EULER EQUATION ESTIMATION ON MICRO DATA

SULE ALAN
University of Essex

KADIR ATALAY
University of Sydney

THOMAS F. CROSSLEY
University of Essex
and
Institute for Fiscal Studies

Consumption Euler equations are important tools in empirical macroeconomics. When estimated on micro data, they are typically linearized, so standard IV or GMM methods can be employed to deal with the measurement error that is endemic to survey data. However, linearization, in turn, may induce serious approximation bias. We numerically solve and simulate six different life-cycle models, and then use the simulated data as the basis for a series of Monte Carlo experiments in which we evaluate the performance of linearized Euler equation estimation. We sample from the simulated data in ways that mimic realistic data structures. The linearized Euler equation leads to biased estimates of the EIS, but that bias is modest when there is a sufficient time dimension to the data, and sufficient variation in interest rates. However, a sufficient time dimension can only realistically be achieved with a synthetic cohort. Estimates from synthetic cohorts of sufficient length, while often exhibiting small mean bias, are quite imprecise. We also show that in all data structures, estimates are less precise in impatient models.

Keywords: Euler Equations, Measurement Error, Instrumental Variables, GMM

1. INTRODUCTION

Since Hall (1978), Euler equations, which are first-order conditions from dynamic optimization problems, have been used extensively to test these models against economic data and as a basis for estimating preference and technology parameters. This tradition includes studies of consumer behavior [see for example, Attanasio

 3267

et al. (1999)]; studies of the investment behavior of firms [for example, Bond and Meghir (1994) or Mulligan (2004)]; and tests of asset pricing theories [for example, Mehra and Prescott (1985)]. The main attraction of the Euler equation approach is that it allows researchers to estimate preference parameters with limited data and without fully specifying the stochastic processes that agents face. For example, in principle, it is not necessary to model consumers' expectations or observe their wealth when estimating the elasticity of inter-temporal substitution (EIS) and the degree of relative prudence via an Euler equation approach. Estimated values of these parameters are central to our thinking about macroeconomic policies such as the UK's temporary VAT cut during the crisis of 2008–9 [Crossley et al. (2009)].

Unfortunately, the advantages of the Euler equation approach may be significantly diminished by practical problems that arise from the nature of the available data. In particular, when estimated on micro data, Euler equations are typically linearized so standard IV or GMM methods can be employed to deal with the measurement error that is endemic to survey data.[1] However, it may be that the approximation bias induced by linearization is as bad as the problems that linearization is intended to solve. Carroll (2001) investigates the estimation of the EIS in an environment with a cross-sectional variation in interest rate and impatient agents. He concludes that cross-sectional estimation of linearized Euler equation is not useful. Ludvigson and Paxson (2001) investigate the estimation of the relative prudence parameter in an environment with a fixed interest rate and reach a similarly negative conclusion. In contrast, Attanasio and Low (2004) argue that with sufficiently long sample periods and enough time-series variation in the inter-temporal price (interest rate), good estimates of the EIS can be obtained with linearized Euler equations. More than a decade after these papers appeared, the issue seems unresolved. Some researchers cite Attanasio and Low to support the continued use of linearized Euler equations (or parameters obtained from them), while others cite Carroll to justify avoiding the estimation linearized Euler equations, or employing parameter estimates obtained from them.[2]

The goal of this paper is to revisit this debate, and move it toward resolution. We numerically solve and simulate six different life-cycle models, and then use the simulated data as the basis for a series of Monte Carlo experiments in which we evaluate linearized Euler equation estimation. Relative to the papers cited about, we extend the available evidence in two key ways. First, we consider a wider range of economic environments (models and parameterizations). Second, we consider realistic data structures. Unlike any papers cited above, we add plausible measurement error to our simulated data, and we sample from the simulated data in ways that mimic available empirical methods (such as using repeated cross-sections to construct a synthetic cohort).[3] We also offer several methodological improvements on these previous studies. We explicitly test the validity and relevance of available instruments, and we consider the full empirical distribution of estimates that emerges from our Monte Carlo experiments, which again gives researchers a more complete guide to the likely performance of these methods.[4] We are able to replicate Attanasio and Low's finding that, with sufficiently long

time series, and sufficient variation in the interest rate, linearized Euler equation estimation can work reasonably well. But more broadly, our findings support Carroll's strong scepticism about linearized Euler equation estimation. We find that the performance of linearized Euler equation estimation declines, with both greater mean bias and a loss of precision, when agents are moderately impatient. A sufficient time dimension can only realistically be achieved with a synthetic cohort. Estimates from synthetic cohorts of sufficient length, while often exhibiting small mean bias, are quite imprecise. Furthermore, our examination of empirical distributions of estimates shows that extreme values arise in all environments. These problems become more prominent in short panels—or when agents face potential credit constraints.

The remainder of the paper is organized as follows. Section 2 reviews the problems associated with linearized Euler equation estimation. Section 3 introduces the six different life-cycle models we study and describes the design of our Monte Carlo studies, and in particular, the different data structures we consider. Section 4 summarizes the problems associated with estimation of different life-cycle models studied. Section 5 presents the results of our Monte Carlo experiments. Section 6 concludes.

2. THE ECONOMETRICS OF EULER EQUATION APPROXIMATION

Consider a standard life-cycle model in which the consumer consumes a single good, has time-separable preferences and holds long and possibly short positions of a single asset. The first-order condition from this problem is[5]

$$U^{'}(C_t) = \beta E_t[(1 + R_{t+1})U^{'}(C_{t+1})], \tag{1}$$

where $U^{'}$ is the marginal utility of consumption, β is the discount factor, and R_{t+1} is the real rate between periods t and $t + 1$. A widely used functional form for the sub-utility function is the isoelastic form:

$$U(C_t) = \frac{C_t^{(1-\gamma)}}{1 - \gamma}, \tag{2}$$

where the parameter γ is the coefficient of relative risk aversion. Interest usually centers on the reciprocal of this parameter, $(1/\gamma)$, the EIS, or on the coefficient of relative prudence, which is $\frac{\gamma+1}{2}$.

Substituting this utility function into equation (1) yields an exact Euler equation:

$$\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} (1 + R_{t+1})\beta = \varepsilon_{t+1} \tag{3}$$

with $E_t(\varepsilon_{t+1}) = 1$, where ε_{t+1} represents the expectation error (the innovation in discounted marginal utility), which the theory implies is orthogonal to variables in the information set at time t. This relationship is the basis of very many estimates of the preference parameters (β, γ) and tests for the validity of the orthogonality

conditions implied by the theory. GMM estimation is based on the orthogonality of the error term 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 specifying the stochastic environment that agents face.

Nonlinear GMM estimation on micro data is inconsistent if the consumption data are measured with error. For example, if we allow for a multiplicative measurement error so that observed consumption is given by

$$C_t^0 = C_t \eta_t, \tag{4}$$

then the exact Euler equation for observable consumption becomes

$$\left(\frac{C_{t+1}^0}{C_t^0}\right)^{-\gamma} (1 + R_{t+1})\beta = \left(\frac{\eta_{t+1}}{\eta_t}\right)^{-\gamma} \varepsilon_{t+1}. \tag{5}$$

The problem is that the composite error term does not have a conditional expectation of unity, even if we assume that ϵ_t and η_t are independent,[6]

$$E_t\left[\left(\frac{\eta_{t+1}}{\eta_t}\right)^{-\gamma} \varepsilon_{t+1}\right] = E_t\left(\frac{\eta_{t+1}}{\eta_t}\right)^{-\gamma} E_t(\varepsilon_{t+1}) = E_t\left(\frac{\eta_{t+1}}{\eta_t}\right)^{-\gamma} \neq 1. \tag{6}$$

It is now widely accepted that household-level consumption data is likely to be very noisy. For example, Runkle (1991) estimates 76% of the variation in the growth rate of food consumption in the PSID is due to measurement error, Shapiro (1984) arrives at an even higher estimate of 92% noise. Using a procedure that allows for preference heterogeneity, Alan and Browning (2010) obtain an estimate of 86%. Dynan (1993) reports 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 sources are quasi-panels, constructed from cross-section expenditure survey information by taking within-period means following the birth cohorts through time (e.g. taking means over all the 25-year olds in 1 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 that acts very much like measurement error [Deaton (1985)].

One way to deal with measurement error problem is to linearize the equation (3) and use standard linear IV and GMM techniques. In particular, the convention is to assume measurement error in consumption is multiplicative. Naturally, log-linearization will move such measurement error into the additive residual. Measurement error results in residuals with a MA(1) structure, but variables lagged twice (or more) are still valid instruments. Following the steps in Carroll (2001),

we get

$$\Delta \log C_{h,t+1} = \alpha + \frac{1}{\gamma} \log(1 + R_{t+1}) + e_{h,t+1} \qquad (7)$$

for the first-order approximation. The constant term α contains the discount factor (β) and the means of the higher order moments consumption growth and interest rates (for the first-order approximation, this would be the second and higher moments). The residual term contains (i) the true innovation in marginal utility (or expectation error) between t and $t+1$, ε_{t+1}, (ii) the measurement errors at t and $t+1$, and (iii) an approximation error composed of variation in the higher moments of consumption growth and interest rates (conditional on past information).

3. RESEARCH DESIGN

In order to investigate the problems associated with the estimation of approximate Euler equations, we compare six different life-cycle models of consumption which correspond to different economic environments studied in the literature. In all six models, preferences are time-separable and within period utility is isoelastic with the coefficient of relative risk aversion set to 4 (so that the EIS is 0.25).

Agents face two types of income shocks, permanent and transitory. The income process of agent h is

$$Y_{h,t} = P_{h,t} \varpi_{h,t}, \qquad (8)$$

where $\varpi_{h,t}$ is an i.i.d. lognormal transitory shock with unit mean and a constant variance $e^{\sigma_{\varpi}^2} - 1$ and $P_{h,t}$ is permanent income which follows a log random walk process:

$$P_{h,t} = G P_{h,t-1} Z_{h,t}, \qquad (9)$$

where $Z_{h,t}$ is an i.i.d. lognormal permanent shock with unit mean and a constant variance $e^{\sigma_z^2} - 1$. In our simulations, we set σ_{ϖ} to 0.1 and σ_z to 0.05, these values are in line with those used in the literature [they are identical to those used in Attanasio and Low (2004)] and experiments with other values give similar results. We assume that the innovations to income are independent over time and across individuals so that we abstract from aggregate shocks to income. However, there are aggregate shocks in these environments because realizations of the real interest rate are assumed to be the same across agents. The real interest rate follows an AR(1) process with a mean of 0.03, an AR parameter of 0.6, and a standard deviation of the error of 0.025.[7] In some models, we augment the income process just described with the possibility of a zero income realization (details given below).

Table 1 presents the parameter values assumed for the six models. The models differ by degree of impatience (since we are ignoring income growth, the degree of impatience is determined by the difference between real interest rate and individual's discount rate), by the possibility of a zero income realization, and by the type of borrowing constraint (either the natural borrowing constraint—the agent

TABLE 1. Parameter values

Parameter	Value
Coefficient of risk aversion (γ)	4, 6
Discount rate [$\delta = (\frac{1}{\beta}) - 1$]	0.03 and 0.07
Standard deviation of permanent income shocks (σ_z)	0.05
Standard deviation of transitory income shocks (σ_u)	0.1
Unconditional mean of interest rate process (μ)	0.03
AR(1) coefficient of interest rate process (ρ)	0.6
Standard deviation of interest rate process (σ_ε)	0.025
Probability of zero income (Models 3 and 4)	0.01

TABLE 2. Models

Model	Impatient	Patient	Borrowing constraint	Zero income probability
AL-P	–	Yes	Life time	No
AL-I	Yes	–	Life time	No
C-P	–	Yes	Life time	Yes
C-I	Yes	–	Life time	Yes
D-P	–	Yes	Explicit	No
D-I	Yes	–	Explicit	No

can only borrow what she can pay back with certainty—or an explicit, period-by-period constraint.) Table 2 summarizes the distinguishing features of the six models. Model AL-P is similar to the environment studied by Attanasio and Low (2004). Agents' discount rates are equal to the mean real interest rate (0.03) and there is only the natural borrowing constraint. The important feature of this model is that even though borrowing is allowed up to the natural limit, individuals do not borrow because they are quite patient and so have a strong taste for accumulation. The second model, AL-I, is an impatient version of the first model. In this model, the discount rate of agents set to 0.07. As a result, agents in AL-I borrow especially early in life.-341pt]All right to change enclosures to conform with the {[(...)]} convention?

Model C-P and C-I are motivated by Carroll (2001). In these models, we augment the income process described above by allowing transitory income shocks to take a "0" value in any given period with small probability (specifically, with a probability 0.01).[8] This addition to the model strengthens agents' precautionary motive. Moreover, this assumption, along with backward induction and the fact that marginal utility of zero consumption is infinite means that agents will not borrow. The resulting consumption functions are very steep at low wealth levels. C-I is very close to the "buffer-stock" model studied by Carroll; C-P retains the income process from Carroll's study but assumes more patient agents.

Finally, we examine two versions of the environment first proposed by Deaton (1991). In these models, individuals are explicitly prevented from borrowing.

This assumption (with a lower bound for labor income) leads to a kink in the consumption function. We have two motivations for including these models. First, households with zero assets are observed in real data and these models (particularly the impatient version) can replicate that fact, while the models described above do not. Second, a comparison of D-P and AL-P will help clarify our arguments below regarding the key features of the economic environment.

We solve these six different life-cycle models using standard methods. Further details of our specification, solution, and simulation of the life-cycle models are given in the Online Appendix A. After solving for the consumption function of a generic household for 60 periods for each model, we simulate 60-period consumption paths for 10,000 ex-ante identical agents (households) and discard the first 10 and last 10 periods to minimize the starting and end effects.[9] We then use the resulting simulated data (for each model) as the basis of our Monte Carlo experiments.

Our Monte Carlo experiments draw repeatedly from the simulated population of agents (i.e., consumption paths) described above. This is done 999 times [Davidson and MacKinnon (2004)]. In each case, a sample of 1,000 agents is drawn from the population of 10,000 with replacement. However, we mimic three different data structures. The first is a long panel. Each of the 1,000 agents is followed for 40 periods. As emphasized by Chamberlain (1984) and others [see Attanasio and Low (2004)], the orthogonality conditions implied by theory hold with long T. Thus, a long panel is a best case scenario for Euler equation estimation. However, typical lengths of household panels are much shorter, and very long panels will inevitably suffer from attrition and other problems. We therefore also consider shorter panel data sets with $T = 14$. This length roughly mimics the available PSID data on food expenditure from 1974 to 1987 which has been much used in Euler Equation estimation [see Alan and Browning (2010), and the references therein].[10]

Finally, one way in which researchers have tried to get around the short length of available panels is to construct synthetic cohorts from repeated cross-sectional surveys. Some repeated cross-sectional surveys, such as the Family Expenditure Survey (UK) or the Consumer Expenditure Survey (United States) are available for many years. They do not allow for individual agents to be followed over time. However, cohorts defined by fixed characteristics can be followed by time and, as shown by Browning et al. (1985) and Deaton (1985), this allows for the estimation of linear models (of individual agent behavior) in differences at the aggregate cohort level. Thus, such data can be used to estimate linearized Euler equations [see, for example, Blundell et al. (1994), Attanasio and Weber (1995), Attanasio et al. (1999)].[11]

In our Monte Carlo experiments on synthetic cohort estimation, we construct a synthetic cohort of length $T = 40$ from our simulated populations. For our synthetic cohort, we draw, with replacement, 1,000 agents for *each* period ($t = 1, 2, 3 \ldots 40$); thus, the synthetic cohort is constructed from different samples of individuals for each period, as would be the case in constructing a synthetic cohort

TABLE 3. Data structures

	Measurement error	True panel	N	T	Cohorts
Long panel	N	Y	1,000	40	–
Long panel	Y	Y	1,000	40	–
Short panel	Y	Y	1,000	14	–
Synthetic cohort	Y	N	1,000	40	1

from the U.S. Consumer Expenditure Survey or the U.K. Family Expenditure Survey.

For each of these data structures (long panel, short panel, and synthetic cohort), our baseline experiments are conducted twice: first on the simulated data, and then on the simulated data with measurement error added. Experiments on the simulated data without measurement error allow us to isolate the effect of approximation bias and are comparable to previous papers in the literature. So they are a useful starting point. Of course, in the absence of measurement error, and with true panel data, there is no reason to work with the linear approximation to the Euler equation, as one could estimate the exact Euler equation by nonlinear GMM (as noted above, to use a synthetic cohort requires the linear approximation). When we add measurement error to the simulated data, we create a scenario which mimics the one faced by researchers using actual consumption data.[12]

The measurement error that we add to the simulated data at the individual agent level is i.i.d log normal with a unit mean and a variance of 0.004 such that the approximately 75% of the period to period variance in consumption growth is due to the noise, close to the estimate for the PSID in Runkle (1991).[13] When working with the unadulterated simulated data, we lag instruments once, but when we add measurement error to the simulated data, we lag consumption instruments twice, because i.i.d measurement error in consumption levels induces measurement error with an MA(1) structure in consumption changes. We also use twice-lagged consumption growth and income instruments when working with synthetic cohorts, because of the measurement error in cohort means induced by sampling. All the different data structures we consider are summarized in Table 3.

In each of our experiments, we consider two things. As in previous papers in this literature, we examine the actual estimates in repeated samples. We report the mean finite sample bias in 999 replications, and also a nonparametric confidence interval for this quantity, as well the mean standard error across the sample and mean percentage bias. Further, we go beyond the previous literature and look directly at the properties of the instrument in the simulated data. The instruments we consider are the ones used extensively in the literature: lagged interest rates, lagged consumption growth, and lagged income.[14] We consider the validity of these instruments—whether they are uncorrelated with the residual in the log-linearized Euler equation (7)—and the relevance of these instruments—the strength with which they predict the endogenous variables in the estimating equation.

Note that because we know the true value of the preference parameter (γ) in our simulated data, we can use the true parameter value to construct the true residuals in the linearized Euler equation (plus a constant). That is, inverting equations (7) and evaluating at $\gamma = 4$ gives

$$\Delta \log C_{h,t+1} - \frac{1}{4}\log(1 + R_{t+1}) = \alpha + e_{h,t+1} \tag{10}$$

for the first-order approximation. All the terms on the left-hand side of equations (10) are observed in our simulated data, so we can calculate the right-hand side for each observation in the data. Those quantities will not be mean zero (because they contain the constant from the linearized Euler equation). As described in Section 2, the variation in those terms across individuals and time comes from innovations to marginal utility, measurement error, and approximation error where the latter comprises variation in higher order moments in consumption growth and interest rates. The constructed residuals on the right-hand side of equations (10) may or may not be correlated with the instruments. The measurement error we add in our simulations is orthogonal to the instruments. Theory indicates that the expectation error should be orthogonal to lagged variables, but this is not necessarily the case for the approximation error. With the constructed residuals in hand, the issues is open to direct empirical investigation.

Our instrument validity test is then a t-test obtained from the regression of these constructed residuals on our instruments. The null hypothesis is that the residuals are uncorrelated with the instruments: a significant t-statistic would suggest that instruments are not valid. Estimation of linearized Euler equations (with the standard instruments) will not be a promising strategy if we get many rejections of this null hypothesis. We report the fraction of t-statistics in the repeated samples of our Monte Carlo experiments that exceed 1.96 in absolute value (corresponding to a 5% test). It is important to note that this is not a standard over-identification test. Because we can construct the appropriate residuals, we can test the validity of the instruments even when the equation of interest is just identified.

In each sample in each experiment, we also conduct the "first-stage" regression of the endogenous regressor, contemporaneous interest rate, on the instruments to examine instrument relevance (that is, to check for weak instruments).[15] The null hypothesis of this test is instruments are jointly weak. Estimation of linearized Euler equations (with the standard instruments) can only be a promising strategy if we get many rejections of this null hypothesis. Stock and Yogo (2005) calculated the critical values for this test as a function of the number of included endogenous regressors, the number of instrumental variable, and the desired maximal bias of the IV estimator relative to OLS. In the case for one endogenous variable, allowing for a maximum relative bias of 10% compared to OLS, and at the 5% significance level and with three of instruments (as in our experiments with the first-order linearized Euler equation), the critical value is 9.08 for Linear GMM, and 6.46 for Limited Information Maximum Likelihood (LIML).

In our experiments, realizations of the interest rate process are common to all agents. This means that we have an aggregate variable on the right-hand side of equations (7) and the usual formulas may significantly underestimate the standard error of the estimates [Moulton (1986)]. Accordingly, we cluster the standard error on the time period.

Finally, we note that when estimating on data simulated from models D-P and particularly D-I, the Euler equation does not hold in situations where the explicit borrowing constraint is binding. Thus, we discard the periods in which agents do not carry forward assets between periods.[16] Note that the conditional expectation expressed in Equation (1) which underpins Euler equation still holds conditional on this selection: we are selecting on the basis of a variable in the information set.

4. THE CONSUMPTION FUNCTION AND LINEARIZED EULER EQUATIONS

Before turning to the Monte Carlo results, it is useful to develop some hypotheses regarding when the estimation of linearized Euler Equations is likely to perform well, and when it is likely to fail.

The potential problem with the approximate Euler equation (7) is that variation in the higher order moments of consumption growth (and consumption growth and interest rates) that are subsumed into the residual term may be correlated with lagged variables, leaving researchers without any valid instruments. For example, there is no theoretical reason that the lagged consumption growth should be uncorrelated with conditional skewness [Carroll (2001)]. We refer to the resulting inconsistency of the estimate as approximation bias.

Of course, this is more likely to be a problem the larger (and more variable) the higher moments are. For example, a comparison of the first-order approximations in Section 2 [equation (7)] shows that the key omitted variable in the first-order linearized Euler equation is the variance of consumption growth, which may be correlated with the lagged consumption growth.

The models we consider are homogeneous with respect to the stochastic processes. For a given model, all agents face the same income and interest rate processes. Across models, the interest rate process is the same, and the only difference across models in the income process is the small positive probability of zero income realization in models three and four. Thus, differences in the variance of consumption growth are driven not by these stochastic processes but rather by the sensitivity of log consumption to realizations of the state variables. This in turn is controlled by the semi-elasticity of consumption function (policy rule) with respect to the state variables. The semi-elasticity with respect to cash-on-hand turns out to be critical. A large semi-elasticity implies that shocks to cash-on-hand will pass through to greater variability in consumption growth. This has two consequences. First, from the point of view of estimating interest rate responses, greater variability in consumption growth coming from income shocks means more noise and hence less precision. Second, when the higher order moments of
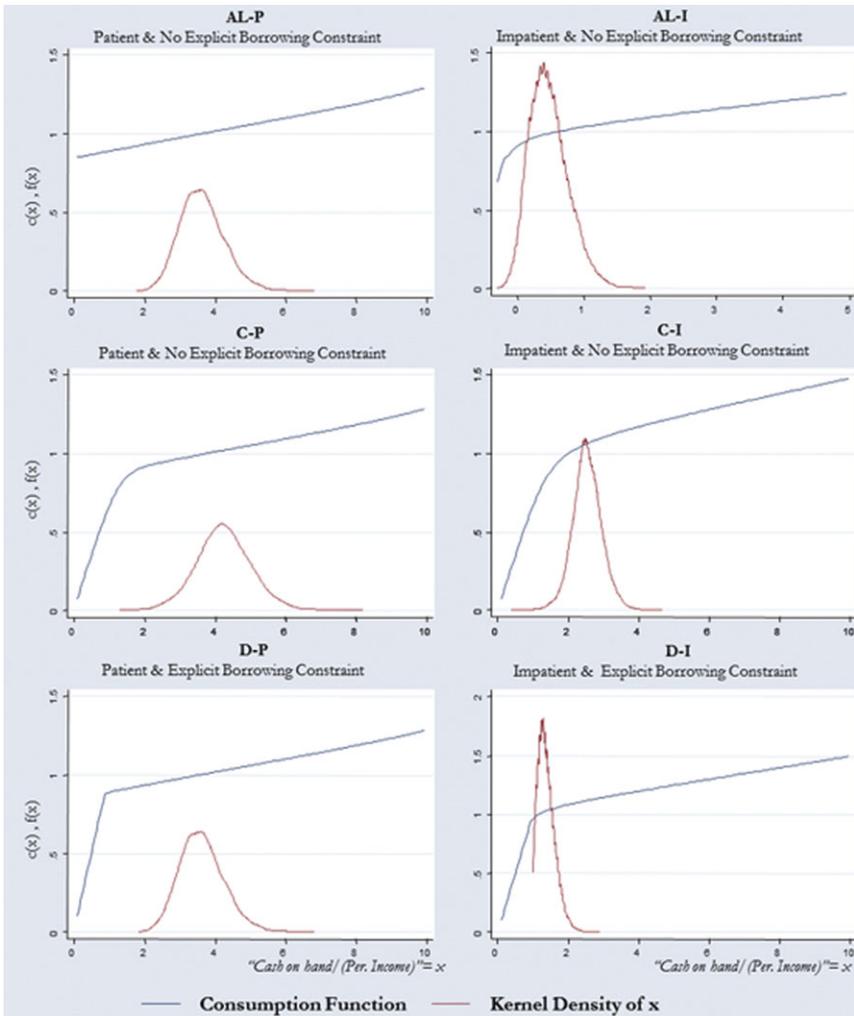
FIGURE 1. Consumption functions and distribution of cash-on-hand. Consumption Function and distribution of normalized cash-on-hand at the age of 40. *X*-axis is cash-on-hand to permanent income ratio.

consumption growth are larger, there is more scope for them to vary and potentially correlate with the instruments.[17] Thus, we expect that when the semi-elasticity is large, estimates based on linearized Euler equations will certainly be less precise, and they may be more biased.

Figure 1 illustrates these key characteristics of our models. For each model, the numerically solved consumption function for age 40, and the (simulated) distribution of normalized cash-on-hand at the same age are plotted. Of course,

the consumption function also depends on the interest rate. We solve the model for a stochastic interest rate but for comparability we simulate all six models with a common vector of interest rate realizations. In all our simulations, the interest rate realization at age 40 is 0.033 (recall that the long run average of the interest rate process is 0.03).

In most of the models we study, the policy functions (that is, consumption functions) are nonlinear, and broadly similar: steep (and curved) at low cash-on-hand levels but much flatter (and near-linear) at high cash-on-hand levels. The consumption function for AL-P is distinctive in that it is linear, while the consumption for D-P is distinguished by a sharp kink.

What determines the degree of approximation bias is not the overall shape of the policy function (which is common across many models), but the semi-elasticity of the consumption function in the region of the state space in which agents operate. It is clear from the figures that the cash-on-hand distributions of patient models (models AL-P, C-P, and D-P) are located at higher cash-on-hand levels. As a result, for agents in these models, the steep parts of the consumption function are irrelevant and their weighted average semi-elasticity is low. On the contrary, in impatient models (models AL-I, C-I, and D-I) agents accumulate very little wealth. They operate on the steep part of their consumption function, and the weighted average semi-elasticity is high. The consumption growth of these agents will thus be much more sensitive to shocks to cash-on-hand, and the resulting higher variability of consumption growth gives greater scope for approximation bias.[18]

A comparison of AL-P and D-P illustrates our point. These models are superficially quite different in the sense that borrowing is allowed in AL-P but not in D-P and the shapes of the consumption functions for these two models are very different. However, in the region of the state space where agents operate, as indicated by the empirical density of cash-on-hand, the consumption functions are very similar, and this is reflected in similar performances of linearized Euler equations in these environments.

5. MONTE CARLO RESULTS

5.1. Baseline Experiments

We begin our discussion of the results of the Monte Carlo experiments with Table 4. Table 4 reports evidence on instrument validity and instrument relevance. As described above, instrument validity is assessed by regressing the error term from the linearized Euler equation (plus a constant) on the instruments. The error term is calculated (without estimation) from the simulated interest rate and consumption growth, and the true value of the preference parameter, γ (see equation 10). Instrument relevance is assessed by regressing the endogenous explanatory variable—here the interest rate—on the instrument set. Again the instrument set includes lagged interest rate, lagged consumption growth, and lagged income for

TABLE 4. Instrument validity and relevance results, 999 replications

| | | Instrument validity | | | Instrument relevance |
| | | First order: % $\Delta\log C^h_{t+1} - \frac{1}{4}\log(1+R_{t+1}) = \alpha + e^h_{t+1}$ | | | (First stage) |
Model		Lag interest rate	(Twice) lag consumption growth	Lag income	F-stat Mean [10%, 90%]
AL-P	Base				
	[2.5%, 97.5%]	[−1.93, 1.88]	[−2.14, 2.59]	[−0.97, 2.54]	6,999
	"Fraction of $|t| > 1.96$"	0.042	0.072	0.076	[6,996, 7,003]
	Measurement error				
	[2.5%, 97.5%]	[−0.20, 3.95]	[−1.84, 2.56]	[−1.83, 2.28]	5,523
	"Fraction of $|t| > 1.96$"	0.031	0.067	0.053	[5,521, 5,526]
AL-I	Base				
	[2.5%, 97.5%]	[−2.83, 2.62]	[−8.88, −0.65]	[−4.01, 2.02]	7,458
	"Fraction of $|t| > 1.96$"	0.092	0.432	0.789	[7,450, 7,467]
	Measurement error				
	[2.5%, 97.5%]	[−0.11, 5.31]	[−4.16, 1.19]	[−4.15, −1.16]	7,146
	"Fraction of $|t| > 1.96$"	0.088	0.376	0.801	[7,143, 7,151]
C-P	Base				
	[2.5%, 97.5%]	[−1.96, 2.53]	[−7.3, 1.08]	[−1.01, 3.63]	7,361
	"Fraction of $|t| > 1.96$"	0.097	0.132	0.084	[7,343, 7,378]
	Measurement error				
	[2.5%, 97.5%]	[−2.30, 0.32]	[−2.48, 2.29]	[−0.53, 2.42]	7,078
	"Fraction of $|t| > 1.96$"	0.054	0.101	0.064	[7,058, 7,098]
C-I	Base				
	[2.5%, 97.5%]	[−2.17, 1.95]	[−7.78, 7.72]	[−3.05, 6.80]	7,073
	"Fraction of $|t| > 1.96$"	0.102	0.358	0.470	[7,056, 7,091]
	Measurement error				
	[2.5%, 97.5%]	[−4.88, 3.99]	[−3.19, 1.41]	[−0.71, 2.76]	7,056
	"Fraction of $|t| > 1.96$"	0.128	0.265	0.204	[7,032, 7,089]
D-P	Base				
	[2.5%, 97.5%]	[−1.87, 2.28]	[−6.40, 1.60]	[−1.01, 2.37]	6,999
	"Fraction of $|t| > 1.96$"	0.037	0.073	0.065	[6,996, 7,003]
	Measurement error				
	[2.5%, 97.5%]	[−1.09, 0.97]	[−4.41, 2.43]	[−0.87, 1.63]	5,527
	"Fraction of $|t| > 1.96$"	0.021	0.093	0.010	[5,522, 5,530]

the data without measurement error, and twice-lagged consumption growth, the lagged interest rate, and lagged income for the data with measurement error.

The final column of Table 4 reveals that in these models, using the first-order approximation as a basis for estimation, the standard instrument set has very good predictive power for the interest rate. There is no issue of weak instruments. The table also shows that the lagged interest rate is generally a valid instrument. In impatient models (AL-I, C-I, and D-I), we have significant validity problems with the

TABLE 4. Continued

		Instrument validity			Instrument relevance
		First order: % $\Delta \log C_{t+1}^h - \frac{1}{4} \log (1 + R_{t+1}) = \alpha + e_{t+1}^h$			(First stage)
D-I	Base				
	[2.5%, 97.5%]	[−2.07, 2.49]	[−9.06, −0.38]	[−6.09, −2.32]	6,961
	"Fraction of $|t| > 1.96$"	0.092	0.832	0.989	[6,907, 7,011]
	Measurement error				
	[2.5%, 97.5%]	[−3.9, 2.34]	[−2.84, 0.89]	[−4.06, 0.75]	6,681
	"Fraction of $|t| > 1.96$"	0.130	0.224	0.139	[6,572, 6,776]

Notes:

1. All other numbers are result of 999 Monte Carlo replications.
2. Panel length is 40 periods.
3. For the first-order approximation instruments, we use the twice-lagged consumption instrument with the lag interest rate and lag income.
4. Instrument validity test is a t-test obtained from the regression of constructed residuals on instruments.
   a. Instrument validity columns report for each instrument. (i) t-statistic values at 2.5% and 97.5%, (ii) the fraction of t-stats with absolute value greater than 1.96 (critical value at 5% significance level).
   b. Instrument relevance column reports the Cragg–Donald F (CDF) statistic from the first stage of IV. For the first-order approximation interest rate is the only endogenous variable and lagged interest rate, (twice) lagged consumption growth, lagged income are instruments. Mean values of CDF are reported. CDF values at 10% and 90% are in parentheses.
   c. Stock and Yogo (2002)—test for weak instrument = bias of two-stage estimation relative to OLS is greater than 10%.
      i. Critical Value at 5% significance level when the number of instruments is 3 and endogenous variable is 1 for linear GMM = 9.08, for LIML =6 .46.
5. The measurement error that we add to the simulated data is i.i.d log normal with a unit mean and a variance of 0.004.

other instruments. For example, in AL-I, without measurement error, we find that lagged consumption growth is significantly correlated with the error term in 43% of samples. For C-I, lagged consumption growth is significantly correlated with the error term 36% of the samples. This suggests that lagged consumption growth and lagged income are not valid instruments for the estimation of the first-order linearized Euler equation (even though the model we simulate implies that they must be uncorrelated with the innovation in marginal utility). These results highlight the trade-off between instrument validity and relevance, and supports Carroll (2001) arguments. We find that even instruments work well in the first-stage regressions, they are usually invalid instruments. Interestingly, adding realistic measurement error to consumption appears to slightly improve the situation because it weakens the correlation between these instruments and the error terms (without much cost in terms of instrument relevance). For example, in the C-I model, the frequency with which lagged consumption growth is significantly correlated with the error term falls from 36% to 27%. This could be because the error terms now contain the measurement error (in addition to the variation in approximation error and the innovations to marginal utility) or it could be because, in the presence of measurement error we use second (rather than first) lags of consumption growth in the instrument.

TABLE 5. Estimates of the EIS, 999 replications, first-order approximation, long panel

		Base	With measurement error
	True value EIS: $\left(\frac{1}{\gamma}\right) = 0.25$		
Model			
AL-P	Mean coefficient	0.234	0.279
	[2.5%, 97.5%]	[0.19, 0.28]	[0.207, 0.354]
	(Mean std. error)	(0.078)	(0.106)
	(Mean bias as % of true parameter value)	(−6.4)	(11.6)
AL-I	Mean coefficient	0.218	0.284
	[2.5%, 97.5%]	[0.15, 0.29]	[0.113, 0.496]
	(Mean std. error)	(0.11)	(0.245)
	(Mean bias as % of true parameter value)	(−11)	(−12)
C-P	Mean coefficient	0.237	0.212
	[2.5%, 97.5%]	[0.203, 0.269]	[0.166, 0.283]
	(Mean std. error)	(0.090)	(0.098)
	(Mean bias as % of true parameter value)	(−5.2)	(15.2)
C-I	Mean coefficient	0.216	0.298
	[2.5%, 97.5%]	[0.17, 0.27]	[0.182, 0.653]
	(Mean std. error)	(0.123)	(0.189)
	(Mean bias as % of true parameter value)	(−13.6)	(19.2)
D-P	Mean coefficient	0.235	0.274
	[2.5%, 97.5%]	[0.19, 0.29]	[0.196, 0.354]
	(Mean std. error)	(0.078)	(0.105)
	(Mean bias as % of true parameter value)	(−6)	(9.6)
D-I	Mean coefficient	0.213	0.209
	[2.5%, 97.5%]	[0.115, 0.319]	[0.145, 0.292]
	(Mean std. error)	(0.192)	(0.194)
	(Mean bias as % of true parameter value)	(−14.8)	(−16.4)

Notes:

1. Panel length is 40 periods.
2. For each model and estimation strategy, the first number is the mean value of the estimate of the EIS in 999 Monte Carlo replications. The second number, in brackets is 2.5% and 97.5% of EIS estimates. Third number is the mean robust standard errors clustered on period. Fourth number is the mean bias as the percentage of true parameter value.
3. Online Appendix Figure A1 presents the full empirical CDF of EIS estimates.

Table 5 reports the estimates of the EIS that result from estimating the first-order linearized Euler equation on these models with the full set of instruments. The left-hand column gives the results for unadulterated consumption data, the second column gives the results when realistic measurement error is added to the data. The true value of the EIS is 0.25. For each model, and in each column, we report four numbers. First, at the top, the mean estimate in 999 samples. Second, in

square parentheses, we report semi-parametric confidence intervals and reporting the values corresponding to the 2.5% lowest and highest estimates. Third in round parentheses, is the mean standard error of the estimate across the samples. Finally, at the bottom is the mean of the percentage bias across the samples.[19]

In the left-hand column, without measurement error, we see that the mean bias is negative in all models, The first-order linearization of the Euler equation tends to lead to underestimates of the EIS. As expected the problem is much more severe in impatient models. In patient models, the mean bias is around 6% of the true value. In impatient models, this rises to 15% of the true value.

For a more complete understanding of the distribution of the Monte Carlo results of EIS estimates across the models, we plot the empirical CDF of EIS estimates across six models in the Online appendix. From the cumulative distribution, one can read off the percentage of the sample that is more or less than a particular value, allowing us to complete assessment of the effect of different models on the full distribution of estimates. From panel A of Figure A1, it is clear that patient models have a narrower distribution near the true EIS value confirming that the patient models are doing better than impatient models. The CDFs reveal that for all models, estimates quite far from the true value arise is a significant number of replications. In the absence of measurement error, these are usually on the low side, but with measurement error there are many more high estimates.

The right-hand column of Table 5 reveals that measurement error sometimes worsens things, but not always. There is a loss of precision in all cases, but this is sometimes small. The mean percentage bias sometimes rises, and sometimes falls. The pattern of the estimation strategy performing much better in models with patient agents is no longer as sharp. This can also clearly be seen in the panel B of Figure A1 in the online appendix.

We would summarize the result of these section as follows. With a long panel and sufficient variation (recall that $T = 40$ here, and the autocorrelation coefficient in the interest rate process is 0.6), the linearized Euler equation is modestly successful in recovering the EIS. Across a range of quite different models, and with realistic measurement error in consumption, the mean bias in estimates of the EIS ranges from 10% to 19% (see the second column of Table 5). This replicates the findings for Attanasio and Low (2004). However, note that in impatient models, the estimates are quite imprecise. For example, with the data from the AL-I model, with realistic measurement error almost 2.5% of the estimates of the EIS are above 5 (where the true value is 0.25). For the AL-I and CL-I models, even with a long panel, the typical confidence interval (for a given sample) includes both 0 and 1 in more than 10% of samples (see online appendix Table A6 for the exact number).[20]

These baseline estimates are based on a true annual panel of 40 years in length. While this is useful "best case" scenario, given the cost of panel surveys, as well as problems with attrition, this is not realistic. In the next subsection, we consider two more realistic data structures.

5.2. Alternative Data Structures

In Tables 6 and 7, we report experiments with two more realistic data structures. As described in Section 3, these are a short annual panel of 14 years, and a synthetic cohort observed for 40 years. In this subsection, our intention is to be as realistic as possible and so we always experiment with consumption data which includes measurement error. To facilitate comparison, we present results based on a long true panel with measurement error (repeated from the previous section) alongside the results for more realistic data structures.

Table 6 reports our findings on instrument validity and instrument relevance with these alternative data structures. There are two points to note. First, problems with the validity of lagged consumption growth and lagged income are less apparent with the synthetic cohort. Thus, it is largely the individual (cross-sectional) variation in these instruments, rather than the time-series variation, that undermines validity in the (true) panel data. Second, turning to instrument relevance, the instruments are much weaker at the cohort level and in many case are below the critical value of the Cragg–Donaldson F-test (recall that critical value for one endogenous and three instruments at the 5% significance level is 9.58 for Linear GMM and 6.46 for LIML). For this reason, when estimating the approximate Euler equation in these experiments, we used LIML. LIML is known to perform better when instruments are weak [see, e.g., Andrews and Stock (2005), Hahn et al. (2003)].[21]

Table 7 summarizes the estimates from these Monte Carlo experiments. As in Table 5, we report the mean estimates (in 999 replications), the semiparametric confidence interval, the mean standard error, and the mean percentage bias. Examining Table 7 shows that estimation on the short panel is, in general, not successful. The extent of mean bias, particularly on the impatient models, is unacceptable. For example, when the data are generated by C-P or C-I, the mean bias is 33% and 38% of the true parameter value, respectively. These results suggest that 14 annual observations—as are afforded by the PSID—are not adequate for linearized Euler equation estimation.

In terms of mean bias, the synthetic cohort produces results that compare quite well to the long panel. The cost though is a loss of precision: the mean standard error is about twice as large in the synthetic panel. Here, we are estimating with a single cohort, and some additional precision might come from following multiple birth cohorts through the data. The results also highlight that, for all data structures, linearized Euler equation estimates offers poor performance when consumers are impatient. The results for the D-I model are the least favorable of all the models we consider and are particularly bad for synthetic cohort estimation.[22] In Figure 2, we present the full empirical CDF of EIS estimates for each combination of model (data generating process) and data structure (long panel, short panel, and synthetic cohort). From Figure 2, we see that using short panels causes the distribution of EIS estimates to shift inwards and to the left, indicating an increasing concentration of estimates with lower values. For example, in patient models, over 10% of

TABLE 6. Alternative data Structures first-order approximation with measurement error, instrument validity, and relevance results

			Instrument validity		Instrument relevance
		"[2.5, 97.5] of t statistic" and "fraction of t-stats bigger than 1.96 (2.024)"			F-stat
Model		Lag interest rate	Twice lag consumption growth	Lag income	Mean [10%, 90%]
AL-P	Long panel	[−0.20, 3.95]	[−1.84, 2.56]	[−1.83, 2.28]	5523
		0.031	0.067	0.053	[5521, 5526]
	Short panel	[−2.55, 3.41]	[−3.27, 1.71]	[−2.48, 2.78]	497
		0.082	0.161	0.102	[496 , 510]
	Synthetic cohort	[−2.05, 2.41]	[−2.39, -0.63]	[−2.15, −1.39]	6.48
		0.042	0.060	0.040	[5.42 , 8.10]
AL-I	Long panel	[−0.11, 5.31]	[−8.88, −0.65]	[−4.15, −1.16]	7146
		0.088	0.376	0.801	[7143, 7151]
	Short panel	[0.70, 2.24]	[−3.15,2.71]	[−3.88, 1.56]	497
		0.096	0.298	0.745	[496 , 500]
	Synthetic cohort	[−3.61, 1.46]	[−4.48, −0.70]	[−3.75, 0.408]	7.97
		0.092	0.212	0.279	[6.36 , 12.9]
C-P	Long panel	[−2.30, 0.32]	[−2.48, 2.29]	[−0.53, 2.42]	[−2.30, 0.32]
		0.054	0.101	0.064	0.054
	Short panel	[−1.23, 2.14]	[−1.84, 2.95]	[−3.58, 1.81]	491
		0.065	0.126	0.163	[483, 495]
	Synthetic cohort	[−2.84, 1.64]	[−14.3, 2.76]	[−1.42, 1.97]	7.90
		0.085	0.113	0.039	[6.36, 15.23]
C-I	Long panel	[−4.88, 1.99]	[−3.19, 1.41]	[−0.71, 2.76]	7056
		0.128	0.265	0.204	[7032, 7089]
	Short panel	[−7.88, 11.59]	[−7.92, 1.93]	[−1.47, 2.62]	491
		0.182	0.235	0.234	[485, 497]
	Synthetic cohort	[−3.78, 2.79]	[−4.60, 0.89]	[−4.52, 2.41]	7.70
		0.148	0.284	0.305	[6.37, 14.87]
D-P	Long panel	[−1.09, 0.97]	[−4.41, 2.43]	[−0.87, 1.63]	5527
		0.021	0.093	0.010	[5522, 5530]
	Short panel	[−2.05, 2.05]	[−3.28, 1.56]	[−2.12, 2.76]	501
		0.70	0.167	0.084	[496, 509]
	Synthetic cohort	[−2.20, 2.21]	[−2.17, −0.70]	[−2.10, −0.40]	6.46
		0.052	0.059	0.023	[5.39, 8.05]
D-I	Long panel	[−3.9, 2.34]	[−2.84, 0.89]	[−4.06, 0.75]	6681
		0.130	0.224	0.139	[6572, 6776]
	Short panel	[−13.7, 0.64]	[−9.06, 7.41]	[−1.41, 7.87]	504
		0.319	0.706	0.234	[498, 511]

EIS estimates are lower than 0.1 (indicating a relative risk aversion of 10). In all models, we observe that, even when mean biases are moderate, extreme values such as EIS values smaller than 0.1 or bigger than 1 are observed frequently. When we construct a confidence interval for each estimate, in all models and either of

TABLE 6. Continued

	Instrument validity			Instrument relevance
	"[2.5, 97.5] of *t* statistic" and "fraction of *t*-stats bigger than 1.96 (2.024)"			*F*-stat
Model	Lag interest rate	Twice lag consumption growth	Lag income	Mean [10%, 90%]
Synthetic cohort	[−3.02, 2.10] 0.103	[−2.34, 0.98] 0.203	[−2.06, 1.75] 0.089	6.36 [3.17, 11.67]

Notes:

1. Table reports the first-order approximation results using the simulated data with measurement error. The measurement error that we add to the simulated data is i.i.d log normal with a unit mean and a variance of 0.004.
2. For long panels, panel length is 40 periods ($T = 40$), for short panels $T = 14$, and for synthetic cohort, $T = 40$. All other numbers are result of 999 Monte Carlo replications.
3. For long-panel and short-panel data structures, we use the twice-lagged consumption instrument with the lag interest rate and lag income. For the synthetic panel data, we also twice-lagged the income.
4. Instrument validity test is a *t*-test obtained from the regression of constructed residuals on instruments.
   a. For long panel and short panel data structures, instrument validity columns report for each instrument the fraction of *t*-stats with absolute value greater than 1.96 (critical value at 5% significance level).
   b. For synthetic panel data, instrument validity columns report for each instrument the fraction of *t*-stats with absolute value greater than 2.024 (critical value at 5% significance level).
5. Instrument relevance column reports the Cragg–Donald *F* (CDF) statistic from the first stage of IV. For the first-order approximation, interest rate is the only endogenous variable and lagged interest rate, twice-lagged consumption growth, (twice-) lagged income are instruments..
   a. Mean values of CDF are reported. CDF values at 10% and 90% are in parentheses. Stock and Yogo (2002)—test for weak instrument = bias of two-stage estimation relative to OLS is greater than 10%. Critical value at 5% significance level when the number of instruments is 3 and endogenous variable is 1 for LIML = 6.46.

the realistic data structures, the confidence interval associated with almost every estimate (that is, sample) contains either 0.1 or 1.[23]

The EIS measures the response of consumption to a change in the inter-temporal price. To get an economic sense of the bias and imprecision we are reporting, consider the temporary VAT cut that the UK government implemented in 2008 in response to deteriorating economic conditions. The aim of the VAT cut, from 17.5% to 15%, was to stimulate household consumption. Assuming full pass through and ignoring unrated goods, a cut from 17.5% to 15% would be a reduction in the current inter-temporal price of consumption of 2.5/117.5 = 2.1%. Holding the marginal utility of consumption constant to focus on the (Frisch) substitution effect, the EIS assumed in our baseline simulations (0.25) implies a boost to current consumption of about half a percent (0.25 × 2.1%). An EIS of 0.1 would imply a much smaller bump to final consumption of less than a quarter of a percent; an EIS of 1 would imply a much larger increase to final consumption of over 2%. (for comparison, real household final consumption expenditure fell by 3.5% in 2009, after rising by an average of about 3.5% per year over the previous decade [Office of National Statistics (2017)]. The point is that, with realistic data structures, most of the replicate samples in our Monte Carlo study would suggest either was possible.

TABLE 7. Alternative dta structures, first-order approximation with measurement error estimates of EIS, 999 Replications, LIML

Model		Mean coefficient	[2.5, 97.5]	Mean standard error	Mean bias as % of true parameter value
AL-P	Long panel	0.279	[0.217, 0.363]	(0.106)	(11.6)
	Short panel	0.211	[0.038, 0.495]	(0.336)	(−15.6)
	Synthetic cohort	0.218	[0.079, 0.360]	(0.224)	(−12.8)
AL-I	Long panel	0.284	[0.113, 0.454]	(0.245)	(13.6)
	Short panel	0.204	[0.046, 0.486]	(0.356)	(−18.4)
	Synthetic cohort	0.222	[0.063, 0.370]	(0.339)	(−11.2)
C-P	Long panel	0.212	[0.166, 0.283]	(0.098)	(15.2)
	Short panel	0.332	[0.123, 0.634]	(0.366)	(32.8)
	Synthetic cohort	0.203	[0.097, 0.467]	(0.232)	(−18.8)
C-I	Long panel	0.298	[0.182, 0.653]	(0.189)	(19.2)
	Short panel	0.345	[0.190, 0.516]	(0.285)	(37.6)
	Synthetic cohort	0.297	[0.142, 0.429]	(0.335)	(18.8)
	Long panel	0.274	[0.196, 0.354]	(0.105)	(9.6)
	Short panel	0.208	[0.035, 0.498]	(0.335)	(−16.8)
	Synthetic cohort	0.214	[0.049, 0.373]	(0.223)	(−14.4)
	Long panel	0.209	[0.145, 0.292]	(0.194)	(−16.4)
	Short panel	0.183	[0.032, 0.438]	(0.251)	(−26.8)
	Synthetic cohort	0.200	[0.054, 0.378]	(0.401)	(−20)

Notes:

1. For long-panel and short-panel data structures, we use the twice-lagged consumption instrument with the lag interest rate and lag income. For the synthetic panel data, we also twice-lagged the income.
2. For long panels, $T = 40$; for synthetic cohort, $T = 40$.
3. Estimations are done by LIML.
4. Figure 2 presents the full empirical CDF of EIS estimates.

Similarly, with additive inter-temporal preferences, the EIS is the inverse of the coefficient of relative risk aversion. An EIS of 1, 0.25, and 0.1 implies a coefficient of RRA of 1, 4, and 10, respectively. Most economists will recognize that, holding time preference constant, a consumers with RRA of 1, 4, and 10 behave very differently (with respect to portfolio choice, insurance demand, etc.).

5.3. Alternative Estimation Strategies

Here, we discuss two alternative estimation strategies to linearized Euler equation.[24] First, a few empirical studies [Dynan (1993), Ludvigson and Paxon (2011)] use second-order approximation instead of linearized Euler equation. In the Appendix A2, we examine the performance of this estimation strategy in our simulated environments. Our Monte Carlo results show that the second-order
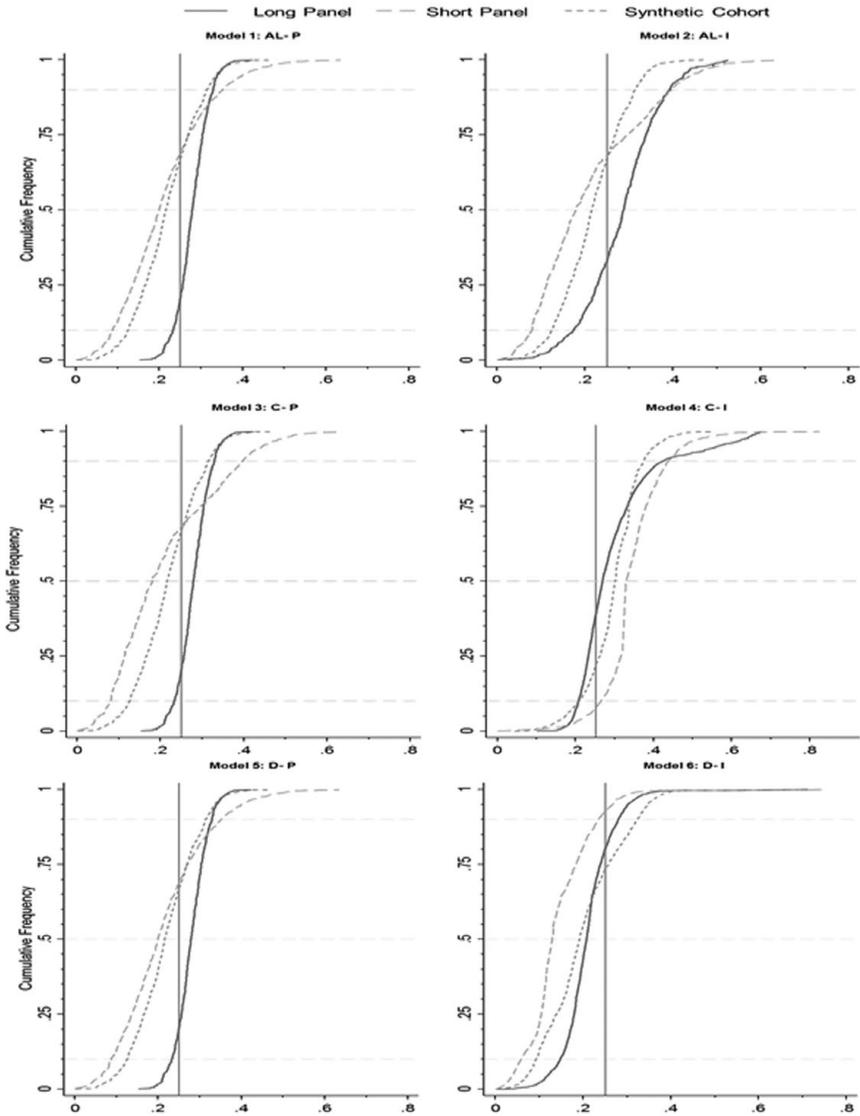
FIGURE 2. Cumulative density functions of EIS estimates with measurement error using different data structures. The vertical reference line is at the true value of EIS (0.25). Detailed statistics of Monte Carlo results are presented in Table 7.

approximation to the Euler equation does not provide a superior basis for estimation, and in fact suffers from a weak instrument problem: the standard instruments do not have useful predictive power for the second-order term.

Second, an alternative strategy is estimating the full nonlinear Euler equation [Equation (3)] using the generalized method of moments methodology introduced

by Hansen (1982), and hence eliminating problems associated with approximation. Appendix A3 discusses the results of two alternative nonlinear GMM estimation on the consumption data with measurement error. The first non-linear GMM is the Exact GMM estimator introduced by Hansen (1982). While this estimator avoids the problems associated with approximation, it reintroduces the problems associated with measurement error in nonlinear models. The second estimator is the GMM-D procedure proposed by Alan et al. (2009), which takes explicit account of the measurement error in consumption. We find that in our environments neither nonlinear GMM estimator outperforms linearized Euler equation estimation.

In real data, it may be more difficult to identify constrained households. Therefore, as a final investigation, we repeat our analysis for D-I without excluding constrained households. The results are reported in the online appendix Tables A4 and A5. Table A4 shows that problems with the validity of instruments are more prevalent when we do not exclude constrained households. Table A5 further shows that mean bias is larger and the estimates have higher variance when we do not exclude constrained observations.

6. CONCLUSION

Two key parameters in macroeconomics are the sensitivity of consumption to interest rates (inter-temporal substitution) and to uncertainty (prudence). A common approach to estimating these parameters is based on Euler equations. Often these nonlinear first-order conditions are log-linearized so that linear instrumental variables methods can be used to deal with measurement errors in micro data on consumption. However, linearization may itself induce an (approximation) bias so that the usefulness of the Euler equation approach has been debated. We have explored this issue with a series of simulation studies, and our findings offer a reconciliation of the debate. On the one hand, our results confirm the finding of Attanasio and Low (2004) that with sufficiently long time series and sufficient variation in the interest rate, linearized Euler equation estimation works reasonably well. We strongly affirm their emphasis on the central role that variation in the inter-temporal price must play in estimating the elasticity of consumption with respect to that price.

More broadly, though, our findings support Carroll (2001) and others in this literature who are skeptical of the usefulness of linearized Euler equation estimation. We find that the performance of linearized Euler equation estimation declines, with both greater mean bias and a loss of precision, when agents are moderately impatient. Perhaps more importantly, linearized Euler equation estimation does not appear to work well on any realistic data structure. A sufficient time dimension can only realistically be achieved with a synthetic cohort, and our experiments suggest that estimates from synthetic cohorts of sufficient length, while often exhibiting small mean bias, are very imprecise.

In light of these findings, we suggest that researchers may need to pursue other methods for obtaining the key parameters of consumers' inter-temporal choice

problem. The "semi-structural" methods of Alan and Browning (2010), Alan et al. (forthcoming), and Druedahl and Jorgensen (2016) are one promising direction.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit http://dx.doi.org/ S1365100518000032.

NOTES

1. While aggregate data may avoid the measurement error in micro data, there is substantial evidence that estimating these models with aggregate data can lead not only to biased parameter estimates but also to false rejections of the underlying models [see, for example, Attanasio and Weber (1993)].

2. For example, recent papers that use log-linearized Euler equations include Dogra and Gorbachev (2016), Huang et al. (2015), Mazzocco (2007), and Mudit and Ravi (2016). On the other hand, recent papers that cite Carroll (2001) against the use of linearized Euler equation include Gomes and Issler (2017), Gayle and Khorunzhina (2016), Adams et al. (2016), and Benito and Mumtaz (2009).

3. Attanasio and Low are explicit that they are not trying to study realistic data structures.

4. The issue of instrument relevance (whether instruments are weak) has been studied for consumption Euler equations estimated on aggregate data [Yogo (2004)], but not, to our knowledge, in the case of approximate Euler equations estimated on micro data.

5. Of course, if the agent has access to several assets, and she is not at a corner, one can derive a similar condition for each of these assets.

6. Note that η_t is not in the agent's information set at time t and cannot be taken outside the conditional expectation.

7. Here, we are following the literature: Attanasio and Low (2004), Alan et al. (2009), and Alan and Browning (2010) all assume an AR(1), and both Attanasio and Low (2004) and Alan and Browning (2010) show that this specification is supported by the data.

8. Carroll (2001) specifies zero-income probabilities of 0.01, 0.03, and 0.05 in alternative experiments. Attanasio and Low (2004) allow for zero income with a probability of 0.05 in one of their robustness checks.

9. Dropping initial periods, to allow the asset distribution to settle down to the ergodic distribution, is the right thing to do when simulating an infinite horizon model. When simulating a life-cycle model, it is not really necessary, but we do this to ensure that our negative results for linearized Euler equation estimation are not driven by the initial periods in the simulations.

10. Although the PSID began in 1969 and continues, the food variables expenditure are very hard to interpret prior to 1974, and food-related questions are suspended for several years after 1987. This case illustrates typical practical problems with true panels. It is not sufficient for the panel to continue for many periods. It is also necessary that the consumption information be collected continuously and in a consistent fashion.

11. Note that this is not the same as estimating an Euler equation on aggregate consumption data, the individual-agent-level equation is explicitly aggregated, avoiding the problems identified by Attanasio and Weber (1993). As pointed out in Attanasio and Low (2004), estimation on synthetic cohorts requires an equation that is linear in parameters. Estimation on aggregate data requires a household-level equation that is also linear in variables. Attanasio and Low (2004) emphasize that aggregation to the cohort level can have additional benefits (in addition to long T); in particular, there can be some averaging out of the measurement error in household level consumption data. However, there are downsides to aggregation as well. While no variation in aggregate-level instruments (the lagged interest rate) is lost, potentially useful within-cohort variation in other instruments (lagged consumption

growth, lagged income) is lost. Moreover, as shown in Deaton (1985), constructing cohort means year by year from fresh samples is subject to sampling error, and this sampling error effectively induces a cohort-level measurement error in the means. Thus, it is of interest to compare the performance of linearized Euler equation estimation on synthetic cohort data to the estimation of the same equation on true panels of different lengths.

12. Again, when we work with synthetic cohort, this measurement error will be averaged out, but a second source of measurement error (the sampling variation in cohort means) is introduced.

13. This is the same measurement error structure as assumed in Alan et al. (2009). We also repeated our experiments with a measurement error variance that implied that approximately 50% of the period to period variance in consumption growth is due to the noise (i.e. with a smaller measurement error variance.) The results were similar to those obtained under our baseline measurement error assumption, and so for sake of brevity they are not reported, but are available upon request.

14. We use the same instrument set as Attanasio and Low (2004). When we introduced measurement error, we use twice-lagged consumption growth instead of lagged consumption growth.

15. This is the F statistics from the first stage when there is only one endogenous regressor.

16. On average (across simulated samples), 6% of observations are deleted from the sample. We also obtained results without excluding these observations, and we discuss those results at the end of Section 5.3.

17. If the higher moments of consumption growth were large but varied neither through time or across agents, then of course they would be subsumed in the intercept of the linearized Euler equation, and could not be correlated with the instruments. The point here is that larger second and higher moments give greater scope for potential correlation.

18. In the working paper version of the paper, we propose a simple nonparametric statistic (effective semi-elasticity) to summarize this relevant curvature of consumption functions. This statistic integrates average elasticity of policy function over the ex-post distribution of the key state variable (cash-on-hand). Hence, it captures both the semi-elasticity of the policy function per se, and the ex-post relevance of different parts of the state space (and hence different parts of the policy function). We use this measure to predict models for which linearized Euler equations are likely to be a poor basis for estimation. Our results show that effective semi-elasticity is a good predictor of failure within classes of models, less good for very different models. Details of the calculation of effective semi-elasticity can be found in the working paper.

19. We also calculated the median percentage bias and root-mean-squared error in each case. Median percentage bias measures were very similar to the mean percentage bias and are available on request. The root-mean-squared error which encompasses both the mean bias and the variance of the estimator are reported in the working paper version of the paper.

20. In our experiments, all agents face the aggregate interest rate, and the lagged aggregate interest rate is a very strong instrument. In practice, there may be individual variation in the inter-temporal price which agents face which is not easily observable to the researcher. On the other hand, our experimental setup presumes aggregate exogenous shocks to interest rates. Part of the argument in Carroll (2001) is that in an equilibrium of a standard closed model, the interest rate is not exogenous. We nevertheless think that it is reasonable to proceed on the basis that there are shocks to aggregate interest rates that can be thought of as exogenous to typical consumers is a given country. Given the results in Table 5, we did experiment with the just identified model (with the lagged interest rate as only instrument). This gave very similar results and the details are available upon request.

21. Results with GMM are very simillar and available upon request.

22. Estimating D-I model on synthetic cohort data poses the problem that we cannot identify which agents carry forward zero wealth. These results are relying on an unreaslistic (best-case) scenario where the researcher is able to exclude constrained individuals from the sample. In Section 5.3, we examine the case in which we include the constrained individuals in synthetic cohort estimation.

23. See online appendix Table A6 for the exact numbers.

24. See online appendix for detailed discussion and results.

REFERENCES

Adams, A., L. Cherchye, B. De Rock, and E. Verriest (2014) Consume now or later? Time inconsistency, collective choice, and revealed preference. *American Economic Review* 104(12), 4147–4183.

Andrews, Donald W. K. and J. H. Stock (2005) Inference with Weak Instruments. Cowles Foundation discussion paper 1530.

Alan, S., O. Attanasio, and M. Browning (2009) Estimating Euler equations with noisy data: Two exact GMM estimators. *Journal of Applied Econometrics* 24, 309–324.

Alan, S and M. Browning (2010) Estimating intertemporal allocation parameters using simulated residual estimation. *Review of Economic Studies* 77, 1231–1261.

Alan, S., M. Browning, and M. Ejrneas (2016) Income and consumption: A micro semistructural analysis with pervasive heterogeneity. *Journal of Political Economy* (forthcoming).

Attanasio, O. P. and G. Weber (1993) Consumption growth, the interest rate, and aggregation. *Review of Economic Studies* 60, 631–649.

Attanasio, O. P. and G. Weber (1995) Is consumption growth consistent with intertemporal optimization? Evidence from the consumer expenditure survey. *Journal of Political Economy* 103(6), 1121–1157.

Attanasio, O. P., J. Banks, C. Meghir and 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 (2004) Estimating Euler equations. *Review of Economic Dynamics* 7(2), 405–435.

Benito, A. and H. Mumtaz (2009) Excess sensitivity, liquidity constraints, and the collateral role of housing. *Macroeconomic Dynamics* 13(3), 305–326.

Blundell, R., M. Browning, and C. Meghir (1994) Consumer demand and the life-cycle allocation of household expenditures. *Review of Economic Studies* 61, 57–80.

Browning, M., A. Deaton, and M. Irish (1985) A profitable approach to labor supply and commodity demands over the life-cycle. *Econometrica* 53, 503–544.

Bond, S. and C. Meghir (1994) Dynamic investment models and the Firm's financial policy. *Review of Economic Studies* 61(2), 197–222.

Carroll, C. (2001) Death to the log-linearized consumption Euler equation! (and very poor health to the second-order approximation). *Advances in Macroeconomics* 1(1), 1003–1003.

Chamberlain, G. (1984) Panel data. In Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics*, vol. II, pp. 1247–1318. Amsterdam: Elsevier North-Holland.

Crossley, T. F., H. Low, and M. Wakefield (2009) The economics of a temporary VAT cut. *Fiscal Studies* 30(1), 3–16.

Davidson, R. and J. G. MacKinnon (2004) *Econometric Theory and Methods*. New York: Oxford University Press.

Deaton, A. (1985) Panel data from time series of cross sections. *Journal of Econometrics* 30, 109–126.

Deaton, A. (1991) Saving and liquidity constraints, *Econometrica* 59(5), 1221–1248.

Dogra, K. and C. Gorbachev (2016) Consumption volatility, liquidity constraints and household welfare. *Economic Journal* 126, 2012–2037.

Druedahl J. and T. H. Jorgensen (2016) Estimating Dynamic Economic Models with Non-Parametric Heterogeneity. Mimeo.

Dynan, K. E. (1993) How prudent are consumers? *Journal of Political Economy* 101(6), 1104–1113.

Gayle, W. R. and N. Khorunzhina (2016) Micro-level estimation of optimal consumption choice with intertemporal nonseparability in preferences and measurement errors," *Journal of Business & Economic Statistics* (just-accepted).

Gomes, F. and J. Issler (2017) Testing consumption optimality using aggregate data. *Macroeconomic Dynamics* 21(5), 1119–1140.

Hahn, J., J. Hausman, and G. Kuersteiner (2004) Estimation with weak instruments: Accuracy of higher order bias and MSE approximations. *Econometrics Journal* 7(1), 272–306.

Hall, R. E. (1978) Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy* 86, 971–987.

Hansen, L .P. and K. J. Singleton (1982) Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica* 50(5), 1269–1286.

Huang, K., K. X. D. Liu, and J. C. Zhu (2015) Temptation and self-control: Some evidence and applications. *Journal of Money, Credit and Banking* 47, 581–615.

Ludvigson, S. and C. Paxson (2001) Approximation bias in linearized Euler equations. *Journal of Economics and Statistics* 83(2), 242–256.

Mazzocco, M. (2007) Household intertemporal behaviour: A collective characterization and a test of commitment. *Review of Economic Studies* 74(3), 857–895.

Mehra, R. and E. C. Prescott (1985) The equity premium: A puzzle. *Journal of Monetary Economics* 15(2), 145–161.

Moulton, B. (1986) Random group effects and the precision of regression estimates, *Journal of Econometrics* 32, 385–397.

Mudit K. and S. Ravi (2016) Elasticity of intertemporal substitution in consumption in the presence of inertia: Empirical evidence from a natural experiment. *Management Science* (forthcoming).

Mulligan, C. (2004) What do aggregate consumption Euler equations say about the capital-income tax burden? *American Economic Review* 94(2), 166–170.

Office of National Statistics (2017) Consumer trends, UK; Quarter 4(Oct to Dec) 2016, released 31 March 2017 and accessed at: https://www.ons.gov.uk/economy/nationalaccounts

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: Some evidence from panel data *Economic Letters* 14(1), 93–100.

Stock, J. H. and Yogo, M. (2005) Testing for weak instruments in linear IV regression. In D. W. K. Andrews and J. H. Stock (eds.), *Identification and Inference for Econometric Models. Essays in Honor of Thomas Rothenberg*, pp. 80–108. New York: Cambridge University Press.

Yogo, M. (2004) Estimating the elasticity of intertemporal substitution when instruments are weak. *The Review of Economics and Statistics* 86(3), 797–810.