to step 3 and revise the choice of $(\gamma, \delta, \sigma_\eta, \sigma_\varepsilon, \rho_{\varepsilon r})$.

Thus a simple model with two preference parameters can be estimated using data on (noisy) consumption levels and interest rates. What if we now observe more?

4.4 Using income and asset information.

SRE relies on specifying the conditional distribution of the expectations errors. All of the above only uses consumption and interest rate information. The strength that our method shares with the Euler equation approach is that we do not have to specify the income or interest rate processes. Generally, however, one would expect to achieve better results (in terms of identification of heterogeneity and in terms of precision) by using the observed income series for each household. As we saw in section 2, the lognormality assumption is a poor one if there are sometimes very low income realizations. Our own feeling is that identification in non-standard situations requires more

Parameter	Value
Coefficient of Relative Risk Aversion, γ	4
Discount rate, δ	0.05
mean r , μ	0.05
AR(1) coefficient of r	0.6
Standard deviation of interest rate shocks, σ_ϵ	0.025
mean income innovation, z	1
Standard deviation of permanent income innovation, σ_z	0.02
Standard deviation of transitory income innovation, σ_ϵ	0.1
Permanent income growth G	1
Standard deviation of measurement error, σ_η	0.03 (50% noise)

Table 3: Parameter Values

information. If we observe income realizations and sometimes there is very low income then we might condition the variance (or higher moments) of the expectations error distribution on that. In particular, if income has a unit root then the current shock has a permanent component and should have a powerful effect on consumption. Similarly, if we think that agents are sometimes liquidity constrained so that the Euler equation does not hold, then we need to observe asset information. In general, the Euler equation will only hold if positive assets (or assets above some debt limit) are carried forward. We can model this in an SRE framework but we leave this for future work.

5 Small sample properties.

In this section we present small sample results on GMM estimation of exact and approximate Euler equations and our Simulated Residual Estimation (SRE) method. We remind the reader that one of the most important conclusions that we take from the recent literature is that the estimation problem here is inherently a small sample one (see the discussion at the end of section 2); hence we do not present any asymptotic results and rely on Monte Carlo simulations alone. We use the same simulation environment as described in section 3. We generate data using a standard life cycle model in which a consumer maximizes expected utility subject to the intertemporal budget constraint. Details are given in the Appendix. Table 3 gives the parameters (the same as for model 9 in table 1) that we use.

For GMM we use continuously updated GMM (see Hansen, Heaton and Yaron (1996)) to remove any dependence on the normalization. For the exact form, we estimate the preference parameters β and γ using the following orthogonality condition on the error term:

$$E_{h,t} \left[\left(\frac{C_{h,t+1}}{C_{h,t}} \right)^{-\gamma} (1 + r_{h,t+1})\beta - 1 \right] = E_{h,t} [\varepsilon_{h,t+1} - 1] = 0 \quad (23)$$

The instruments taken are the constant and lagged real rate¹⁵, Our second empirical model is the approximate Euler equation:

$$\ln \left(\frac{C_{h,t+1}}{C_{h,t}} \right) = \alpha + \frac{1}{\gamma} r_{t+1} + e_{h,t+1}$$

where we use the same instruments used for the exact GMM estimation. Since this is a panel data model with potential correlation across individuals, the weighting matrix should be constructed accordingly. In practice, we found that a flexible weighting matrix that allows for all possible correlations do not work well in terms of convergence.

The final estimator we use is SRE. The errors used to generate simulated consumption paths are assumed to be distributed with a unit mean lognormal distribution and the auxiliary parameters are obtained by estimating the approximate Euler equation.

In our Monte Carlo experiments, we investigate the small sample properties of GMM on the exact Euler Equation, GMM on the first order approximation and SRE, both with and without measurement error. We perform four sets of experiments. We assume that the econometrician has panel data on consumption and estimates the preference parameters by pooling all individuals together. The baseline experiment is for 10 ex-ante identical households followed for 40 periods and no measurement error. The number of replications for all experiments is 10,000. The second set of results increases the number of households to 20, holding the number of time periods constant. The third set of results take the baseline case and reduce the number of time periods to 15. The basic motivation behind this experiment is to establish how well the estimators perform in a situation (fairly realistic) in which we

¹⁵We also experimented with different instruments such as lagged consumption growth and lagged income growth. Results with these instruments (in addition to lagged interest rate) are worse than the results we present here.

Environment				Exact GMM		AGMM	SRE		
	T	N	σ_η	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\sigma}_\eta$
1	40	10	0	.950	3.65	4.72	.954	3.71	.004
				[.950]	[3.55]	[4.29]	[.955]	[3.45]	[.004]
				(.004)	(.637)	(3.19)	(.006)	(1.58)	(.003)
2	40	20	0	.951	3.59	4.70	.953	3.70	.004
				[.950]	[3.54]	[4.37]	[.953]	[3.64]	[.002]
				(.003)	(.427)	(1.33)	(.004)	(.517)	(.001)
3	15	10	0	.948	1.06	1.66	.959	3.36	.007
				[.948]	[.768]	[1.40]	[.957]	[2.65]	[.004]
				(.004)	(1.20)	(13.02)	(.020)	(2.26)	(.004)
4	40	10	.03	.942	2.79	3.41	.950	3.80	.026
				[.946]	[2.55]	[2.84]	[.953]	[2.73]	[.027]
				(.012)	(1.01)	(44.5)	(.039)	(3.04)	(.007)
Notes. True values are $\gamma = 4$ and $\beta = 0.952$ Values are means, in square brackets we present medians and in brackets the standard deviations.									

Table 4: Small Sample Results: Means, Medians and Standard Deviations of Sampling Distributions

have a short panel. Finally, in the fourth experiment we add moderate measurement error to the consumption paths in the baseline model; specifically with our parameter values, half of the observed standard deviation of first differenced log consumption is noise.

Table 4 presents the sampling distributions of the three estimators for our five experiments. In the absence of measurement error and with a long panel (environment 1), GMM using the exact Euler Equation and SRE perform very similarly with both giving estimates of the CRRA that are biased slightly downward. Exact GMM yields a lower standard deviation and a more symmetric sampling distribution. Both perform somewhat better than the approximate GMM estimator. GMM on the approximate equation (AGMM) results in an upward bias in the estimate of the CRRA, with the divergence between the mean and median indicating some very large values. As a by-product of the SMD estimation we obtain the estimates of the standard deviation of the measurement error. Of course, since there is no measurement error, our estimates are necessarily biased but usually close to zero.

Increasing the number of cross-section units resulted in an even lower mean estimate of CRRA for the exact GMM whereas the approximate GMM and the SRE do not show any changes in means. On the other hand the standard deviations of the distributions for all three estimators are substantially lowered. For the third experiment, we see that decreasing the number of time periods from 40 to 15 leads to some substantial changes. First, as expected, the dispersion of all of the estimators rises considerably. The standard deviations obtained from all estimators have gone up with the most dramatic increase for the approximate GMM estimator (from 3.2 to 13). Second, both of the GMM estimators exhibit serious downward bias in the estimates of the CRRA, whereas the SRE seems to be relatively stable (with again a downward bias). This suggests that SRE has small- T properties even in the absence of measurement error.

In the last experiment, we allow for moderate measurement error in the consumption measurement. The first feature of the estimates given in the Table is that measurement error of this order leads to a downward bias in the exact GMM estimator. While the approximate GMM is expected to yield a result similar to the first experiment it is rather surprising that it yields much lower mean and median for the CRRA estimates. Also note that the dispersion of the sampling distribution increases dramatically. This result is particularly disappointing for the approximate model since the approximation is chosen to deal with multiplicative measurement error. The SRE estimator is relatively unaffected by the presence of measurement error, except that the variance of the estimators is somewhat higher. Additionally SRE appears to recover the measurement error variance very precisely.

The conclusion we draw for these Monte Carlo results are that in a very specific context and using the same model for the SRE as in actually generating the data, SRE does at least as well as exact GMM when there is no measurement error and long panels and considerably better if we have a short panel or measurement error. Additionally, SRE always dominates approximate GMM for the estimation of the CRRA.

6 Estimates from the PSID.

6.1 Sample selection.

In this section we present an application of SRE using the Panel Study of Income Dynamics (PSID) from the United States. The survey contains annual information on food at home and food at restaurants. Our sample covers the periods between 1974 and 1987. We exclude households from the poverty sample and households that do not report in all 14 years. We also exclude any households that may be liquidity constrained in any period; we do this using the conventional indicator of having liquid assets equal to at least two months of income. This results in the exclusion of very many households. Then we exclude households with a head aged 65 or over in 1987. Finally, we excluded households with extreme consumption changes from one year to another. Specifically, households whose consumption more than doubled or halved between any years were excluded from the sample. The final sample has a total of 102 households.

Although this is a very small sample, it is ideally suited to our purposes since it is very 'clean' in terms of the model. In future work we shall explore how to incorporate liquidity constrained households, large changes in household structure and large changes in circumstances that lead to significant changes in consumption.

We divide the sample into two education groups since we expect that the level of impatience and aversion to consumption fluctuations may vary across different education groups. Households whose head has less than 12 years of education are labeled as "less educated" (30 households) and those with more than 12 years are labeled as "more educated" (72 households). We allow for variation over time in household size, marital status and number of children. We also allow for variation in the years of education of the head as long as such variation remains within the broad education category we define.

Table 5 presents basic statistics and demographic changes in our sample. Most of our household heads are white and married. Not surprisingly, we see clear demographic changes over the sample period: Number of people living in the same family unit and number of children increased considerably from 1974 to 1987 for the more educated (and relatively younger) group. We observe slightly opposite trend for the less educated group; a slight decrease in number of children living in the household from 1974 to 1987. Finally, majority of our households are male headed. Although we present uncondi-

	More Educated (70.6%)	Less Educated (29.4%)
real food exp.	8974	8944
consumption growth	.018	−.004
head age 1974	30.1	34.6
# of children 1974	.68	1.53
# of children 1987	1.38	1.13
# in family unit 1974	2.58	3.60
# in family unit 1987	3.38	3.43
pct. married 1974	82%	90%
pct. married 1987	86%	93%
pct. black	13%	30%
pct. male head	93%	97%
Numbers (non-percentage ones) are mean values over all households		
and all years in the sample. 1984 prices.		

Table 5: Summary Statistics of PSID sample

tional means of consumption growth for both education group here, in our estimations of Euler equations we will control for these demographic changes.

6.2 Estimation of the homogeneous model.

We begin by estimating a model in which the discount factor and the CRRA are assumed to be homogeneous within each education group. In order to obtain the auxiliary parameters we estimate the usual first order approximation to the consumption Euler equation by OLS. Since the related empirical literature has mostly relied on instrumental variable estimation of the log-linearized Euler equation we also present the IV estimates for comparison purposes. In both cases we have included the first difference of number of children, family size and marital status to control for the change in demographics.

In order to be consistent with our Monte Carlo experiments we also estimate the nonlinear Euler equation in which the demographics enter exponentially. We use twice lagged interest rates and twice lagged consumption growth as instruments for the IV and nonlinear GMM estimations. Table 6 presents the IV estimates and the OLS auxiliary parameter estimates used for the SRE. The estimates are the constant, the coefficient on the real rate, the second to fourth moments of the residuals, the AR parameter in a re-

Estimated parameter	Less educated		More educated	
	OLS	IVE	OLS	IVE
Constant	-.006	-.025	.009	.016
	(.016)	(.027)	(.011)	(.014)
Coefficient on real rate	.14	0.55	.262	.135
	(.470)	(.887)	(.326)	(.461)
std of residuals	.248	.252	.266	.261
skewness of residuals	-.570	-.608	-.412	-.275
kurtosis of residuals	4.94	5.05	4.34	3.83
AR(1) coefficient of residuals	-.003	-.008	-.01	-.01
	(.044)	(.039)	(.029)	(.031)
Correlation btw residuals and Δr_{t+1}	.055	.027	-.006	.001
Nonlinear GMM Results				
CRRA	1.31		.820	
	(2.84)		(2.93)	
EIS	.763		1.22	
	(.982)		(2.55)	
Discount Factor	.908		.924	
	(.322)		(.189)	

Table 6: Auxiliary Parameter Estimates

gression of the residuals on their lagged values and the correlation between the regression residuals and interest rate shocks. We also present results for the exact nonlinear Euler equation. The results for the approximate IVE are typical for this literature. In particular, the coefficient on the real rate is very imprecisely determined with confidence intervals of $[-1.19, 2.29]$ and $[-0.77, 1.04]$ for the less educated and more educated respectively. There appears to be no significant negative auto-correlation in the residuals. Finally the OLS and IV estimates are (statistically) similar, reflecting the weakness of our instruments. The exact GMM results presented at the bottom of Table 6 look plausible although they all suffer from extremely low precision.

For SRE we assume that the expectational errors distribution can be parameterized using a mixture of two lognormal distributions, both with a mean of unity. In order to identify the three parameters of the mixture distribution (the two variances and the mixing probability) we use the standard deviation, skewness and kurtosis of the OLS regression residuals obtained from

the PSID sample. For the measurement error we take a unit mean lognormal distribution and estimate its standard deviation within the SRE procedure. The autoregressive coefficient of regression residuals is used to identify the percentage of noise in the sample. Since we do not observe any significant correlation between the interest rate shocks and regression residuals we disregard it in the estimations. We also confirmed the absence of aggregate shocks by checking the correlation between regression residuals and the time dummies¹⁶. It is important to note that SRE can be modified to incorporate such shocks through, for example, common shocks to permanent income. We leave this for our future work.

Our mixture model has six model parameters (the two preference parameters, the measurement error variance, and the parameters of the mixed expectational error distribution) and six auxiliary parameters. Table 7 presents SRE estimates of the discount factor, the coefficient of risk aversion and the percentage of noise in the consumption growth data for each education group.¹⁷ We present results for both the lognormal and the mixture of lognormal assumptions¹⁸. For each set of estimates we performed a Kolmogorov-Smirnov test to see how well we match the regression error distribution and consumption growth distribution for each education group. As can be seen, the simple log-normal specification for the expectational errors leads to very different estimates of all the parameters. The difference is substantial for the less educated group; estimated discount factor for the single lognormal is a great deal lower than that of the mixture of lognormals. The estimate of the CRRA with the single lognormal is much higher than that of the mixture for the less educated group whereas the result is reversed for the more educated group.

The assumption of the expectational error distribution seems to matter a lot for the measurement error variance estimation. The single lognormal assumption yields much higher noise estimates for both education groups. For the (preferred) mixture model, the discount rate estimates are 5.6% and 4.7% for the less educated and the more educated groups respectively (we present confidence bands below). Given that the average real rate from the sample period is approximately 2.1%, the discount factor estimates suggest

¹⁶The presence of aggregate shocks invalidates the use of time dummies as instruments since they will be correlated with the expectational errors (see Runkle 1991).

¹⁷We do not present standard errors since the model is nonlinear; tests of hypotheses of interest will be given below.

¹⁸For the log-normal model we drop the skewness and kurtosis as auxiliary parameters.

	Less educated		More educated	
Parameters	LN	Mixed	LN	Mixed
Discount factor, β	.694	.947	.998	.955
CRRA, γ	3.67	1.20	1.29	1.95
EIS, $(1/\gamma)$.272	.833	.775	.513
ME std σ_η	.07	.01	.05	.02
(noise)	(28%)	(4%)	(19%)	(8%)
K-S Test: residuals	.13	.81	.25	.88
K-S Test: C. growth	.03	.78	.10	.83
Notes: K-S test are the p-values for Kolmogorov -Smirnov tests for the equality of the auxiliary model's residual and consumption growth distribution (obtained from PSID and that from simulated data.				

Table 7: SRE Parameter Estimates

a fairly high degree of impatience for both education groups. Turning to the results for EIS, we note that there is a substantial difference between the IV and the SRE estimates. Also note that the difference is not as big between SRE and nonlinear GMM estimates for the less educated group. The most important substantive finding is that the less educated have a lower CRRA. We turn now to the precision of the estimates of the preference parameters.

To conduct inference on the parameter estimates of interest we note that the standard asymptotic Wald tests are inappropriate for nonlinear models. Therefore we adopt a quasi-likelihood ratio test procedure and perform tests of ‘significance’ for the discount factor and CRRA parameters separately. To do this, we define a grid for the parameter in hand and perform SRE for each point on the grid. For example, we fix the discount factor at different values from 0.1 to 1.1 and then estimate the CRRA (and the other parameters) by SRE at each of these points. Since this gives an equation with one degree of over-identification, the appropriate weighting matrix for the criterion should be used. This matrix is the inverse of the covariance matrix of the auxiliary parameters. We obtain this using a nonparametric bootstrap on the original PSID data, in which we re-sample households (that is, we re-sample consumption paths of length 14). Using the inverse of this covariance matrix as the weighting matrix, the minimized function value has a $\chi^2(1)$ distribution, under the null that the parameter is equal to that value.

The results for these tests for the two preference parameters are presented in the top panel of Figure 1. Note that the $\chi^2(1)$ statistics are zero at the point estimates; the horizontal line gives the 5% cut-off. Both figures clearly suggest that the distribution of the estimates are decidedly asymmetric for both education groups (confirming the invalidity of using conventional standard errors). The discount factor estimate is more precise for the less educated group relative to the more educated. The lower confidence band for the more educated is very wide whereas upper band is fairly tight. For both education groups, a point estimate of unity for the discount factor is rejected. The 95% confidence interval for the discount factors are [0.75, 0.97] and [0.19, 0.99] for the less and more educated respectively.

For the CRRA estimates the most important finding is that we decisively reject values below unity or very high values for the more educated. The confidence intervals for the CRRA are [0.87, 2.95] and [1.42, 3.70] for the less and more educated respectively. This translates into intervals of [0.34, 1.15] and [0.27, 0.70] for the EIS which are much tighter than for the GMM estimators. This reinforces our earlier analysis that SRE delivers more precise estimates than the linearized and the nonlinear model on the same (noisy) data.

The results presented here are encouraging. We find that the discount factor can be estimated with some precision, even though we have very small sample sizes. The point estimate for the more educated is higher than that for the less educated, which accords with widespread priors. We find that both groups have a higher discount rate than the mean real rate in our data so that both groups display impatience. We also find reasonable estimates for the CRRA with the more educated displaying more aversion to fluctuations (more risk aversion). This could be taken as evidence against the hypothesis that the EIS is independent of the level of consumption (the iso-elastic form). All of this is for the model with the same parameters for everyone within an education group. We now go on to consider whether the discount rate is the same for everyone with the same education.

6.3 Heterogeneity in the discount factor.

Of all the features that empirical analysis using micro data has to address, heterogeneity is the most important. In this section we present a method to identify the heterogeneity in the discount factor within each education group. We choose to concentrate on heterogeneity in the discount rate rather than

the CRRA since that has been the principal focus in the previous literature; see, for example, Carroll and Samwick (1997). Our approach to identifying the distribution of discount factors begins with the observation that there are persistent differences between households in their consumption growth. To show this we take means over time of consumption growth for each household.

In Figure 3 we present the distributions of mean consumption growth for our two education groups. Two features of these distributions merit attention. First, the distribution for the more educated is to the right of that for the less educated. This is reflected in the higher discount factor found for the former in the homogeneous case. Second, within each education group there is significant heterogeneity. For example, for the more educated the mean consumption growth is about 0.01 (-0.004 for the less educated group) but some households have an average consumption growth of more than 0.1 per year (so that consumption in the final year is four times that of the initial year) and others have almost -0.1 (consumption in the final year is one fifth that of the initial year). One reason for these differences is that different households have different realizations of the expectations errors and some have persistently pleasant shocks. In the SRE modelling this is captured by our use of simulated residuals, with some simulated households having long runs of good or bad draws. The other possible source of variation, if we assume that everyone has the same CRRA and the same measurement error structure, is differences in the discount factor. That is, more patient households have higher expected consumption growth.

To implement an estimator allowing for heterogeneity we first assume that the discount factor is normally distributed across the population with mean μ_β and standard deviation σ_β .¹⁹ In the simulation model we draw values of this, β_h , for each h and then construct household specific paths using the analogue of equation (20) and simulated residuals $\varepsilon_{h,t+1}^S$:

$$\frac{C_{h,t+1}^S}{C_{h,t}^S} = \left\{ \frac{1}{\beta_h(1+r_{t+1})} \varepsilon_{h,t+1}^S \right\}^{-\frac{1}{\gamma}} \quad (24)$$

To estimate the additional parameter σ_β , we require one more auxiliary parameter. The obvious candidate is some measure of the dispersion of the

¹⁹ An alternative would be to impose that the discount factor is below unity by taking a distribution with support $[0, 1]$ (the Beta is an obvious candidate). We choose not to do this since previous investigators have found evidence that some households have discount factors above unity.

time averages of consumption growth discussed in the previous paragraph. Specifically, we include a household specific fixed effect in the OLS regression of consumption growth on the real rate and take the standard deviation of these fixed effects as our new auxiliary parameter.

We begin by testing whether the homogenous model is rejected by the data. To do this we take the just identified seven parameter heterogeneity model and set $\sigma_\beta = 0$. The $\chi^2(1)$ statistics for homogeneity are 32 and 47 for the less educated and more educated respectively. Thus we decisively reject the homogeneity assumption. Given this rejection we go on to estimate the extent of the heterogeneity. In Figure 3 we present the results for a grid search over the standard deviation of the discount factor for each of the education groups. The horizontal line gives the 5% cut-off for the standard deviation. The estimation results are presented in Table 8. The principal

Parameters	Less educated	More education
Mean discount factor, μ_β	.87	.95
Std of discount factor, σ_β	.101	.086
CRRA, γ	1.68	2.03
EIS, $(1/\gamma)$.595	.493
ME std, σ_η (noise)	.03(11%)	.04(16%)

Table 8: Heterogeneity SRE estimates

features of these results are: the mean discount rate is very similar to the homogenous model for the more educated group; there is a significant dispersion of the discount rate with 95% of less educated households between 0.05 and 1.15 and 95% of more educated households between 0.79 and 1.12; there is no significant impact of allowing for heterogeneity on the estimates of the CRRA and the measurement error variance. The results indicate a fairly wide distribution for discount factors, even allowing for education. In future work we shall explore how this distribution correlates with observable fixed factors.

7 Conclusions.

There is widespread agreement that, given currently available data, we cannot accurately estimate the parameters of intertemporal allocation using

GMM on exact or approximate Euler equations. Our reading of this literature and our own results is that this is a small sample (strictly, short panel) problem. The alternative seems to be to move to full structural modelling. In the current state of the art this is cumbersome, fragile and unable to deal with significant heterogeneity. To circumvent these problems, we present a novel estimation procedure that combines some of the advantages of the Euler equation and structural modelling approaches. This procedure is based on simulating expectation errors; we refer to it as Simulated Residual Estimation (SRE). The principal advantage of SRE is that it allows us to estimate preference parameters without having to specify the underlying economic environment explicitly. We develop variants of the basic procedure that allow us to take account of measurement error in consumption, the ‘news’ in interest rate realizations and heterogeneity in discount factors.

A Monte Carlo investigation of the small sample properties of the SRE estimator indicates that it dominates GMM estimation of both exact and approximate Euler equations in the case when we have short panels and noisy consumption data. To complement the Monte Carlo results, we present an illustrative empirical application to two samples drawn from the PSID. The results are very encouraging even though we have small sample sizes. We find that we can estimate the parameters of intertemporal allocation much more precisely than with a conventional GMM on a log-linearized model. For example, we find that the 95% confidence interval for the EIS is $[0.27, 0.70]$ for the more educated whereas the linearized GMM confidence interval is $[-0.38, 0.90]$. Moreover, the parameter estimates seem quite reasonable. For example, we find discount factors that are less than, but close to unity, with a higher discount factor for the more educated group. We also find that the more educated have a higher CRRA. Finally, we present evidence that there is a significant heterogeneity in the discount factor, in both the statistical and substantive sense.

SRE relies on specifying the conditional distribution of the expectations errors. All of the above only uses consumption and interest rate information. The strength that our method shares with the Euler equation approach is that we do not have to specify the income process nor the processes for other relevant variables. Generally, however, one would expect to achieve more identification and better precision by using the observed income series for each household. Similarly, if we think that agents are sometimes liquidity constrained so that the Euler equation does not hold, then we need to observe asset information. In general, the Euler equation will only hold if positive

assets (or assets above some debt limit) are carried forward. We can model this in an SRE framework but we leave this for future work.

There are a number of further avenues to explore. One of these is to allow for conditionally heteroscedastic expectations errors in the estimation step; that is, allowing that the marginal utility of expenditure is a martingale and not a random walk. One particularly important facet of this is to allow for persistent heterogeneity in expectations error variances across agents. This will require the use of income and asset information. For example, low levels of beginning of period assets will be associated with high expectations error variances. As another example, we showed in section 2 that the distributional assumption we make (a mixture of two log normals) is a poor one if there are sometimes very low income realizations. If we observe income realizations then we can condition the variance (or higher moments) of the expectations error distribution on that. We also plan to use cross-section differences in interest rates due to differences in marginal tax rates and the wedge between borrowing and lending rates. The ultimate goal of this analysis will be the development of credible estimates of the variation in preference parameters across agents. Based on this we can then address whether cross-section variations in the EIS are due to (persistent) heterogeneity or to differences in wealth levels. This will lead on to a systematic exploration of alternative forms for the utility function (including the vexed question of whether the iso-elastic assumption is tenable). In the empirical illustration presented in this paper we have largely followed the literature in our modelling assumptions so as to highlight the SRE procedure. In future work on the PSID we shall present analyses based on larger samples with unbalanced panels and some agents only being in the sample for a short period. It will also be important to take coherent account of the fact that food is a sub-component of total expenditure and also to take more careful account of changes in the demographic composition of the household.

To conclude: Simulated Residual Estimation provides a procedure for estimating the parameters of intertemporal allocation without the need for full structural modelling. In estimation we can allow for 'classical' measurement error in consumption which leads to a good deal of bias in exact Euler equation GMM estimation. Furthermore the SRE procedure is flexible enough to allow us to consider a number of extensions to the conventional model. In this paper we have only considered allowing for heterogeneity in discount factors; in future work we plan to develop some of the other issues discussed above.

References

- Altonji, J. G. and A. Siow, (1987) "Testing the Response of Consumption to Income Changes with Noisy Panel Data", *Quarterly Journal of Economics*, 102, 293-328.
- Alvarez, Javier, Martin Browning and Mette Ejrnæs, (2002), "Modelling income processes with lots of heterogeneity", CAM working paper 2002-01, University of Copenhagen.
- Amemiya, Y. (1985), "Instrumental variable estimator for the nonlinear errors-in-variables model", *Journal of Econometrics*, 28(3), 273-289
- Attanasio, O. P. , J. Banks, C. Meghir, G. Weber, (1999) "Humps and Bumps in Lifetime Consumption", *Journal of Business and Economic Statistics*, 17(1), 22-35.
- Attanasio, O. P. and H. Low, (2002) "Estimating Euler Equations", mimeo, University College London.
- Blundell, R. and R. Smith (1991), "Initial conditions and efficient estimation in dynamic panel data models", *Annales d'Economie et de Statistique*, 20/21, 109-123.
- Browning, M., M. Ejrnæs, (2001) "Consumption and Children", CAM working paper 2002-06, University of Copenhagen.
- Browning, M. and A. Lusardi, (1996) "Household Saving: Micro Theories and Micro Facts", *Journal of Economic Literature*, December, 1797-1855.
- Carroll, C., (2001) "Death to the Log-Linearized Consumption Euler Equation!", *Advances in Macroeconomics*, Forthcoming.
- Carroll, C. and A. Samwick, (1997) "The Nature of Precautionary Wealth", *Journal of Monetary Economics*, 40, 41-71.
- Chamberlain, G. (1980), "Analysis of Covariance with Qualitative Data", *Review of Economic Studies*, 47, 225-238.
- Deaton, A., (1991) "Saving and Liquidity Constraints", *Econometrica*, 59(5), 1221-1248.
- Dynan K. E., (1993) "How Prudent Are Consumers?", *Journal of Political Economy*, vol. 101, no:6, 1104-1113.
- Gallant, A. R. and G. Tauchen, (1996) "Which Moments to Match?", *Econometric Theory*, 12:657-681.
- Gourieroux, C., A. Monfort and E. Renault, (1993) "Indirect Inference", *Journal of Applied Econometrics*, 1993, 8:S85-S118.
- Gourinchas, P.O. and J. A. Parker, (2001) "Consumption Over the Life Cycle", *Econometrica*, Forthcoming.

Hall, R. E., (1978) “Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence”, *Journal of Political Economy*, December, vol. 86, 971-987.

Hall, G. and J. Rust, (2002) “Econometric Methods for Endogenously Sampled Time Series: The Case of Commodity Price Speculation in the Steel Market”, Cowles Discussion Paper 1376, University of Yale.

Hansen, L .P. and K. J. Singleton, (1982) “Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models”, *Econometrica*, 50(5), 1269-1286.

Hansen, L. P., J. Heaton and A. Yaron, (1996) “Finite Sample Properties of Some Alternative GMM Estimators”, *Journal of Business and Economic Statistics*, 14(3):262-280.

Hubbard, R. G., J. Skinner and S. P. Zeldes, (1995) “Precautionary Saving and Social Insurance”, *Journal of Political Economy*, vol. 103. no:2, 360-399.

Ludvigson, S. and C. Paxson, (2001) “Approximation Bias in Linearized Euler Equations”, *Journal of Economics and Statistics*, vol. 83, no:2, 242-256.

Runkle, D. E., (1991) “Liquidity Constraints and The Permanent Income Hypothesis”, *Journal of Monetary Economics*, 27, 73-98.

Shapiro, M. D. (1984), “The Permanent Income Hypothesis and the Real Interest Rate”, *Economics Letters*, 14, 93-100.

Wooldridge, J. (2000), ”The initial conditions problem in dynamic non-linear panel data models with unobserved heterogeneity”, mimeo, Michigan State University.

A The consumption function

We assume the utility function is intertemporally additive and the sub-utilities are iso-elastic (with no durables and no demographics). The problem of the generic consumer h at time t is:

$$\begin{aligned} \max E_t \left[\sum_{j=0}^{T-t} \frac{(C_{h,t+j})^{1-\gamma}}{1-\gamma} \frac{1}{(1+\delta)^j} \right] \\ \text{s.t. } A_{h,t+j+1} = (1+r_{h,t+j})A_{h,t+j} + Y_{h,t+j} - C_{h,t+j} \end{aligned}$$

where C is non-durable consumption, A is assets, Y is stochastic labor income and r is the stochastic real interest rate. We assume finite life and end of life T is certain. The discount rate δ and the coefficient of risk aversion γ are positive. Our generic consumer has no bequest motive so that $A_{T+1} = 0$. The stochastic process driving labor income is taken to be that described in equation (11). We assume that the innovations to income are independent over time and across individuals i.e. we assume away aggregate shocks to income. Individuals can use only one asset to smooth their consumption against these idiosyncratic income shocks. The return on this asset (interest rate) is generated by a stationary AR(1) process:

$$r_{h,t+1} = (1-\rho)\mu + \rho r_{h,t} + \epsilon_{h,t+1} \quad (25)$$

where μ is the unconditional mean, ρ is AR(1) coefficient with $0 < \rho < 1$, and ϵ_{t+1} is assumed to be *iid* Normal with mean zero and standard deviation σ_ϵ .

Following Deaton(1991), the budget constraint is re-defined as

$$X_{h,t+j+1} = (1+r_{h,t+j+1})(X_{h,t+j} - C_{h,t+j}) + Y_{h,t+j+1} \quad (26)$$

where $X_{h,t+j} = A_{h,t+j} + Y_{h,t+j}$ (cash on hand). The income process is nonstationary which makes the problem harder to solve since the range of possible income values is large. Instead, we redefine all the relevant variables in terms of their ratios to current income and solve for the consumption to income ratio. By doing this we reduced the number of state variables to two, namely the cash on hand to income ratio and the interest rate. Moreover, we obtain an iid income process which can be approximated by standard Quadrature

methods. Given this redefinition of the relevant variables, the Euler equation can be written as

$$\theta_t(w_t, r_t)^{-\gamma} - \frac{1}{(1 + \delta)} E_t [(1 + r_{t+1})\theta_{t+1}(w_{t+1}, r_{t+1})^{-\gamma} z_{t+1}^{-\gamma}] = 0 \quad (27)$$

where $\theta_t = \frac{C_t}{Y_t}$, $w_t = \frac{X_t}{Y_t}$.

The problem is solved via policy function iteration using the terminal value condition. At the terminal date T , consumption is function of only cash on hand and since the bequest motive is assumed away $\theta_T = w_T$.

For the income process, we use a 10 point Gaussian Quadrature and following Tauchen (1986) we approximate the interest rate process by forming a 10 point first order discrete Markov process. We use a cubic spline to approximate the consumption function at each iteration. Since we solve a finite life problem, we obtain T consumption-to-income ratio functions $\{\theta_1(w_1, r_1), \dots, \theta_T(w_T)\}$.

Table 3 reports the parameter values used in the solution and the simulation of the model described above. The agent is allowed to borrow the amount he can pay back with certainty. In the infinite life case this would correspond to the borrowing limit of $\frac{\min Y}{\max r}$. The discount rate and the mean interest rate are chosen to be equal in order to prevent consumers to quickly go towards the borrowing constraint. When the discount rate is large relative to the interest rate, consumers borrow close to the maximum possible amount. Then the movement of consumption is largely driven by income and the identification of interest rate impact on consumption growth becomes very difficult.

We initialize the algorithm with the consumption rule at the end of life $c_T(x_T) = x_T$. The constraint on borrowing is that at the end of the life person should pay back all his outstanding debt. In practice this constraint will never bind since the utility function satisfies the Inada conditions which implies that zero consumption is never chosen. Instead we will observe very impatient individuals getting very close to the borrowing limit, whereas it will be irrelevant for the patient ones. Since we do not assume an explicit borrowing limit as in Deaton (1991), the consumption functions are continuously differentiable. In fact, in our case where agents have iso-elastic preferences and income uncertainty, consumption functions are strictly concave. In order to solve the problem, we define an exogenous grid for the cash on hand to income ratio: $\{x_j\}_{j=1}^J$. It is important to adjust the grid as the

solution goes back in time. The algorithm finds the consumption that makes the standard Euler equation hold for each value of x and r . In practice, we took 500 points for x and 10 points for r . After obtaining c_{T-1} , we use a cubic spline to approximate $c_{T-1}(x_{T-1})$ for each r . After obtaining the consumption functions for each age, we simulate life time consumption paths using the intertemporal budget constraint and generating random draws for income and interest rate. Generated paths differ due to different realizations of income and interest rates for each individual.

For our Monte Carlo experiments we generate 80 period consumption paths for ex ante identical consumers. Individuals are assumed to face the same interest rate series. Therefore individuals' consumption paths differ due only to different income realizations. Although it is possible to allow for cross section variation in the interest rate we believe it is more plausible to assume only time series variation (to take proper account for correlation across individuals' expectational errors). We perform several experiments by using different time period (T) and number of individuals (N) in the estimations.

B Simulating the lognormal distribution.

Before presenting the full optimization algorithm, we have to digress a little and discuss how to simulate draws from a lognormal distribution with a mean of unity. In the optimization routine for any simulation estimator it is important to keep the draws constant from iteration to iteration, otherwise the optimization routine becomes unstable. We can simulate a lognormal by taking:

$$X \sim \exp(a + bN(0, 1))$$

where $N(0, 1)$ denotes the standard Normal. The mean and variance of X are given by:

$$\begin{aligned} \mu_X &= \exp(a) \sqrt{\exp(b^2)} \\ \sigma_X^2 &= \exp(2a) \exp(b^2) (\exp(b^2) - 1) \end{aligned}$$

To ensure that the mean is unity we need to impose:

$$a = -\frac{b^2}{2}$$

Thus if we simulate draws from a lognormal with mean 1 and a standard deviation of σ_X we use:

$$X \sim \exp\left(-\frac{\ln(1 + \sigma_X^2)}{2} + \sqrt{\ln(1 + \sigma_X^2)}N(0, 1)\right)$$

In the algorithm below, the procedure is to draw a matrix vector of standardized Normal variables and then to use this formula to give a lognormal with unit mean and varying standard deviation.

Figure 1: Inference on Discount Factor and CRRA (PSID)

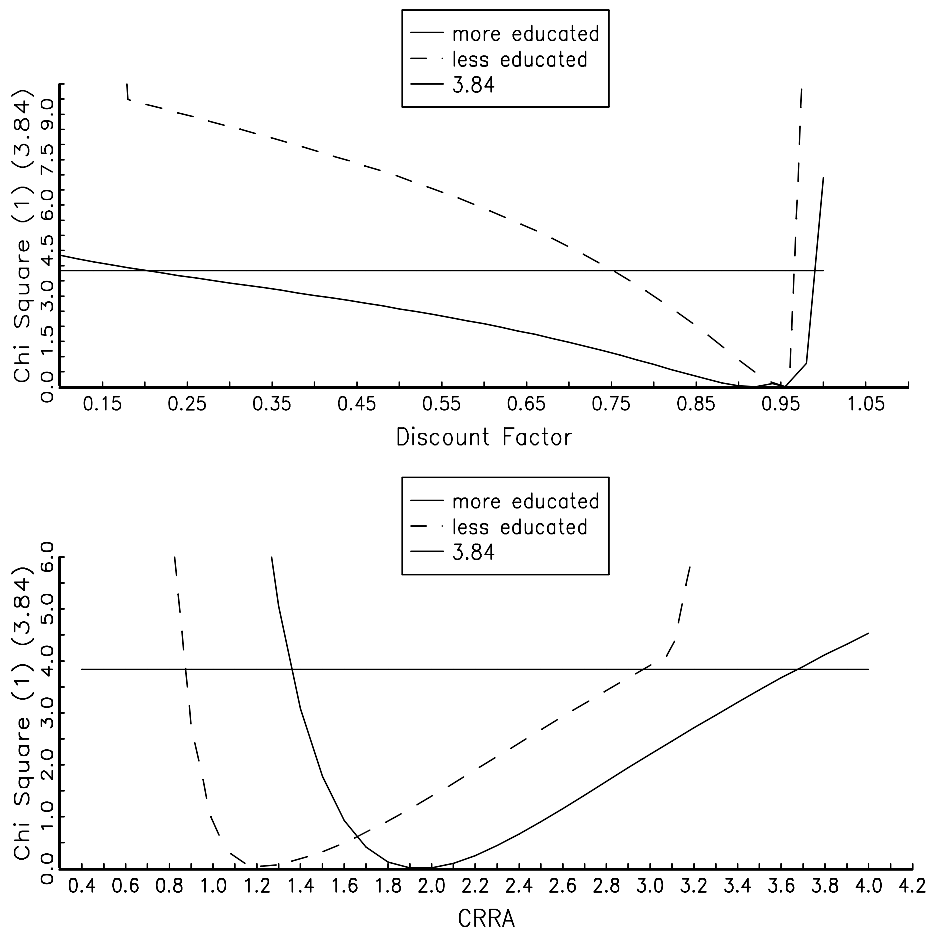


Figure 2: Distribution of Mean Consumption Growth (PSID)

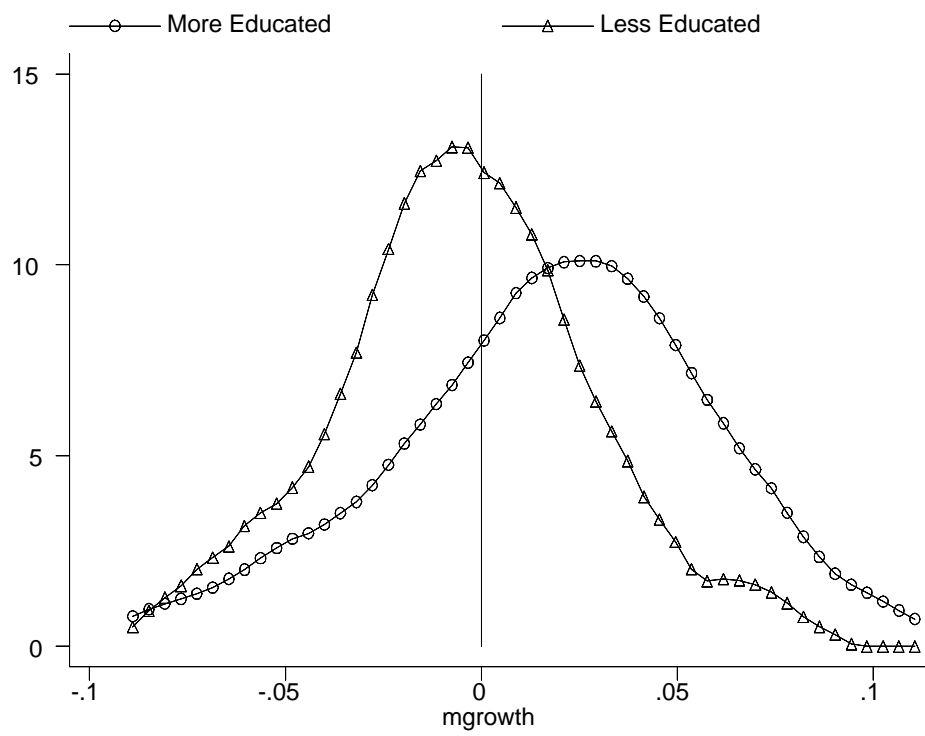


Figure 3: Test of Heterogeneity in Discount Factor (PSID)

