UNIVERSITÀ DEGLI STUDI DI TORINO

Modelling Movements in Individual Consumption:

A Time Series Analysis of Grouped Data

By Orazio P. Attanasio and Margherita Borella[1]

University College London, UK, Institute for Fiscal Studies, UK, NBER, USA and CEPR, UK;

University of Torino, Italy, CeRP – Collegio Carlo Alberto, Italy, and Netspar, The Netherlands.

We characterize the time series properties of group-level consumption, income and interest rates using micro data. We relate the coefficients of MA representations to structural parameters of theoretical models of consumption behaviour. Using long time series of cross sections to construct synthetic panel data for the UK, we find that for high-educated individuals the Euler equation restrictions are not rejected, the elasticity of intertemporal substitution is higher than one and there is evidence of "excess smoothness" of consumption. Low-educated individuals, conversely, exhibit excess sensitivity of consumption to past income and elasticity of intertemporal substitution not statistically different from zero.

JEL: E2, D1, C3

Keywords: Consumption, Euler Equations, Life Cycle Model

---

1. Introduction

In the last 25 years, many empirical studies of consumption behaviour focused on some version of an Euler equation for intertemporal optimization and used the over-identifying restrictions implied by such an equation to estimate structural parameters and test the model. These studies, too numerous to be listed here, used both aggregate and individual level data, reaching different conclusions about the validity of the model and about the magnitude of the structural parameters that can be identified within such a framework. Most of the implications of such theoretical structure are restrictions on the time series properties of household consumption. Moreover, unless one is willing to make strong assumptions about the nature of intertemporal trades available to consumers (like that of complete markets), one can only obtain consistent estimates of the model's parameters from the orthogonality conditions implied by an Euler equation exploiting large-T asymptotics. Therefore, the absence of complete markets poses aggregation problems that make the use of aggregate data problematic and consistency problems for any econometric evidence based on micro data that do not cover a sufficiently long time horizon.

The results one obtains from an empirical study of consumption behaviour based on Euler equation estimation over a long time period (including parameter estimates and tests of the overidentifying restrictions of the model), ultimately depend on the time-series properties of consumption, income and interest rates. While much is known about the time-series properties of aggregate consumption in its relation to income, interest rates and other aggregate variables,[2] the same is not true for the time series properties of individual level consumption or even appropriately aggregated quantities. The latter point is important for the results one gets from Euler equation aggregation, as shown by Attanasio and Weber (1993).

In this paper, we propose a new methodology to analyse the time series properties of consumption expenditure, jointly with those of income and other (possibly aggregate) variables of interest, using a long time-series of household level data. Our approach fills an important gap in the existing literature. Individual longitudinal data have been used to study the time-series properties of hours and earnings by authors such as Lillard and Willis (1978), MaCurdy (1982), Abowd and Card (1989), Moffitt and Gottshalk (1995) and Meghir and Pistaferri (2004), among others. These studies model only labour market variables. Moreover, they focus on the dynamic properties of purely idiosyncratic components and treat aggregate shocks as nuisance parameters that are eliminated, together with deterministic life cycle effects, in preliminary regressions. Our exercise, instead, studies mainly consumption and in this respect constitutes an original contribution. In addition, the availability of a long time series of cross sections allows us to focus on business cycle fluctuations rather than removing aggregate shocks. We study how a number of household level and aggregate variables co-vary and estimate the parameters of a time series model that fits their behaviour over time. This is a second element of novelty of our exercise.

By constructing the theory-consistent possibly non-linear aggregates, we can avoid aggregation biases that have been shown to be important in the study of the dynamic properties of consumption (see Attanasio and Weber, 1993).

We represent the time series behaviour of group-level consumption, income and interest rates as an MA process whose parameters, given some assumptions, we can estimate. These parameters will summarize the time series properties of the variables we study. Given a set of

---

[2] The papers by Sargent (1978), Hall (1978), Flavin (1981), Blinder and Deaton (1985), Campbell (1987), West (1988), Campbell and Deaton (1989), Caballero (1990), Quah (1990) are only some examples of this literature.

identifying restrictions, we will be able to identify patterns of dynamic dependence and some contemporaneous correlations. Having characterized these properties of group-level consumption, income and interest rates, we map the patterns of correlations that emerge from the data to those implied by different theoretical models. In particular, we show how to relate the coefficients of our statistical model to several models that have been studied in the consumption literature. In this sense, our approach provides a unifying framework that can be used to assess and compare different pieces of evidence that have been accumulated in the literature. Moreover, our focus on specific properties of the time series processes for consumption and other variables of interest provides insight on the features of a model that might be able to fit the data when simpler versions are rejected.

We show what are the restrictions that the Euler equation for consumption imposes on our MA representation of grouped consumption, income and interest rates and we show how, if these restrictions are not violated, one can use estimates of the restricted MA process to estimate structural preference parameters, as is typically done in the Euler equation literature. Our procedure makes it explicit which time series features of the data lead to potential rejections of the over-identifying restrictions implied by Euler equations and which lead to specific estimates of preference parameters. Moreover, our approach allows us to identify aspects of individual preferences that, while not leading to violations of the orthogonality conditions implied by an Euler equation, might be important determinants of individual behaviour.

But our approach can also go beyond what is learned from the estimation of Euler equations. As is well known, the Euler equation is silent about contemporaneous correlations between consumption innovations and other sources of innovation to variables that are relevant for the dynamic problem solved by the consumer, namely income and interest rates. The Euler equation does not provide, without additional equations, a consumption function, which is a mapping from innovations to consumption. And yet our approach can identify some of the contemporaneous correlations and we can check whether these estimated correlations are consistent with different versions of the model. Our contribution is therefore also directly related to a strand of the literature on permanent income with rational expectations that has looked at the implications of the life cycle/permanent income model for the time series behaviour of consumption and income. Some of the relevant papers in this literature include Sargent (1978), Flavin (1981), Campbell (1987) and Campbell and Deaton (1989), West (1988), Hansen, Roberds and Sargent (1991). These papers worked with versions of the model that would deliver a closed-form solution for consumption as a function of permanent income and test the restrictions implied by this closed-form solution on the bi-variate representations of consumption and income. Similar results, however, can be obtained by log-linearization.

Many of the papers in this literature concluded that consumption is 'excessively smooth', in that consumption does not react enough to permanent innovations to income. We pursue, using micro data, a similar strategy. In particular, we study the restrictions imposed by the life cycle model on the parameters of our time series model that govern the contemporaneous correlation between income and consumption. Following Attanasio and Pavoni (2011), we stress that we can have situations in which the Euler equation for consumption is satisfied and there is no evidence of 'excess sensitivity' of consumption and yet one can observe that consumption does not move enough in reaction to changes in permanent income. Attanasio and Pavoni (2011) stress how this excess smoothness is related to the violation of the intertemporal budget constraint *with a single asset* and the presence of insurance of idiosyncratic shocks.

Our approach, therefore, can be seen as giving a unifying framework that uses simultaneously the implications of the data covariance structure to estimate structural parameters

and test the orthogonality restrictions implied by the model (as in the Euler equation approach), and studying how innovations to income are reflected (or not) in changes in consumption.

The use of large-T asymptotics is an important distinguishing feature of our approach. While we remove deterministic trends, we do not remove business cycle aggregate shocks. Indeed, our approach can be described as an attempt to model these aggregate shocks and is therefore based on large-T asymptotics.[3] Without relying on large-T asymptotics, our approach would not be able to encompass the Euler equation approach, which, in general, cannot identify the parameters of interest without a long time horizon.

Relying on large-T asymptotics implicitly forces us to take a stand on what is a 'sufficiently' long time period. As always with this type of issues, it is difficult to give an explicit and general answer. The need for a long time period arises from the necessity to average out expectational errors over time. If the time period, therefore, covers several business cycles (as is in the case in our application), one should feel reassured. The Montecarlo exercises performed in Attanasio and Low (2002) indicate that quarterly data over a 20 or 30 years horizon should be sufficient to identify the parameters of interests with an Euler equation with sufficient precision.

A small, but increasing, number of papers emphasizes the presence of heterogeneity in preference parameters, which can explain a number of regularities. Blundell, Browning and Meghir (1994), Attanasio and Browning (1995), Crossley and Low (2011) and Alan and Browning (2010) all find a non-constant elasticity of intertemporal substitution. An advantage of using household-level data is that they enable us to stratify our sample according not only to the year of birth of the household head, as is typically done in the studies using pseudo panels, but also to some measure of lifetime wealth, such as the education level, thus allowing for an important source of heterogeneity to have an impact on preference parameters.[4]

The data on which we apply our approach are from the UK Family Expenditure Survey. The FES is, to the best of our knowledge, the longest time series of cross sections containing exhaustive and detailed information on consumption, its components and several other variables of interest. Unfortunately, the data do not have a longitudinal dimension, as each individual household appears only once. To study the dynamics of our system, therefore, we rely on synthetic cohort techniques, of the type proposed by Deaton (1985). As the theoretical model on the background of our analysis is the life cycle model, it is natural to study consumption in its relation to age and to divide the sample to form year of birth cohorts of individuals that are followed over time. To capture differences in covariance structures (and hence preference parameters) among different groups of the population we also form groups according to the education levels. We use surveys from 1974 to 2000 to form quarterly observations of consumption and income at the cohort level.

The lack of a longitudinal component to our data imposes an important limitation to our work, in that we are forced to ignore pure idiosyncratic variability and to focus instead on the dynamics of group averages. In this sense, the group shocks become really the focus of our

---

[3] Another strand of the literature has focused on the time series properties of disaggregated business cycles. These studies, including Watson and Engle (1983), Quah and Sargent (1994), Forni and Reichlin (1996), focus on the time series properties and aim at characterizing the number of common factors and modeling their dynamic effects on the sectors considered. Another set of papers that are related to ours are those by Cunha et al. (2004) and Cunha et al. (2005) as well as Blundell et al. (2004). In these papers, the authors use schooling, income and consumption data to identify the shocks that inform individual choices.

[4] It could be argued that education levels are choices that individuals make and should not be considered as exogenous features bestowed upon individuals. This is certainly true. However, as we discuss below, under certain conditions, one can model consumption behavior *once the education choices are made*.

empirical study. It should be stressed once more, however, that even if we are forced to aggregate the data, we can control the aggregation directly and construct the non-linear transformations of the data implied by the theory before aggregating them. Having said this, however, it is clear that, should longitudinal data on consumption be available, an interesting extension of our study would be to estimate our model on such data. Such an extension would allow the possibility of identifying the dynamics of purely idiosyncratic shocks.

We find that, at the cohort level, the restrictions implied by the Euler equation are not rejected in our data. In particular we find no evidence of 'excess sensitivity' of consumption to predictable changes in income. When we form finer groups according to the education level, we find substantial differences between households whose heads completed at most compulsory schooling and households whose heads attained higher education levels. For the latter groups, we still do not find evidence of excess sensitivity of consumption to income. The estimated elasticity of intertemporal substitution for higher educated households is higher than (but not statistically different from) one and we do find evidence of "excess smoothness". For the less educated group (that is, households heads with at most compulsory school completed) results are very different: we do find evidence of excess sensitivity of consumption, and our estimates of the elasticity of intertemporal substitution are close to (and not statistically different from) zero.

The rest of this paper is organized as follows. In section 2 we present our statistical model and discuss its identification and estimation. In section 3 we present a simple theoretical framework that can be used to interpret the results. We also show how to derive the so-called 'excess sensitivity' and 'excess smoothness' tests. After describing the data sources in section 4, we discuss the estimation procedure in section 5. In section 6 we report the estimates of our statistical models and in section 7 we show what interpretation they have in terms of the models presented in section 3. Section 8 concludes the paper.

2.   The Methodology

As stressed in the introduction, the main aims of this study are two. First, we want to identify innovations to consumption and other variables that might be interpreted as the determinants of consumption and model their covariance structure. Second, we want to use this covariance structure to shed some light on the plausibility of alternative theoretical models of consumption. As the micro data we work with lack any longitudinal dimension, we are forced to use average grouped data to estimate any dynamic model. This means that we are not able to model idiosyncratic persistence, but only persistence at the group level. The advantage of using grouped data lies in their longitudinal (or time) dimension, which allows to identify the time series moments of the variables of interest.

Notice that we form group exclusively because of the need to have a longitudinal dimension. The assumption that is made on the groups, to be able to use the econometric approach we adopt, is that their membership is constant. This allows us to follow over time groups of 'similar' individuals. Once groups are formed, one can estimate the time series properties for each group, allowing heterogeneous parameters, or impose that the parameters are the same across groups. In what follows, we consider two criteria to form groups: the household head year of birth and educational attainment. The former variable is clearly fixed and exogenous.[5]

---

[5] Although one could think of reasons why the fixed membership is violated for year of birth cohorts: differential mortality is one obvious problem, which we will avoid by excluding individuals at the end of the life cycle.

In the case of education, it can be argued that it is an endogenous choice. However, as long as we are willing to assume that the educational status does not change after a certain age, this is still a valid criteria to form groups and to estimate heterogeneous parameters for the members of these groups. Of course, the variation of parameters across education groups should not be given a causal interpretation: it is not education that would make a certain individual more or less risk averse or more or less patient. On the contrary, it is quite possible that individuals with certain preferences choose a certain level of education. However, one can proceed and estimate our statistical model and map it into structural parameters *for that particular group*.

Our model is effectively a model of group shocks. However, as we construct group averages from our micro sample, the variables we observe are affected by two sources of variability. On the one hand, we have genuine group specific shocks. On the other, as we do not observe group population means but estimate them by our sample averages, our data are affected by "measurement error" or, more precisely, sampling variation, arising from the limited size of our sample. Moreover, as we do not observe consumption growth at the individual level, but estimate averages of levels before taking time differences, the "measurement error" we are discussing induces an MA(1) structure into our system. In this setting, this type of variability constitutes a nuisance which can, however, be controlled for given the information on the sample structure and given the information on within cells variability (see Deaton, 1985).

In this section we sketch the main features of our approach. In particular, we write down the statistical model that we estimate and discuss the identifying assumptions we make.

2.1.   *The statistical model*

In what follows we consider three observable variables, x_{hgt}, y_{hgt} and r_t, where the index h denotes the individual household, the index t the time period and the index g the group (assumed to be of fixed membership in the population) to which household h belongs. For simplicity, we sketch the model for three univariate variables, but it can be easily extended to three vectors of variables.

The x variables represent choice variables for the individual household. In the basic specification we estimate, x is non-durable consumption. In general, however, by adding more equations to the system, x could include several variables such as, for instance, different components of consumption, hours of work and participation rates. These variables are affected by all the shocks present in the system.

The y variables are household (or group) specific but are assumed not to be determined by individual choice, at least at the frequency we are considering. In the specific model we estimate, we consider y to be income, while in more general models it could include, for instance, wage rates for male and females. The y variables are affected by all the shocks in the system with the exception of the shocks specific to the variables in the first group (the choice variables), which can be thought of as taste shocks.

The r variables are also not determined by individual choices and, moreover, do not vary across individuals or cohorts. We think of this kind of variables as prices. In the model we estimate below, we have a single variable of this type, the interest rate, but we could consider many, such as the prices of individual commodities.

To take into account the possibility of stochastic trends (and because of the structural interpretation we give to the covariance structure that we estimate), we take the first differences of the household specific variables. We test for the presence of unit roots and we mention this evidence below. The statistical representation we propose to estimate is the following:

$$\Delta x_{hgt} = \sum_{j=0}^{q_{xx}} \alpha_{jg}^{xx} u_{hgt-j}^{x} + \sum_{j=0}^{q_{xy}} \alpha_{jg}^{xy} u_{hgt-j}^{y} + \sum_{j=0}^{q_{xr}} \alpha_{jg}^{xr} u_{t-j}^{r}$$

$$\text{(1)} \qquad \Delta y_{hgt} = \sum_{j=0}^{q_{yy}} \alpha_{jg}^{yy} u_{hgt-j}^{y} + \sum_{j=0}^{q_{yr}} \alpha_{jg}^{yr} u_{t-j}^{r}$$

$$r_{t} = \sum_{j=0}^{q_{rr}} \alpha_{jg}^{rr} u_{t-j}^{r}$$

where the shocks u_{hgt}^{x}, u_{hgt}^{y} and u_{t}^{r} are assumed to be uncorrelated over time and with a diagonal variance-covariance matrix at each time t.

Notice that the variables in (1) can be non-linear transformations of the variables of interest. Although in the absence of longitudinal individual data it will not be possible to estimate (1) directly and we will be forced to aggregate it over groups, as we work with repeated cross sectional data, we can control the aggregation directly. This is important because theoretical models often imply relationships among non-linear functions of variables.

We adopt the convention that a generic variable (observable or unobservable) z_{gt} is the population group average of z_{hgt}, that is $z_{gt} \equiv E_{gt}[z_{hgt}]$. Taking population averages over all h belonging to group g of system (1) we can re-write it without the subscript h:

$$\Delta x_{gt} = \sum_{j=0}^{q_{xx}} \alpha_{jg}^{xx} u_{gt-j}^{x} + \sum_{j=0}^{q_{xy}} \alpha_{jg}^{xy} u_{gt-j}^{y} + \sum_{j=0}^{q_{xr}} \alpha_{jg}^{xr} u_{t-j}^{r}$$

$$\text{(2)} \qquad \Delta y_{gt} = \sum_{j=0}^{q_{yy}} \alpha_{jg}^{yy} u_{gt-j}^{y} + \sum_{j=0}^{q_{yr}} \alpha_{jg}^{yr} u_{t-j}^{r}$$

$$r_{t} = \sum_{j=0}^{q_{rr}} \alpha_{jg}^{rr} u_{t-j}^{r}$$

The variables on the left-hand side of system (2) are not observable: we do not have data on the entire population, but, at each time period, we have samples drawn from the population. We can therefore produce estimates of these variables using the averages on samples that in each time period are composed of different individuals. A consequence of the fact that our samples are of limited size is that the variables that we can use as estimates of the left-hand-side variables of system (2) will vary not only because of the presence of the shocks but also because of the variability of our samples. However, the effect that this type of variability has on the total variability and on the autocorrelation properties of our variables of interest can be easily controlled for by using the information on the within cell variability and on cell size. To be specific, if we denote with $\tilde{z}_{gt}$ the sample equivalent of the population mean z_{gt} we have:

$$\text{(3)} \qquad \Delta \tilde{z}_{gt} = \Delta z_{gt} + \Delta \eta_{gt}^{z} \qquad\qquad \text{with } z = x, y$$

where η_{gt}^{z} reflects the difference between population and sample group averages. The variance of η_{gt}^{z}, which we assume to be independent of the changes in the innovations we are interested in modelling, can be estimated consistently from the within cell variability and the cell size. Hence system (2) can be written as:

$$(4) \quad \Delta \tilde{x}_{gt} = \sum_{j=0}^{q_{xx}} \alpha_{jg}^{xx} u_{gt-j}^{x} + \sum_{j=0}^{q_{xy}} \alpha_{jg}^{xy} u_{gt-j}^{y} + \sum_{j=0}^{q_{xr}} \alpha_{jg}^{xr} u_{t-j}^{r} + \eta_{gt}^{x} - \eta_{gt-1}^{x}$$

$$\Delta \tilde{y}_{gt} = \qquad\qquad \sum_{j=0}^{q_{yy}} \alpha_{jg}^{yy} u_{gt-j}^{y} + \sum_{j=0}^{q_{yr}} \alpha_{jg}^{yr} u_{t-j}^{r} + \eta_{gt}^{y} - \eta_{gt-1}^{y}$$

$$\tilde{r}_{t} = \qquad\qquad\qquad\qquad\qquad \sum_{j=0}^{q_{rr}} \alpha_{jg}^{rr} u_{t-j}^{r}$$

As mentioned above, the terms η_{gt}^{z} (with $z = x, y$) represent the difference between sample and population means induced by limited sample sizes. We assume that this type of 'measurement error' (or unobserved heterogeneity) present in the group specific variables at time t is orthogonal to the unobservable random variables u. We can easily compute the time-varying variances of these terms using our micro data, and estimate system (4) conditionally on them. We discuss in detail the estimation methods we use in section 5.

Before estimating the group averages in (4), we remove all deterministic trends from the variables of interest. These are likely to be affected by time, group and age effects. Within the framework of a life cycle model, natural groups to consider are those formed on the basis of the year of birth of the household head. Therefore, group effects are essentially cohort effects. As year of birth, age and time are linked by a linear relationship, it is impossible to disentangle them. In what follows we label all deterministic trends in the data as 'age and cohort' effects and remove them by regressing the variables of interests on cohort dummies and a polynomial in age.[6]

It should be stressed that our first step regressions do not include time dummies: as business cycle shocks (either common across groups or not) are the focus of the study, we do not want to remove them. Removing group specific age polynomials, however, does remove deterministic trends from our data, which may, but do not need to, be interpreted as a combination of age and cohort effects. What we do not remove are business cycle fluctuations that are instead typically removed in studies that introduce time dummies in the first step regressions and focus on idiosyncratic shocks. We should stress that the fact that our pseudo-panel is unbalanced and that we only follow two cohorts does not constitute an issue as we have a relatively large number of time periods for *each* cohort.

Having estimated the parameters of such a regression, we interpret the group averages of the regression residuals as deviations of the average group data from the deterministic trends present in the data. It is the group sample averages of these residuals, which we denote with $\tilde{z}_{gt}$, that we model and study.

As stressed above, because we lack panel data, we can only study dynamic models by using grouped observations. That means that we can only study the dynamics of the averages of the generic variables $\tilde{z}_{gt}$. Any purely idiosyncratic persistence, embedded in η_{gt}^{z}, cannot be recovered by our methodology. This might be a serious problem in evaluating the importance of precautionary savings or similar phenomena. Nothing much can be done about this except

---

[6] While this label is arbitrary, disentangling the various effects is not the aim of this study that focuses, instead, on modelling the innovations to income and consumption at the business cycle frequency. Strictly speaking non linear age effects can be identified from time and cohort effects. However, we would not know how to interpret them.

noticing that to get a handle on persistence at the individual level, a genuine panel dimension is needed as it will be necessary to observe the covariance between individual variables in subsequent time periods.[7]

The MA structure in equation (4) is reasonably flexible and general, but it does impose some important and strong restrictions that are used to identify the model. In particular, for the contemporaneous effect of shocks, we need to impose the triangular structure described above for identification. In particular, we assume that shocks to individually determined variables, such as non-durable consumption, do not affect the variables that are assumed to be given to the individual households. This assumption is standard in the multivariate VAR literature. In our context, it can be justified by interpreting shocks to consumption as taste shocks which, therefore, would not affect the budget constraints, and, therefore income. Innovations to income, instead, should affect (via the budget constraint) consumption. As we discuss in the next section, the way in which these innovations enter consumption, however, is dictated by the model – which gives rise to the excess smoothness tests.

It should also be stressed that we wrote system (4) imposing the triangular restriction also on the lags of the relevant shocks, but only for ease of notation. This specific restriction is not necessary for identification. As we discuss in the next section, theory imposes several (over-identification) restrictions on the lag coefficients, which we will exploit to give a structural interpretation to our time series model. Because the triangular restrictions on lagged coefficients are not necessary for identification (and indeed are not imposed in estimation), all our results on the estimation of structural parameters and on tests of excess sensitivity are independent of the specific assumption which is made for the identification of the contemporaneous correlations.[8]

In addition to these restrictions, we make the normalization assumption that the coefficients on the own residuals are equal to one. For instance, we assume that α^{xx}_0 is equal to one.

3. A structural interpretation of the statistical model.

The starting point for a structural interpretation of the parameters in system (4) is the life cycle model, interpreted as a flexible parameterization of a dynamic optimization problem in which the decision unit is the household.

We start with the simplest version of the life cycle model as an example of a way in which a theoretical framework can be used to impose restrictions on the parameters of model (4). We then complicate the model to introduce a number of realistic elements. We neglect deterministic trends (including age effects) as well as family size effects from all the variables in our analysis. This implicitly assumes that the age and family size effects removed in the first step of our estimation procedure capture completely the effect of demographic variables and that these are

---

[7] Having said that, however, it should be stressed that even a relatively short genuine panel might be sufficient to estimate a model which requires large-T asymptotics if repeated panels are available: one can then group the cross moments and follow their dynamics over time.

[8] The restrictions on the lagged covariances used to identify the structural taste parameters and test the restrictions of the model are equivalent to the orthogonality conditions used by an Euler equations for the same purpose. On the other hand, the results on the excess smoothness that we discuss below, which depend on the identification of contemporaneous correlations (or on the way income innovations are translated into consumption innovations), may be affected by the specific normalization used.

considered as deterministic. As we focus on business cycle frequencies, we do not think that this assumption is particularly strong.[9]

3.1. A simple version of the life cycle model

A very simple version of the life cycle model implies the following system of equations for a generic individual:

$$(5) \qquad \log(\lambda_t) = E_t[\log(\lambda_{t+1})] + E_t(r_{t+1}) + k_t + \varepsilon_{t+1}$$

$$(6) \qquad \log(\lambda_t) = \log(U_x(x_t, \varphi_t))$$

where the variable k_t is a function of the discount factor and of higher moments of the expectational error ε_{t+1}. U_x is the marginal utility of (non-durable) consumption, which is assumed to depend on consumption and a vector of observable and unobservable variables φ_t. λ_t is the marginal utility of wealth and represents the effect of all present and future variables relevant for the optimization problem faced by the individual. r_{t+1} is the interest rate. The specification in equations (5) and (6) also assumes intertemporal separability, in that the marginal utility of consumption at t does not depend on variables from other time periods.

If one considers the fact that equations (5) and (6) refer to a single generic household, it is clear why, even in such a simple framework, aggregating such an equation across groups of households would generate group specific fixed effects. These could arise if, for instance, there are systematic differences across groups in the discount factors, higher moments of the expectational errors, or in the unobserved components of φ_t. Indeed, our results disaggregated by education level will show these effects to be important.

If we assume that the interest rate and k_t are constant over time and that the observable component of the vector φ_t contains only deterministic variables that can be captured by the deterministic trends removed in our first step, equations (5) and (6) have very simple and strong implications for the model in (4).

First, one can simplify the model considerably eliminating the last equation (that refers to the interest rate). Furthermore, if the specification of the utility function is such that the marginal utility can be approximated by a linear function of log consumption (as it is the case, for instance, for a CRRA utility function), from equations (5) and (6) one can see that changes in log consumption can be related to the expectational error ε_{t+1} and therefore should not exhibit any serial correlation and should be uncorrelated with any information available at t. This is the celebrated 'random walk'[10] result, stressed by Hall (1978). In our model, it translates into restrictions on the coefficients on lags of all shocks of the consumption equation. The so-called

---

[9] If one thinks that the age polynomials used in the first step are not sufficient to remove the effect of demographic variables, and is willing to retain the assumption that they are deterministic, these variables can be used in the first step regressions.

[10] Or martingale, to be precise.

tests of 'excess sensitivity' of consumption to predictable components of income take the form in our model of tests on the lagged income shocks being good predictors of consumption.

3.2.   The elasticity of intertemporal substitution

A first and very important generalization of the model is to consider time varying interest rates. This extension is of particular interest as it allows one to estimate the elasticity of intertemporal substitution.

Allowing for a time variable interest rate involves including the third equation in model (4), so that one can measure the correlation between innovations to interest rates and consumption (and other variables)[11]. If we consider an asset whose rate of return is the same across groups and that is widely held, than equations (5) and (6) induce a set of additional restrictions on system (4). Let's define $A_g^{xx}(L) = \sum_{j=1}^{q} \alpha_{jg}^{xx} L^j$, and analogously for $A_g^{xr}(L)$ and $A_g^{rr}(L)$. Substituting equation (6) into (5), one can see that the theoretical model implies a relationship between the expected value of (log) marginal utility of consumption and the interest rate. If the utility function is isoleastic with a coefficient of relative risk aversion γ, the log marginal utility of consumption will be proportional to log consumption so that equation (5) and (6) impose the following restriction on system (4):

$$(7) \qquad \gamma_g (A_g^{xx}(L)u_{gt+1}^x + A_g^{rx}(L)u_{gt+1}^r) = A_g^{rr}(L)u_{gt+1}^r.$$

As this has to hold for every possible realization of the residuals, the restrictions on the coefficients of system (4) are that:

$$(8) \qquad A_g^{xx}(L) = 0 \text{ and}$$

$$(9) \qquad \gamma_g A_g^{xr}(L) = A_g^{rr}(L).$$

The second set of restrictions implies that, as long as the interest rate is predictable, one can identify the coefficient of relative risk aversion.

Should one encounter a rejection of these restrictions, several alternative specifications are possible depending on the nature of the rejection. If a significant coefficient on the lagged 'consumption shocks' in the consumption equation were found, a possible explanation would be the possibility of an unobservable component in the vector of preference shifters φ_t. Such a component, which, for lack of a better term, we label 'unobserved heterogeneity', captures those aspects of preferences that are not directly modeled and that are likely to be important for consumption. The time series properties of consumption innovations would then be clearly affected by the time series properties of such a term. The fact that the restrictions about the proportionality of the coefficients on the interest rate lagged innovations is violated might be an

---

[11] One can either assume that the interest rate is the same for all groups or allow for differences in intertemporal prices induced, for instance, by differences in marginal tax rates across groups. The latter approach, however, involves the necessity of measuring group specific interest rates.

indication of differences in interest rates and/or risk aversion across groups. If one more lag in the interest rate innovation enters the system even this possibility can be tested against more general misspecifications.

3.3. *Excess smoothness*

The restrictions we have discussed so far are derived from the orthogonality conditions implied by the Euler equation for consumption derived from equations (5) and (6). These conditions, together with a set of intertemporal budget constraints (and initial and terminal conditions for assets) pin down the allocation of consumption over the life cycle. When it is possible to derive a closed form solution for consumption (as is the case, for example, with quadratic utility and constant interest rates) then the solution imposes restrictions on the coefficients of system (4) that relate income shocks to consumption. When a closed form solution for consumption that pins down the relationship between income and consumption innovations is not available, one can rely on approximate solutions, of the type developed by Campbell (1994) and used, among others, by Blundell, Pistaferri and Preston (2008) and Attanasio and Pavoni (2011).

The restrictions that the life cycle/permanent income model imposes on the contemporaneous correlation between consumption and income relates to the fact that consumption should react to news about permanent income in a way that is mediated by the intertemporal budget constraint and that depends on the information that current income shocks give about future income. These are the type of restrictions that were studied by Flavin (1981), Campbell and Deaton (1989), West (1988), Quah (1990) and Hansen, Roberds and Sargent (1991) (HRS) among others. HRS, in particular, stress that given the Euler equation, the intertemporal budget constraint imposes testable restrictions on the response of the (change of) non-durable consumption to shocks to income whose violation has been interpreted as 'excess smoothness' of consumption. It is worth comparing the HRS approach to the specification we have proposed.

HRS show that, in a simple version of the permanent income model, the model gives rise to a representation of the following type:

$$\Delta y_{gt} = w_{gt}^y + \beta_{1g}^{yy} w_{gt-1}^y + \beta_{0g}^{yx} w_{gt}^x + \beta_{1g}^{yx} w_{gt-1}^x$$

$$\Delta x_{gt} = \qquad\qquad\qquad w_{gt}^x + \beta_{1g}^{xx} w_{gt-1}^x$$

(10)

$$E[w_{gt}^y] = E[w_{gt}^y w_{gt-1}^y] = E[w_{gt}^x] = E[w_{gt}^x w_{gt-1}^x] = E[w_{gt}^y w_{gt}^x] = 0$$

$$Var(w_{gt}^y) = \sigma_{wy}^2; \;\; Var(w_{gt}^x) = \sigma_{wx}^2$$

HRS stress that an implication of the theory is that this representation is *not* a Wold representation for the joint time series of consumption and income. Such a representation, assuming that consumption is a martingale (imposing the Euler equation) can be used to test the restrictions imposed by the intertemporal budget constraint. The test proposed does not require the specification of the information set observed by the consumer.

The structure of this representation is quite similar to ours, except that we have the opposite triangular structure, that is in the system (4) consumption is allowed to depend on all

shocks, while income is not allowed to depend on the consumption shock. Under special circumstances, the two representations are equivalent. If, for instance, all the lag coefficients in our consumption equation are zero, one can map one specification into another. We need such a restriction so that, in our simple model, consumption is a martingale. However, in more general circumstances in which the HRS specification allows for lags in the consumption equation (maybe originated by temporal non-separabilities), the two specifications impose different restrictions on the data.

In a recent contribution, Attanasio and Pavoni (2011) discuss how violations of the intertemporal budget constraint can arise in a situation in which the Euler equation is satisfied but consumption is partly insured in a model with moral hazard and hidden assets. In particular, Attanasio and Pavoni (2011) stress the difference between the restrictions that imply the lack of correlation between predicted income and predicted consumption and the restrictions that involve the contemporaneous correlation of income and consumption. The latter can arise even with an Euler equation holding, if the intertemporal budget constraint *with a single asset* is violated, maybe because it ignores state contingent transfers that insure part of permanent shocks.

In our model, the intertemporal budget constraint with a constant interest rate (and the Euler equation) implies that:[12]

$$(11) \qquad \sum_{j=0}^{q} \alpha_{jg}^{xy} R^{j} \equiv \alpha^{xy}(R) = \alpha^{yy}(R) \equiv \sum_{j=0}^{q} \alpha_{jg}^{yy} R^{j}$$

where $R = (1/1+r)$ and r is the interest rate. Campbell (1987), West (1988), Campbell and Deaton (1989) and Hansen, Roberds and Sargent (1991) report results on versions of this test obtained from aggregate time series on income, consumption and saving that imply that consumption responds too little to innovations in income, a result that has been labelled the excess smoothness of consumption. In our framework, this result would imply that the left-hand-side of (11) would be less than the right hand side.

4.   Data

The data used in the estimation are drawn from the UK Family Expenditure Survey (FES) from 1978, first quarter, to 2005, first quarter. In this survey, about 8,000 families in the UK are interviewed each year and they are asked to fill diaries in which they record all the expenditures they make for two weeks. The survey records also information on demographic and labour supply variables for each member of the family. From these figures, it is possible to reconstruct total family income. We use these data to construct consumption and income grouped data at a quarterly frequency.

The FES has been widely used in the research on consumption. It has several advantages, among which we should mention the fact that it is available over a long time period, which, as we discussed above, is crucial for our identification strategy. Moreover, the quality of the data seems, at least until very recently, very high. Tanner (1998) and more recently Brewer et al. (2006) show

---

[12] In some of our models the coefficient on lagged income shocks are zero. This implies that consumption is a martingale. Hansen Roberds and Sargent (1991) stress that without imposing the martingale property implied by the Euler equation for consumption the intertemporal budget constraint does not impose restrictions on the time series properties of savings and consumption.

that, when aggregated, the FES reproduces closely the dynamics of National Account consumption data.

We select a sub-sample of the FES. In particular, we select all married or cohabitating couples, living in England, Scotland or Wales; we include all households whose head is active in the labour market but exclude those who are self-employed. This exclusion is motivated by the fact that both consumption and income measures might be problematic for household headed by a self-employed individual. The implicit assumptions that is made here is that self-employment status is constant over time or that transition in and out of self employment are not related to other variables of interest.

We further select families in order to define two seven-year-of-birth groups: the first cohort consists of families whose head was born between 1947 and 1953, while the second one whose head was born between 1954 and 1960. Estimates for the second (and younger) cohort are carried out starting from 1980, first quarter. We did not define cohorts using a narrower interval (say, 5 years) because we need samples large enough when creating cohort means to capture meaningfully the variability of shocks at the quarterly level. We only considered these two particular cohorts because of the need to use a time period that is long enough, as discussed in Section 5. We have a total of 13,512 observations for cohort 1, and 11,095 for cohort 2. Table 1 reports more details.

<Table 1 around here>

The variables we use in estimation are: non-durable consumption and disposable income, as well as prices for non-durable consumption computed using the weights available from the FES. Non-durable consumption is defined as the sum of: food, alcohol and tobacco, fuel, clothing, transportation costs and services. Disposable income is measured at the household level and is defined as the sum of net earnings and benefits.

In what follows, we present two sets of results. In the first set, groups coincide with year of birth cohorts. In the second, we also use the education level of the head of the households to define finer groups: hence each cohort is further divided into two education groups, "compulsory" and "more than compulsory" schooling. This definition of groups deserves a discussion. At some point in the life cycle education is a choice, which depends on individual abilities and tastes, as well as, possibly, resources. As seen from the beginning of the life cycle, therefore, education and consumption should be considered as jointly determined. However, if one assumes that at some point education choices are made and cannot be reversed, one can model consumption from that point onwards. We can therefore model the consumption of individuals with different levels of education and consider groups defined not only by year of birth cohorts but also by education achievements. Groups of households with different levels of education will command different levels of resources, will face income processes with different levels of risk and variability, might have access to different asset markets and might have different taste parameters. The choice of education early in life might be driven by all these factors (for instance, in a Roy model, individuals with comparative advantage in activities that require a college degree will select in the high education choice). Of the various factors listed only differences in tastes might create problems for our strategy. If we assume that choices are driven by factors other than tastes (such as ability or parental resources), our strategy is fine, as we will be following the behaviour of two different groups with constant membership and characterize their consumption behaviour. If taste parameters, for instance risk aversion and discount factors, are heterogenous in the population, we might have problems with our entire strategy because of aggregation issues (and not only when we divide our sample in different education groups). Our strategy of forming education groups, however, would still be correct if education choices sort

the population in two groups whose preferences are homogenous. For instance, if high-risk aversion individuals select into low education, then our procedure of forming groups by education would yield consistent estimates, while considering year of birth groups and ignoring education groups would actually be problematic.

In order to remove deterministic trends and seasonal effects, for each cohort/education group, using individual level data, we regress (the logarithm of) each variable of interest on a second order polynomial in age and on quarterly dummies, as well as on the family-specific McClements equivalence scale.

$$(12) \qquad z_{hgt} = \delta_g + \sum_{i=1}^{3} \alpha_{gi} q_{gi} + f^g(t-g) + \beta_g \cdot \Phi_{hgt} + \tilde{z}_{hgt}$$

where $f^g(t-g)$ is the group specific polynomial in age (obtained as time minus year of birth), q's are quarterly seasonal dummies, Φ_{hgt} is a vector of household specific variables including the McClements equivalence scale, a dummy for married couples and a dummy for the head of the household being an employee. δ_g are the group specific intercepts. Equation (12) is estimated by OLS[13]. The implicit assumption here is that the consumption and income innovations $\tilde{z}_{hgt}$ are uncorrelated with the variables on the right-hand side of equation (12), such as age, seasonal dummies and the family composition we use. The cohort average of the residual, $\tilde{z}_{gt}$, reflects both genuine time variation in group averages and measurement error arising from the limited sample sizes in computing the group averages, as we discuss in the next section.

Figures 1 and 2 show the time series of the cohort average residuals obtained from the preliminary regressions for (log) consumption and (log) disposable income. The series for the first cohort starts in 1978, first quarter, while for the second and younger cohort, represented with a dashed line in both graphs, it starts in 1980, first quarter. Both series end in 2005, first quarter. Both for income and consumption the series are quite volatile. In the case of income, the two series are positively and significantly correlated (correlation coefficient = 0.38, p-value 0.0001). In the case of consumption, the correlation is positive but much smaller at 0.1379 (p-value 0.17).

<*Figures 1 and 2 around here*>

Having obtained a cohort/education level measure of consumption and income depurated from deterministic trends, we conduct a preliminary investigation of their time series properties. In particular, we perform standard time series unit root tests such as the GLS Dickey-Fuller and the Kwiatkowski, Phillips, Schmidt, Shin (KPSS) tests. They give indication of the existence of a unit root in all the series considered, i.e. non-durable expenditure and disposable income for all groups considered.[14] This justifies our decision to model the first differences of these series.

---

[13] More efficient estimates could be obtained by controlling for the heteroskedasticity induced by different cell sizes and within cell variances.

[14] Results are reported in table A1 in the Appendix. For consumption and income, we perform the test on the aggregated data without decomposing the shocks from the variation induced by the sample variation. However, as the samples from which the aggregates are constructed are independent, this last source of variation (the η_{gt} in equation 4) is independent over time (as it represents the difference between the population and sample mean of a given variable in a given quarter). This implies that this component does not have a unit root. Our evidence of unit

The interest rate used in the estimation is the 3-month treasury bill rate, from which we subtract the inflation rate in the consumer price index constructed from the FES, to obtain a real rate. The real interest rate has also been de-trended, using the Hodrick-Prescott filter with a smoother set to 1600 as we deal with quarterly data.[15] Unit roots test on the interest rate indicate that this variable is I(0).

5. Estimation

There are several ways in which one can estimate the model (4). By making assumptions on the distribution of the shocks that enter the system (4), it is possible to compute the likelihood function associated with a given sample and estimate the parameters of interest by maximizing such a function. Alternatively, it is possible to avoid making specific functional form assumptions, and use a method of moment estimator. In particular, one could compute variances, covariances and autocovariances of the series of interest and minimize the distance between the sample moments and those implied by the parameters of the model.

While the latter method is potentially more attractive, we experienced a variety of numerical problems in its implementation, when trying to identify separately the parameters of the sampling error process – $Var_{gt}\left(\eta_{gt}^x\right)$ and $Var_{gt}\left(\eta_{gt}^y\right)$ – and the parameters of interest from an economic point of view ($\alpha_{jg}^{xx}, \alpha_{jg}^{xy}, \alpha_{jg}^{yy}$, and the variances of the shocks u_g^x and u_g^y). These problems, as we illustrate below, arise from working with relatively small samples while trying to disentangle two different processes (and are not necessarily a sign of mis-specification).

The difficulties we had with the method of moments, when considering explicitly the measurement error structure implied by the sampling error, can be illustrated even in a very simple model. Let's consider a univariate version of (4) for consumption with just one lag.

$$(13) \qquad \Delta\tilde{x}_{gt} = u_{gt}^x + \alpha_{1g}^{xx}u_{gt-1}^x + \eta_{gt}^x - \eta_{gt-1}^x.$$

To identify the parameters of this model we write the following moments for group g.

$$(14) \qquad Var\left(\Delta\tilde{x}_{gt}\right) = \sigma_{uxg}^2 + (\alpha_{1g}^{xx})^2\sigma_{uxg}^2 + \frac{1}{T-1}\sum_{t=2}^{T}\left(Var_{gt}(\eta_{gt}^x) + Var_{gt-1}(\eta_{gt-1}^x)\right)$$

$$(15) \qquad Cov\left(\Delta\tilde{x}_{gt}, \Delta\tilde{x}_{gt-1}\right) = \alpha_{1g}^{xx}\sigma_{uxg}^2 - \frac{1}{T-1}\sum_{t=2}^{T}Var_{gt-1}(\eta_{gt-1}^x)$$

$$(16) \qquad Var_{gt}\left(\tilde{x}_{gt}\right) \equiv \frac{1}{N_{gt}}\left(\frac{1}{N_{gt}}\sum_{h=1}^{N_{gt}}(\tilde{x}_{hgt} - \tilde{x}_{gt})^2\right) = Var_{gt}(\eta_{gt}^x)$$

where $Var_{gt}\left(\tilde{x}_{gt}\right)$ is the cross sectional variance of average consumption within a group-year cell, N_{gt} is the sample size for a group-year cell, while $Var\left(\Delta\tilde{x}_{gt}\right)$ is the time series variance of (group g) average consumption growth. Notice that sampling errors η_{gt}^x induce heteroskedasticity in

---

roots in consumption and income, therefore, has to come from the shocks in which we are interested and justifies our modeling the first differences of these variables.

[15] We also estimated our models using a first-step deterministically de-trended interest rate: this procedure does not affect our results.

average consumption growth; the time series variances considered in equation (14) is obtained by taking the average over time of the variances.

It is possible to estimate $Var_{gt}\left(\eta_{gt}^{x}\right)$ from the equation (16) and use these estimates to compute the last terms of equations (14) and (15) to obtain an estimate of α_{1g}^{xx} and σ_{uxg}^{2}. However, there is no guarantee that, in small samples, the resulting variance covariance matrix for $\left\{u_{gt}^{x}, u_{gt-1}^{x}\right\}$ is positive-definite and that $\sigma_{uxg}^{2} > 0$.

When considering a multivariate system with several lags, these small sample problems are accentuated as we consider the sampling errors of different quantities (namely income and consumption).

These problems are avoided when implementing a Maximum Likelihood (ML) approach where the likelihood is computed recursively using the Kalman Filter, which we describe in detail in the appendix. The latter guarantees that the relevant variance-covariance matrices stay positive definite. For this reason, we adopted such a method.

The ML is computed under the assumption that the residuals in system (4) are Gaussian. The assumption of normality of the residuals can certainly be criticized. Indeed, Abowd and Card (1989) report some evidence against normality within their framework. We should stress, however, that we are modelling time series, rather than cross sectional variability.[16] More importantly, our approach can be defended on the ground that, as we are modelling only mean and variances of the variables of interest, we are effectively using a quasi-maximum likelihood approach as advocated, for instance, by Watson (1989). To test the robustness of such an argument we implemented the following exercise, which we explain with the help of the univariate model in equation (13) above, but that can be generalized to the multivariate case.

Consider again equation (13) and notice that the sum of two MA(1) processes is still an MA(1). We can then re-write the process for consumption as:

$$(17) \qquad \Delta x_{gt} = v_{gt}^{x} + \delta_{1g}^{xx} v_{gt-1}^{x}$$

where δ_{1g}^{xx} and the variance of v_{gt}^{x} will depend on α_{1}^{xx} and the variances of u_{gt}^{x} and η_{gt}^{x}. If one is not interested in identifying separately these latter set of moments, one can identify δ_{1g}^{xx} and the variance of v_{gt}^{x} from equation (17). We proceeded to do so both by ML (using again a Kalman filter) and a method of moments. The results we obtain for simple univariate models and for multivariate ones, which are available upon demand, are extremely similar with the two methods. This similarity gives us confidence in appealing to a quasi-maximum likelihood argument to justify our use of the Kalman Filter in the estimation of system (4).

When estimating model (4) by Maximum Likelihood, we follow a two-step procedure. We first estimate the variances of the terms η_{gt}^{z}. Remembering that these errors are induced by the finite size of the cells used to compute the cohort averages, we notice that we can estimate variances and covariances of the η_{gt}^{z} terms consistently from the within cell variability and the

---

[16] We are aware, however, that several papers (such as Fagiolo et al., 2008, Fernandez-Villaverde and Rubio-Ramirez, 2006, and Christiano, 2007) that have studied aggregate time series data have found evidence of fat tails that would contradict the assumption of normality. As we discuss below, however, the results we obtain in situations where we can implement GMM methods are very similar to the quasi-ML approach we use.

cell size. We will assume that the idiosyncratic shocks are independent of $\Delta \tilde{z}_{gt}$, the changes in the group-level innovations we are interested in modelling.

Hence we compute the time t variance of the cohort/education group averages η_{gt}^z using the cross-sectional variance for each point in time, corrected for the cell size:

$$(18) \qquad Var_{gt}\left(\eta_{gt}^z\right) = \frac{1}{N_{gt}} \left(\frac{1}{N_{gt}} \sum_{h=1}^{N_{gt}} (\tilde{z}_{hgt} - \tilde{z}_{gt})^2 \right) \qquad z = x,y; \ t=1,\ldots T$$

Analogously, the covariance terms are:

$$(19) \qquad Cov_{gt}\left(\eta_{gt}^x, \eta_{gt}^y\right) = \frac{1}{N_{gt}} \left(\frac{1}{N_{gt}} \sum_{h=1}^{N_{gt}} (\tilde{x}_{hgt} - \tilde{x}_{gt})(\tilde{y}_{hgt} - \tilde{y}_{gt}) \right) \qquad t=1,\ldots T$$

We filter each series of variances or covariances with an MA smoother with a 12-period span. In the computations of the ML, we correct the overall variance covariance matrix with these figures and, as we have about 12,000 observations for each cohort, treat the estimated variances and covariances as known.

6.   Results

In this section, we report our main estimation results. A structural interpretation of the coefficients we estimate is given in the next section. Before we present estimates of the statistical models we discussed in Section 2, we present briefly some of the time series properties of the data we use in estimation.

6.1.   *Univariate models*

We start by reporting, in tables 2 to 4, estimates of univariate MA models for consumption, income, and the interest rate. In particular, table 2 shows estimates of univariate parsimonious MA representations for the logarithm of non-durable expenditure for each cohort, both without and with correction for sampling errors η_{gt}^x. The results on the first four columns of the table correspond to a version of equation (13), while the last four to a version of equation (17), both augmented to consider additional lags. The coefficient on the lag zero shocks (α_{0g}^{xx} or δ_{0g}^{xx}) is normalized to one, and it is not shown in the tables.

<Table 2 around here>

The main interest of this exercise is to check the validity of some of our assumptions; in particular we want to test whether, conditional on time-varying unobserved heterogeneity, the variance of the MA disturbance is constant over time.[17] To this purpose, we split the sample in two periods, from 1978 to 1993 (first quarter) and from 1993(2) to 2005(1), and estimate two separate models for each cohort.[18] We start with a six-lag specification and constrain to zero one lag at a time. With this procedure, we end up selecting an MA(1) process for both cohorts (upper

---

[17] For brevity, in this section we present our results splitting the sample into two cohorts, without further grouping according to the education level of the head.

[18] The second and younger cohort is estimated over the period 1980(1)-1993(1) and 1993(2)-2005(1). Diagnostic tests for white noise on the residuals as well as on the squared residuals of these and subsequent MAs are reported in the appendix, and do not indicate any sign of autocorrelation.

and lower panel, respectively), both when accounting for sampling errors to identify separately the u_{gt}^x and η_{gt}^x processes (left panel, columns i and ii) and when estimating the equivalent of equation (17) (right panel, columns v and vi). For each model, we then test whether *a)* the MA coefficient is constant over time (columns iii and vii) and *b)* in addition, the (square root of the) variance of the MA disturbance is also constant over time (columns iv and viii). The P-values for these likelihood ratio tests are reported in each corresponding column: for example, for cohort 1 and when correcting for measurement error the p-value of restricting the MA coefficient to be constant across periods is 89 per cent, while the p-value for restricting both the coefficients and the (square roots of the) variances to be constant is 98 per cent. The restrictions are also not rejected for the second cohort, with p-values of 90 and 97 per cent respectively.

In the right panel of table 2 we repeat the tests without explicitly taking into account the variation induced by sampling errors: we find that for both cohorts the MA coefficient is constant over the two sample periods (with p-values of 81 and 38 per cent respectively), and that the variances are constant over time, with a p-value of 96 and 15 per cent respectively.

Hence, we find that the assumption of a stable cohort-specific MA process for non-durable expenditure is not rejected by our data.

Table 3 reports results analogous to those in Table 2, but for (the first difference of the logarithm of) disposable income. Also in this case the data always indicate that disposable income is well represented by MA(1) process with a negative coefficient. The tests of constancy of the MA parameter and variances do not reject the null both when we disentangle the variation induced by sampling errors (left panel), and when we do not.

<Table 3 around here>

A negative autocorrelation of the first difference of income has been found in the US literature both with aggregate data (for example by Watson, 1986, on aggregate disposable income) and with micro data on male earnings (MaCurdy, 1982, Abowd and Card, 1989 and, more recently, Meghir and Pistaferri, 2004).[19] We would expect the coefficients obtained when estimating the equivalent of equation (17) to be large in absolute value as it reflects the sum of two MA(1) one of which has a coefficient of -1. However, even when we separately identify the sampling error process, our point estimates for α_{1g}^{yy} are somewhat larger in absolute value than what reported in these studies: for the first cohort, we report an estimate of -0.31 (s.e.=0.2), while for the second cohort an estimate of -0.39 (s.e.=0.13).

Table 4 reports estimates for univariate MA models for the real interest rate. Preliminary tests based on the squared residuals of a linear MA with constant variance (reported in column v in the Table) indicate the residuals being not independent.[20] To overcome this problem, we introduce a number of structural breaks, choosing the time periods based on previous evidence. The real interest rate in the short run is influenced both by the nominal interest rate and by the (expected) inflation rate. To take into account the "great moderation" that took place in the mid-eighties we insert a structural break in 1985, first quarter (see Kim and Nelson, 1999, and McConnell and Perez-Quiros, 2000, for the Great Moderation in the US; Benati, 2008, for a study on the UK). In addition, at the end 1992 the Bank of England started its inflation-targeting policy, hence we include a second structural break at the end of 1992.

---

[19] Other studies, such as Blinder and Deaton (1985) using US aggregate data, find a positive autocorrelation in the first-difference series of disposable income.

[20] Q- and normality tests on the residuals and squared residuals are shown in the Appendix for all the models estimated.

<Table 4 around here>

When allowing for two breaks in 1984(4) and 1993(1), we find that the interest rate dynamics is best represented by a three-lag specification. In Table 4, we report unrestricted estimates for each sub-period: 1978(1)-1985(1) in column (i), 1985(2)-1993(1) in column (ii), and 1993(2)-2005(1) in column (iii). The p-value of the restrictions imposed by specification (iv), which imposes constancy of the MA coefficients in the three sample periods, is 90%. However, imposing the additional restriction of a constant variance, as shown in column (v), leads instead to a rejection of the restrictions at any significance level.[21] Hence, in the subsequent analysis we will model the interest rate as an MA process with constant MA coefficients and two breaks in the variance.

6.2. *Multivariate models*

Having estimated the univariate models, we now move on to the estimation of the main model discussed in Section 2, equation (4). In particular, for each cohort, we consider two variables: non-durable expenditure and disposable income. In addition, we have an equation for the real interest rate, which is common to the two cohorts. We thus perform estimates of the three-equation system for each cohort separately; these systems all identify separately the shocks to consumption and income from the variability induced by sampling variations by using the information on cell sizes and within cell variability as discussed above. In both cases we start from a six-lag specification, constraining to zero one coefficient at a time to find a parsimonious specification: this specification is shown in table 5 for cohort 1 and 2. The results on less parsimonious models are available upon request. In Tables 6 and 7, we re-do the exercise separating households not only by the year of birth of the household head but also but his/her education achievement.

Given the interpretation of the results in terms of excess sensitivity we give in the next section, we paid particular attention to the coefficients on lag 'income shocks' in the specification for consumption. Hence, in table 8, we report for each specification the p-value of the Likelihood ratio test of the restriction that the one-lag income shock in the consumption equation is equal to zero. We discuss these results in Section 7. In addition, we do not attempt to constrain the coefficients of the interest rate shock in the consumption equation: those coefficients will be used to identify the elasticity of intertemporal substitution and will be further discussed in Section 7.

<Table 5 around here>

For both cohorts we do not reject the hypothesis that lagged 'income' shocks are equal to zero in the consumption equation: the p-value of the test that the coefficient is zero is 0.28 and 0.83 for the first and second cohort respectively (see Table 8 in the next section). Whilst in the tables we only report the estimates for the first lag, we also tried adding additional lags, in which case we obtained the same result. Similarly, we also tried to add to the consumption equation lagged values of the 'consumption' shock and failed to reject the hypothesis that these were equal to zero. For this reason, in the table we report the results of a specification that does not contain lagged income and consumption shocks.

As for the interest rate shocks, we selected a six-lag specification for cohort 1, while for the second cohort a three-lag specification is preferred. Allowing for six-lag specification in the interest rate in the equations for cohort 2 does not affect our results. In the consumption

---

[21] The same tests, with virtually identical results, have been carried on a six-lag specification.

equations for the two cohorts, the interest rate shocks are only estimated significantly at lag one for the first cohort, and at lag zero and one for the younger one. This evidence is in line with the univariate models for consumption presented in the second and fourth columns of Table 3, which did not show any prolonged dynamics.

For both cohorts, the contemporaneous shock to income has a significant effect on consumption. We discuss the magnitude of these coefficients in Section 7.3.

As for the income equations, consistently with the evidence on the univariate models, we do not find a very long dynamics; for the first cohort we estimate an MA(0) in the first differences, as the MA(1) coefficient is not estimated to be different from zero. In addition, we fail to reject the hypothesis that, in the income equation, all the coefficients to the interest rate shocks are zero, hence in our parsimonious specification disposable income does not depend on the interest rate. Once again, this evidence is not inconsistent with the evidence on earnings or disposable income presented by other authors for the US.

<Table 6 around here>

<Table 7 around here>

When we estimate the same model separately for households headed by individuals with more than compulsory education and individuals who left school at the age of compulsory education, we obtain the results reported in Tables 6 and 7 as 'lower' and 'higher' education. The main differences with the estimates in Table 5 are:

i) the prolonged dynamics for the interest rate in cohort 1 consumption is found to be significant only for the higher education group, and, more interestingly,

ii) the one-lag income shock enters significantly the consumption equation for both cohorts in the lower education group, while it is found to be non significantly different from zero for the higher educated.

7.   A structural interpretation of the results

Given the estimates of our statistical model, we can now use them to test different versions of the life cycle model and to estimate structural parameters. This is what is typically done in the literature. The difference of our approach is that we make explicit what aspects of the variability we observe in the data identify different components of the model.

7.1.   *Testing the model: Excess sensitivity*

As we discussed above, a simple version of the life cycle model implies that predictable movements to disposable income should not help predicting consumption changes (see Hall, 1978). In our framework, we can easily test whether changes in log consumption exhibit "excess sensitivity" to past (i.e. predictable) income innovations by testing whether lagged income group-level shocks are good predictors of consumption.

To summarize the evidence already reported in tables 5-7, we report in table 8 the p-values of the zero restriction on the lagged income shock in the consumption equation. For both cohorts, without further grouping, the corresponding p-values are 28 and 83 per cent, indicating that neither cohort, as a whole, displays excess sensitivity to income. When we distinguish between higher and lower educated heads of households, we find that for the latter group the p-

values are lower than 5 per cent (0.6 per cent in both cohorts). On the other hand, the p-values are higher than 60 per cent for both cohorts in the higher education group.

<Table 8 around here>

While these findings can be interpreted in different ways,[22] which cannot be effectively distinguished in this framework, it is noteworthy that we find excess sensitivity only for the less educated individuals, who hold less wealth to be used as a collateral and might be more subject to liquidity constraints than the higher educated. This type of finding is also consistent with the evidence reported by Blundell, Pistaferri and Preston (2008) who show that a higher fraction of transitory shocks is reflected in the consumption of low education households than in that of high education households in the US. Alternatively, as emphasised by Carroll (2001) and Carroll and Kimball (2001), impatient consumers who hold little financial wealth (with respect to uncertain human wealth) will engage in precautionary saving, hence behaving like liquidity constrained individuals. Attanasio et al. (1999) indicate that low education households might be more impatient. The elasticity of intertemporal substitution may also play a role: individuals who are less willing to reallocate expenditure between periods to take advantage of interest rates movements end up accumulating less financial wealth, and hence they could be more exposed to liquidity constraints.

7.2. *The elasticity of intertemporal substitution*

In Table 9, we report estimates of the elasticity of intertemporal substitution implied by the restrictions in equation (9) and by the estimates of the coefficients of model (4) we have obtained. To compute those estimates we estimate a restricted version of our system, in which we impose the constraints in (9). We also use the log-likelihood of the restricted model to construct likelihood ratio tests of the restrictions that we impose.

<Table 9 around here>

For the first and elder cohort, we find that the elasticity is estimated at 1.01 with a standard error of 0.26, so that it is statistically different from zero and not from one. For the second younger cohort, we find a lower elasticity equal to 0.63, and, although less precisely estimated with a standard error of 0.39, is also statistically different from zero (with a p-value of 3.5 per cent) and not from one. The restrictions imposed by equation (9) are tested by comparing the value of the unrestricted and restricted likelihoods: the corresponding p-values are reported in the last two rows of table 9 and, for the two cohorts as a whole, indicate that the restrictions imposed by (9) are not rejected. These estimates of the elasticity parameter are consistent with those reported by Attanasio and Weber (1993) and other studies that used the FES.

When we disaggregate each cohort into more homogeneous educational attainments, we find that both the higher educated groups display an elasticity of intertemporal substitution higher than one (1.29 and 1.19 for cohort 1 and 2 respectively): in both cases the standard errors indicate that the parameter is statistically different from zero and not different from one. The p-values of the likelihood ratio tests of the restrictions imposed by equation (9) indicate the restrictions are not rejected.

---

[22] One possibility that is typically made in the literature is the presence of liquidity constraints. Alternatively, one could appeal to non separability between consumption and leisure (Heckman, 1974). To test the latter model we would have to extend our statistical model to include leisure choices and decompose earnings between wages and hours.

Results for the compulsory education groups are rather different. The estimated elasticity is 0.6 for the elder cohort with a standard error equal to 0.31. For the younger cohort, the estimated parameter is equal to 0.42 and it not statistically different from zero. Also in this case, the p-values of the likelihood ratio tests of the restrictions imposed by equation (9) indicate the restrictions are not rejected.

A number of studies emphasise the presence of heterogeneity in the preference parameters: Blundell, Browning and Meghir (1994), Attanasio and Browning (1995), and more recently Crossley and Low (2011) and Alan and Browning (2010) all find a non-constant elasticity of intertemporal substitution. All these studies, with the exception of Alan and Browning (2010), find the elasticity of intertemporal substitution is increasing in consumption/wealth. Mankiw and Zeldes (1991) and Attanasio, Banks and Tanner (2002) find different elasticities among stock-holders and non stock-holders. Guvenen (2006) shows how introducing heterogeneity in elasticities of intertemporal substitution into a calibrated model improves the fit of the model, and provides an example of how policy conclusions based on a calibrated model with constant elasticity can be misleading.

When considering different groups (year of birth cohorts) or groups defined on educational attainment, we have so far assumed that individuals belonging to different groups face the same interest rates. This assumption, which is equivalent to say that the assets whose return we are using in the empirical analysis is available and used by members of all groups, could be violated for several reasons. First, even if an asset with the same nominal interest rate is available to all groups, if marginal tax rates are different across groups, the relevant intertemporal prices would be systematically different across groups. Second, it is possible that the asset we are considering is not available or not widely used by the members of a certain group. If these problems are relevant, the different estimates of the elasticity of intertemporal substitution or the different results on the model rejections would not necessarily be valid. Investigating these issues would be extremely interesting. Unfortunately, the lack of data on portfolio holdings makes this analysis particularly difficult (see, however, the analysis on US data by Mankiw and Zeldes, 1991 and Attanasio, Banks and Tanner, 2002).

7.3.   Markets: Excess smoothness

Using the results we obtained estimating various versions of system (4), and following Attanasio and Pavoni (2011), we can test the restriction in equation (11) against the alternative of excess smoothness. In particular we test the hypothesis that $\alpha^{xy}(R) - \alpha^{yy}(R) = 0$ (no excess smoothness) against the alternative that the same difference is less than 0. We evaluate the terms $\alpha^{zy}(R)$ at a quarterly interest rate equal to 1 per cent. We report these results in Table 10.

<Table 10 around here>

Of particular interest are the results derived from the estimates in table 5 (cohort 1 and 2) and for the highly educated in table 6 (cohort 1) and 7 (cohort 2) where the restrictions implied by the Euler equations are imposed (it should be stressed that, as we mentioned above, these restrictions are not rejected). For the first (and elder) cohort the excess smoothness test equals -0.63, with a standard error of 0.20 and a P-value of less than 1 per cent. For the second (and younger) cohort the excess smoothness test is equal to -0.25, with an estimated standard error equal to 0.14 (the P-value is 4 per cent). Hence we reject the hypothesis that $\alpha^{xy}(R) - \alpha^{yy}(R) = 0$, or in other words we reject the null of no excess smoothness. When disaggregating the two cohorts, the excess sensitivity test is equal to -0.23 for both the high educated groups, with a standard error lower than 0.15 in both cases; the P-values are less than 3

and 1 per cent, respectively. Therefore, for this education group we also reject the hypothesis of no excess smoothness.

We perform the same test for the low education groups, although the restrictions implied by the Euler equation are not imposed, and we find that we do not reject the hypothesis of no excess smoothness only for the low education group of cohort 1 (for which the P-value is 26 per cent), while we find excess smoothness for cohort 2 (with P-value of about 3%).

Our results are in line with those of Attanasio and Pavoni (2011) who find, when using non-durable consumption and a broad definition of disposable income which includes smoothing mechanisms, comparable results.

8. Conclusions

In this paper, we have analysed the time series properties of group-level consumption expenditure and income. The methodology we propose consists in estimating multivariate moving average systems for synthetic panels constructed from time series of repeated cross sections. This approach has the advantage of allowing to take into account explicitly the sampling variation induced by the fact that we rely on time series of repeated cross sections to identify the dynamics of individual measures of consumption and income. Data are drawn from the UK Family Expenditure Survey.

When we define groups based on the year of birth, we cannot reject the hypothesis that lagged consumption, income and interest rate shocks have no effects on consumption changes. This evidence is coherent with the results of 'no-excess sensitivity' of consumption reported on micro data by Attanasio and Weber (1993, 1995), Blundell, Browning and Meghir (1994) and Attanasio and Browning (1995). However, when we define cohort-education groups, we find that those households whose head has at most completed compulsory schooling do display some 'excess sensitivity' of consumption to lagged income. Households whose head has a higher education, on the other hand, do not display any excess sensitivity.

We also use the estimated parameters to identify the elasticity of intertemporal substitution using micro data: when we aggregate our data to form year-of-birth groups, we find estimates consistent with the estimates reported by Attanasio and Weber (1993). When disaggregating according to the educational level, we find important differences in the estimated elasticities: while the better educated cohorts display an elasticity of intertemporal substitution greater than – but not statistically different from – one, estimates for the lower educated are in the range of 0.5 and not statistically different from zero. In addition, the restrictions imposed by our estimation procedure to estimate the elasticity of intertemporal substitution are rejected for the lower educated, while the higher educated easily pass the test.

Our approach can be also used to assess the extent to which innovations to income (or interest rates) are reflected into consumption. The Euler equation approach is typically completely silent about this. Given our estimates of the parameters of the time series model for consumption, income and the interest rate, we do find some evidence of 'excess smoothness' of consumption. That is, we find that consumption reacts to permanent innovations to income in a way that is not consistent with the life cycle model. Our evidence is in line with that in Attanasio and Pavoni (2011) who report evidence of excess smoothness and interpret it, following Hansen, Roberds and Sargent (1991), as a violation of the intertemporal budget constraint *with a single asset*. Attanasio and Pavoni (2011) also find that using income definitions that include smoothing mechanisms, such as social assistance and net taxes, results in milder evidence of "excess smoothness".

Another paper which is related to ours and in particular to our attempt to identify the fraction of innovations in income which is reflected in current and future consumption is that by Blundell, Pistaferri and Preston (2008). These authors use a different and complementary approach, as they exploit information on the evolution of cross-sectional second moments, whilst we study the time series properties of group means and map them into estimates of structural parameters. As we mentioned in Section 7, some of our results are consistent with theirs.

The analysis in this paper was done on grouped data, mainly because of data availability. The English FES provides a long time series of cross sections that can be used to study the time series properties of grouped level data. The availability of individual level data is useful for a variety of reasons, including the fact that one controls the process of aggregation. However, the lack of a longitudinal dimension to the data forces us to use grouped data to study their dynamics. This means that we potentially loose dynamic effects that are purely idiosyncratic as they are washed out by the averaging at the cohort level. The availability of individual level longitudinal data could be very useful to investigate how much of the dynamics of individual consumption we were unable to observe because of the nature of the data available to us.

All of our analysis has considered non-durable consumption as a homogeneous commodity. Our approach, however, could be used to consider richer and more complex models. One could, for instance, consider situations where multiple commodities affect utility in ways that are not perfectly aggregable and combine the analysis of demand systems with the dynamics we analyse here. Or one could combine the analysis of hours and consumption, therefore relaxing the separability assumptions implicitly maintained in our exercise. Finally, one could also think of characterizing the dynamics of participation rates for different groups in an attempt to quantify the importance of extensive and intensive margins in labour supply over the business cycle. In this last application, however, the mapping between the time series properties of the variables analysed and tightly parameterised models of preferences would be much harder.

References

Abowd, J. and D. Card, "On the Covariance Structure of Earnings and Hours Changes," Econometrica 57 (1989), 411-445.

Alan, S. and M. Browning, "Estimating Intertemporal Allocation Parameters using Synthetic Residual Estimation," The Review of Economic Studies 77 (2010), 1231-1261.

Alan, S., O.P. Attanasio and M. Browning, "Estimating Euler Equations with Noisy Data: Two Exact GMM Estimators," Journal of Applied Econometrics 24 (2009), 309-324.

Attanasio, O.P., J. Banks, C. Meghir and G. Weber, "Humps and Bumps in Lifetime Consumption," Journal of Business Economics and Statistics 17 (1999), 22-35.

Attanasio, O.P, J. Banks and S. Tanner, "Asset Holding and Consumption Volatility," Journal of Political Economy 110 (2002), 771-792.

Attanasio, O.P. and M. Browning, "Consumption over the Life-Cycle and over the Business Cycle," American Economic Review 85 (1995), 1118-1137.

Attanasio, O.P. and H. Low, "Estimating Euler Equations," Review of Economic Dynamics 7 (2004), 405-435.

Attanasio, O.P. and N. Pavoni, "Risk Sharing in Private Information Models with Asset Accumulation: Explaining the Excess Smoothness of Consumption," Econometrica 79 (2011), 1027-1068.

Attanasio, O.P. and G. Weber, "Consumption Growth, the Interest Rate and Aggregation," Review of Economic Studies 60 (1993), 631-49.

Attanasio, O.P. and G. Weber, "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," Journal of Political Economy 103 (1995), 1121-1157.

Benati, L., "The 'Great Moderation' in the United Kingdom," Journal of Money, Credit and Banking 40 (2008), 121-147.

Blinder, A.S. and Deaton, A.S., "The Time-Series Consumption Function Revisited," Brookings Papers on Economic Activity 16 (1985), 465-521.

Blundell, R., L. Pistaferri and I. Preston, "Consumption Inequality and Partial Insurance," American Economic Review 98 (2008), 1887-1921.

Blundell, R., M. Browning and C. Meghir, "Consumer Demand and the Lifetime Allocation of Consumption," Review of Economic Studies 61 (1994), 57-80.

Brewer, M., A. Goodman, J. Shaw and L. Sibieta, "Poverty and Inequality in Britain: 2006," IFS Commentary 101, Institute for Fiscal Studies, 2006.

Caballero R., "Expenditure on Durable Goods: A Case for Slow Adjustment," Quarterly Journal of Economics 422 (1990), 727-744.Campbell, J.Y., "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis," Econometrica 55 (1987), 1249-73.

Campbell, J.Y. (1994) "Inspecting the mechanism: An analytical approach to the stochastic growth model," Journal of Monetary Economics 33 (1994), 463-506.

Campbell, J.Y. and A.S. Deaton (1989): "Why Is Consumption So Smooth?," Review of Economic Studies 56 (1989), 357–373.

Carroll, C.D., "Death to the Log-Linearized Consumption Euler Equation! (And Very Poor Health to the Second-Order Approximation)," The B.E. Journal of Macroeconomics 1 (2001), 1-38.

Carroll, C.D. and M.S. Kimball, "Liquidity Constraints and Precautionary Saving," NBER Working Papers 8496, National Bureau of Economic Research, 2001.

Christiano, L. J., "Comment," Journal of Business & Economic Statistics 25 (2007), 143-151.

Crossley T.F. and H. W. Low, "Is the Elasticity of Intertemporal Substitution Constant?," Journal of the European Economic Association 9 (2011), 87-105.

Cunha, F., Heckman, J., and Navarro, S., "Separating heterogeneity from uncertainty an Aiyagari-Laitner economy," Paper presented at the Goldwater Conference on Labor Markets in Arizona, March, 2004.

Cunha, F., J. Heckman and Navarro, S., "Separating Uncertainty from Heterogeneity in Life Cycle Earnings," Oxford Economic Papers 57 (2005), 191-261.

Deaton, A., "Panel Data from Time Series of Cross-Sections," Journal of Econometrics 30 (1985), 109-26.

Elliott, G., T. J. Rothenberg, and J. H. Stock, "Efficient tests for an autoregressive unit root," Econometrica 64 (1996), 813-836.

Fagiolo, G., Napoletano, M. and A. Roventini, "Are output growth-rate distributions fat-tailed? Some evidence from OECD countries," Journal of Applied Econometrics 23 (2008), 639-669.

Fernandez-Villaverde, J. and J. F. Rubio-Ramirez, "Estimating Macroeconomic Models: A Likelihood Approach," NBER Technical Working Papers, National Bureau of Economic Research, 2006.

Flavin, M.A., "The Adjustment of Consumption to Changing Expectations about Future Income," Journal of Political Economy 89 (1981), 974-1009.

Forni, M., and L. Reichlin, " Dynamic Common Factors in Large Cross-Sections," Empirical Economics 21 (1996), 27-42.

Guvenen, F., "Reconciling conflicting evidence on the elasticity of intertemporal substitution: A macroeconomic perspective," Journal of Monetary Economics 53 (2006), 1451-1472.

Hall, R.E., "Stochastic Implications of the Life-Cycle Permament Income Hypothesis: Theory and Evidence", Journal of Political Economy 86 (1978), 971-87.

Hansen, L.P., W. Roberds and T.J. Sargent, "Time Series Implications of Present Value Budget Balance and of Martingale Models of Consumption and Taxes," in L. P. Hansen and T. J. Sargent, eds., Rational Expectations Econometrics (Boulder: Westview, 1991), 121-61.

Heckman, J.J., "Shadow Prices, Market Wages, and Labor Supply," Econometrica 42 (1974), 679-694.

Kim, C., and C. R. Nelson. "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," The Review of Economics and Statistics 81 (1999), 608 – 616.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidl, and Y. Shin, "Testing the null hypothesis of stationarity against the alternative era unit root: How sure are we that economic time series have a unit root?," Journal of Econometrics 54 (1992), 159-178.

Lillard, L.A.and Willis, R.J., "Dynamic Aspects of Earning Mobility," Econometrica 46 (1978), 985-1012.

MaCurdy, T.E., "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," Journal of Econometrics 18 (1982), 83-114.

Mankiw N.G. and S.P. Zeldes, "The consumption of stockholders and nonstockholders," Journal of Financial Economics 29(1991), 97-112.

McConnell, M. M., and G. Perez-Quiros, "Output Fluctuations in the United States: What has Changed Since the Early 1980's," American Economic Review 90 (2000), 1464-1476.

Meghir, C. and L. Pistaferri, "Income variance dynamics and heterogeneity," Econometrica 72 (2004), 1-32.

Moffitt R.A. and P. Gottschalk, "Trends in the Variances of Permanent and Transitory Earnings in the U.S. and Their Relation to Earnings Mobility," Boston College Working Papers in Economics 444 (1995).

Quah, D., "Permanent and Transitory Movements in Labor Income: An Explanation for "Excess Smoothness" in Consumption," Journal of Political Economy 98 (1990), 449-475.

Quah, D. and T.J. Sargent, "A Dynamic Index Model for Large Cross Sections," in J. Stock and M. Watson, eds., Business Cycles, indicators and forecasting (Chicago: University of Chicago Press for NBER, 1994), 285-310.

Sargent, T.J., "Rational Expectations, Econometric Exogeneity and Consumption," Journal of Political Economy 86 (1978), 673-700.

Tanner, S. (1998), "How Much Do Consumers Spend? Comparing the FES and National Accounts," in J. Banks and P. Johnson, eds., How Reliable is the Family Expenditure Survey? Trends in Incomes and Expenditures Over Time, IFS Report 57 (London: IFS, 1998), 67-121.

Watson, M.W., "Univariate Detrending Methods with Stochastic Trends," Journal of Monetary Economics 18 (1986), 49-75.

Watson, M.W., "Recursive Solution Methods from Dynamic Linear Rational Expectations Models," Journal of Econometrics 41 (1989), 65-89.

Watson M.W. and R.F. Engle, "Alternative Algorithms for the Estimation of Dynamic Factor, MIMIC, and Time Varying Coefficient Regression Models," Journal of Econometrics 23 (1983), 385-400.

West, K.D., "The Insensitivity of Consumption to News About Income," Journal of Monetary Economics 21 (1988), 17-34.

Zeldes, S.P., "Consumption and Liquidity Constraints: An Empirical Investigation," Journal of Political Economy 97 (1989), 305-46.

Tables

Table 1 – Cohort definition

Cohort	Year of Birth	Observation period	Median age in first year	Median age in last year	Mean Cell Size	Total number of observations
1	1947-53	1978(1)-2005(1)	28	55	123	13,512
2	1954-60	1980(1)-2005(1)	23	52	110	11,095

Note: data drawn from the FES, 1978(1) – 2005(1)

Table 2 – Parsimonious estimates: change in log non-durable consumption.

	With correction for measurement error				Without correction for measurement error			
	1978(2)-1993(1)	1993(2)-2005(1)	1978(2)-2005(1) break	1978(2)-2005(1) no break	1978(2)-1993(1)	1993(2)-2005(1)	1978(2)-2005(1) break	1978(2)-2005(1) no break
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Cohort 1								
u_{t-1}^x	-0.475	-0.311	-0.423	-0.413	-0.707	-0.763	-0.733	-0.733
(se)	(0.214)	(0.553)	(0.226)	(0.264)	(0.099)	(0.079)	(0.061)	(0.063)
$\sigma(u_{93-05}^x)$		1.497	1.714			4.211	4.223	
(s.e.)		(1.052)	(0.694)			(0.428)	(0.432)	
$\sigma(u_{78-93}^x)$	2.251		2.166		3.788		3.792	
(s.e.)	(0.571)		(0.636)		(0.342)		(0.293)	
$\sigma(u_{78-05}^x)$				1.921				3.982
(s.e.)				(0.666)				(0.258)
Log L	-165.66	-137.53	-303.31	-303.40	-165.03	-137.10	-302.33	-302.44
P-value[a]			0.887	0.981			0.819	0.961
Cohort 2								
u_{t-1}^x	-0.352	-0.741	-0.399	-0.356	-0.562	-0.812	-0.631	-0.673
(se)	(0.401)	(0.184)	(0.193)	(0.208)	(0.099)	(0.112)	(0.076)	(0.069)
$\sigma(u_{93-05}^x)$		3.801	3.312			5.333	5.401	
(s.e.)		(0.770)	(0.847)			(0.543)	(0.531)	
$\sigma(u_{80-93}^x)$	2.346		2.478		3.705		3.730	
(s.e.)	(1.092)		(0.611)		(0.356)		(0.359)	
$\sigma(u_{80-05}^x)$				2.495				4.563
(s.e.)				(0.568)				(0.322)
Log L	-143.81	-148.66	-292.58	-292.70	-141.87	-148.44	-291.28	-293.66
P-value[a]			0.900	0.978			0.381	0.154

Note: parsimonious univariate MA estimates of the change in log non-durable consumption, using cohort data, without/with correction for measurement error. Standard errors in parentheses.

[a] P-value for the Likelihood ratio test when testing the zero restrictions versus the unrestricted models (given by column (i) plus (ii) and (v) plus (vi) respectively)

Table 3 – Parsimonious estimates: change in log disposable income.

	With correction for measurement error				Without correction for measurement error			
	1978(2)-1993(1)	1993(2)-2005(1)	1978(2)-2005(1) break	1978(2)-2005(1) no break	1978(2)-1993(1)	1993(2)-2005(1)	1978(2)-2005(1) break	1978(2)-2005(1) no break
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Cohort 1								
u_{t-1}^{y}	-0.228	-0.629	-0.289	-0.306	-0.572	-0.804	-0.616	-0.638
(se)	*(0.059)*	*(0.282)*	*(0.167)*	*(0.214)*	*(0.084)*	*(0.105)*	*(0.072)*	*(0.062)*
$\sigma(u_{93-05}^{y})$		2.286	1.621			4.235	4.307	
(s.e.)		*(0.924)*	*(0.867)*			*(0.417)*	*(0.407)*	
$\sigma(u_{78-93}^{y})$	1.940		2.047		3.339		3.348	
(s.e.)	*(0.496)*		*(0.570)*		*(0.289)*		*(0.296)*	
$\sigma(u_{78-05}^{y})$				1.991				3.799
(s.e.)				*(0.472)*				*(0.258)*
Log L	-156.65	-137.06	-294.10	-294.24	-157.46	-137.38	-295.82	-297.38
P-value[a]			0.671	0.898			0.375	0.278
Cohort 2								
u_{t-1}^{y}	-0.268	-0.629	-0.365	-0.394	-0.506	-0.779	-0.616	-0.611
(se)	*(0.200)*	*(0.170)*	*(0.155)*	*(0.132)*	*(0.099)*	*(0.065)*	*(0.072)*	*(0.057)*
$\sigma(u_{93-05}^{y})$		2.513	1.910			3.955	4.307	
(s.e.)		*(0.658)*	*(0.684)*			*(0.403)*	*(0.407)*	
$\sigma(u_{80-93}^{y})$	2.754		2.945		3.955		3.348	
(s.e.)	*(0.590)*		*(0.537)*		*(0.379)*		*(0.296)*	
$\sigma(u_{80-05}^{y})$				2.704				4.041
(s.e.)				*(0.464)*				*(0.267)*
Log L	-145.92	-134.70	-281.51	-282.45	-145.27	-134.09	-281.55	-281.51
P-value[a]			0.412	0.455			0.112	0.368

Note: parsimonious univariate MA estimates of the change in log disposable income, using cohort data, without/with correction for measurement error. Standard errors in parentheses.
[a] P-value for the Likelihood ratio test when testing the zero restrictions versus the unrestricted models (given by column (i) plus (ii) and (v) plus (vi) respectively)

Table 4 – Parsimonious estimates: real interest rate

	1978(1)-1984(4) (i)	1985(1)-1993(1) (ii)	1993(2)-2005(1) (iii)	1978(1)-2005(1) (iv)	1978(1)-2005(1) (v)
u^r_{t-1}	0.908	0.678	0.933	0.821	0.681
(s.e.)	(0.146)	(0.227)	(0.159)	(0.099)	(0.089)
u^r_{t-2}	0.704	0.518	0.667	0.575	0.532
(s.e.)	(0.230)	(0.211)	(0.184)	(0.121)	(0.100)
u^r_{t-3}	0.786	-0.169	0.358	0.287	0.325
(s.e.)	(0.177)	(0.195)	(0.154)	(0.108)	(0.077)
$\sigma(u^r_{93-05})$	-	-	0.323	0.326	
(s.e.)			(0.033)	(0.032)	
$\sigma(u^r_{82-92})$	-	0.747	-	0.724	-
(s.e.)		(0.163)		(0.123)	
$\sigma(u^r_{78-81})$	1.169			1.267	0.792[a]
(s.e.)	(0.177)			(0.174)	(0.037)
Log L	-42.5	-37.2	-13.9	-95.6	-128.0
P-value[b]				0.905	0.00

Note: parsimonious univariate MA estimates of the real interest rate: columns (i)-(iii) report unrestricted estimates for the three subperiods 78(1)-84(4); 85(1)-93(1); 93(2)-05(1). Column (iv) reports the restricted model with variance breaks, column (v) the restricted model with a single variance.

[a] Refers to whole sample period, 1978(1)-2005(1)

[b] P-value for the Likelihood ratio test when testing the zero restrictions imposed by specification (iv) versus the unrestricted model (columns i+ii+iii), and by specification (v) versus specification (iv).

Table 5 – Consumption, income and interest rate, cohort 1 and 2, parsimonious

	Cohort 1			Cohort 2		
	Consumption	Income	Interest rate	Consumption	Income	Interest rate
u_t^x	1			1		
(se)						
u_t^y	0.370	1		0.388	1	
(se)	(0.203)			(0.156)		
u_{t-1}^y					-0.289	
(se)					(0.140)	
u_t^r	-0.727		1	-0.083		1
(se)	(0.370)			(0.459)		
u_{t-1}^r	0.910		0.826	0.829		0.777
(se)	(0.443)		(0.091)	(0.527)		(0.108)
u_{t-2}^r	0.939		0.678	-0.136		0.542
(se)	(0.480)		(0.127)	(0.537)		(0.114)
u_{t-3}^r	-0.639		0.268	0.328		0.219
(se)	(0.496)		(0.147)	(0.500)		(0.110)
u_{t-4}^r	0.393		-0.105			
(se)	(0.484)		(0.148)			
u_{t-5}^r	-0.793		-0.244			
(se)	(0.533)		(0.154)			
u_{t-6}^r	-0.009		-0.286			
(se)	(0.418)		(0.139)			
$\sigma(u_{93-05}^z)$	0.851[a]	1.539[a]	0.339	1.179[a]	2.417[a]	0.329
(s.e.)	(0.228)	(0.278)	(0.037)	(0.350)	(0.422)	(0.036)
$\sigma(u_{82-92}^z)$			0.704			0.715
(s.e.)			(0.089)			(0.095)
$\sigma(u_{78-81}^z)$			1.356			1.469
(s.e.)			(0.176)			(0.242)
$z = x, y, r$						
Log L	-474.674			-459.726		

Note: Maximum likelihood estimates of the system (3) in the text, with correction for measurement error. Sample: 1978(1)-2005(1). $\sigma(u_{t1-t2}^z)$ is the estimated square root of the variance of the unobserved component for the period t1-t2, with $z=x, y$, and r. [a] Refers to whole sample period.

Table 6 – Consumption, income and interest rate, cohort 1, by education

	Low education			High education		
	Consumption	Income	Interest rate	Consumption	Income	Interest rate
u_t^x	1			1		
(se)						
u_t^y	1.088	1		0.631	1	
(se)	(0.191)			(0.190)		
u_{t-1}^y	-0.696	-0.541				
(se)	(0.182)	(0.144)				
u_t^r	-1.170		1	-0.963		1
(se)	(0.556)			(0.525)		
u_{t-1}^r	0.995		0.841	0.831		0.826
(se)	(0.739)		(0.099)	(0.706)		(0.090)
u_{t-2}^r	0.792		0.577	1.233		0.657
(se)	(0.718)		(0.106)	(0.700)		(0.129)
u_{t-3}^r	-0.715		0.260	-0.698		0.254
(se)	(0.546)		(0.106)	(0.756)		(0.146)
u_{t-4}^r				0.596		-0.145
(se)				(0.723)		(0.143)
u_{t-5}^r				-1.470		-0.283
(se)				(0.785)		(0.142)
u_{t-6}^r				0.045		-0.285
(se)				(0.577)		(0.134)
$\sigma(u_{93-05}^z)$	0.844[a]	3.383[a]	0.330	0.719[a]	1.584[a]	0.339
(s.e.)	(0.292)	(0.635)	(0.037)	(0.313)	(0.303)	(0.037)
$\sigma(u_{82-92}^z)$			0.720			0.710
(s.e.)			(0.105)			(0.103)
$\sigma(u_{78-81}^z)$			1.421			1.346
(s.e.)			(0.186)			(0.170)
$z = x, y, r$						
Log L	-555.292			-531.514		

Note: Maximum likelihood estimates of the system (3) in the text, with correction for measurement error. Sample: 1978(1)-2005(1). $\sigma(u_{t1-t2}^z)$ is the estimated square root of the variance of the unobserved component for the period t1-t2, with $z=x, y,$ and r.
(a) Refers to whole sample period

Table 7 – Consumption, income and interest rate, cohort 2, by education

	Low education			High education		
	Consumption	Income	Interest rate	Consumption	Income	Interest rate
u_t^x	1			1		
(se)						
u_t^y	0.843	1		0.320	1	
(se)	(0.203)			(0.097)		
u_{t-1}^y	-0.555	-0.489			-0.449	
(se)	(0.208)	(0.124)			(0.123)	
u_t^r	-0.218		1	-0.474		1
(se)	(0.646)			(0.690)		
u_{t-1}^r	1.139		0.779	0.815		0.768
(se)	(0.746)		(0.100)	(0.834)		(0.109)
u_{t-2}^r	-0.570		0.534	0.384		0.531
(se)	(0.804)		(0.113)	(0.829)		(0.112)
u_{t-3}^r	0.100		0.232	0.670		0.188
(se)	(0.723)		(0.103)	(0.742)		(0.111)
$\sigma(u_{93-05}^z)$	1.384[a]	3.830[a]	0.328	0.744[a]	3.832[a]	0.331
(s.e.)	(0.433)	(0.625)	(0.032)	(0.337)	(0.577)	(0.037)
$\sigma(u_{82-92}^z)$			0.717			0.711
(s.e.)			(0.107)			(0.097)
$\sigma(u_{78-81}^z)$			1.470			1.466
(s.e.)			(0.242)			(0.256)
$z = x, y, r$						
Log L	-515.784			-532.121		

Note: Maximum likelihood estimates of the system (3) in the text, with correction for measurement error. Sample: 1978(1)-2005(1). $\sigma(u_{t1-t2}^z)$ is the estimated square root of the variance of the unobserved component for the period t1-t2, with $z=x, y$, and r.
[a] Refers to whole sample period.

Table 8 – Excess sensitivity

	Total	Low education	High education
Cohort 1	0.280	0.006	0.622
Cohort 2	0.829	0.006	0.748

Note: P-values of the Likelihood ratio tests on the coefficients on past income shocks in the consumption equation in the models in tables 5, 6, 7.

Table 9 – Elasticity of intertemporal substitution

	Total	Low education	High education
Cohort 1	1.007	0.600	1.287
(se)	*(0.259)*	*(0.309)*	*(0.341)*
Cohort 2	0.626	0.422	1.194
(se)	*(0.390)*	*(0.478)*	*(0.522)*
P-values[a]			
Cohort 1	0.498	0.191	0.413
Cohort 2	0.964	0.345	0.995

Note: estimates of the elasticity of intertemporal substitution obtained estimating by maximum likelihood the restricted versions of the models in tables 5, 6, and 7: the restrictions imposed are given by equation (9) in the text.
[a] P-value of the likelihood ratio test of the restriction imposed by (9) on the unrestricted models in tables 5, 6, 7.

Table 10 – Excess smoothness

	Total	Low education	High education
Cohort 1	-0.630	-0.073	-0.369
(se)	*(0.203)*	*(0.113)*	*(0.190)*
P- value[a]	0.001	0.261	0.028
Cohort 2	-0.254	-0.227	-0.235
(se)	*(0.145)*	*(0.122)*	*(0.098)*
P- value[a]	0.041	0.033	0.009

Note: Estimates of excess smoothness obtained from the maximum likelihood estimates of the models in table 5, 6, and 7. The discounted sum of the income coefficients has been computed at a quarterly interest rate equal to 1%.
[a] P-value of the one-sided t-test of no excess smoothness, null hypothesis: $\alpha^{xy}(R) - \alpha^{yy}(R) = 0$, against the alternative of excess smoothness: $\alpha^{xy}(R) - \alpha^{yy}(R) < 0$
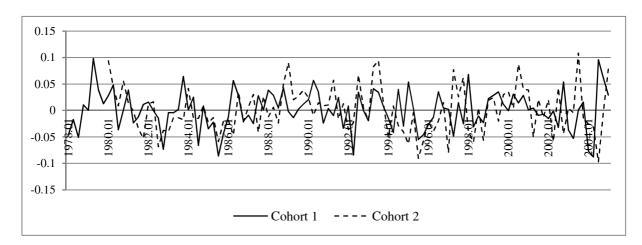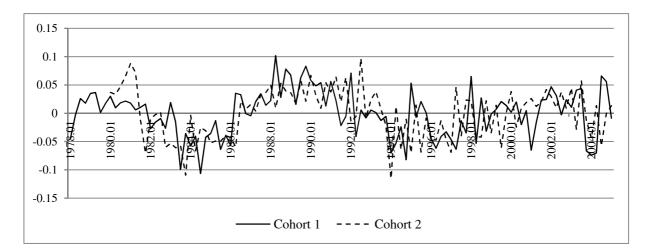
Figure 1 - Log non-durable consumption – residuals from first step regressions

Figure 2 - Log household disposable income – residuals from first step regressions

Appendix

The state space representation of the model. We consider system (4) in the text with one lag; extension to more lags is straightforward. For each group, the system may be written as:

$$\Delta \tilde{x}_t = u_t^x + \alpha_1^{xx} u_{t-1}^x + \alpha_0^{xy} u_t^y + \alpha_1^{xy} u_{t-1}^y + \alpha_0^{xr} u_t^r + \alpha_1^{xr} u_{t-1}^r + \eta_t^x - \eta_{t-1}^x$$

$$(A.1) \qquad \Delta \tilde{y}_t = \qquad\qquad u_t^y + \alpha_1^{yy} u_{t-1}^y + \alpha_0^{yr} u_t^r + \alpha_1^{yr} u_{t-1}^r + \eta_t^y - \eta_{t-1}^y$$

$$\tilde{r}_t = \qquad\qquad u_t^r + \alpha_1^{rr} u_{t-1}^r$$

The system above may be easily written in state space representation, where the state vector is:

$$\xi_{t+1} = \begin{bmatrix} u_{t+1}^x & u_{t+1}^y & u_{t+1}^r & \eta_{t+1}^x & \eta_{t+1}^y & u_t^x & \dots & \eta_t^y \end{bmatrix}'$$

of dimension equal to $(n+s)\times(q+1)+2*n=k$, where q is the number of lags, n is the number of cohort specific variables, s is the number of fixed-across-cohort variables, and the second term in the sum is the number of measurement error terms (which does not depend on the number of lags in the model).

Define:

M i.e. the number of dependent variables in the model (3 in our system);

k the dimension of the space vector, of dimension $M\times(q+1)+2*n$, where q is the number of lags and n is the number of averaged variables, for which we correct for measurement error;

$k_1=M+n$ i.e. the number of variables at time $t+1$ in the space vector plus the measurement error component at time $t+1$.

The state and measurement equations are:

$$(A.2) \qquad \begin{aligned} \xi_{t+1} &= F\xi_t + v_t \\ X_t &= H\xi_t \end{aligned}$$

where X_t is the vector of M dependent variables, H is a $(M\times k)$ matrix containing the α parameters as well as block of zeros, F is a $(k\times k)$ matrix of zero's and one's and v is the state equation disturbance vector, where the first k_1 entries are given by the variables at time $t+1$ in the state vector, and all the other entries are always zero. The variance-covariance matrix Q_t of v is a diagonal matrix apart from the entries in which there is the covariance among the measurement error terms η. Notice that Q_t is time varying, both because the measurement error variances/covariances vary through time, and because the interest rate variances have two breaks. The measurement equation has no noise, so its variance-covariance matrix is equal to zero. The log likelihood function for the unknown set of parameters θ of the model is given by:

$$(A.3) \qquad \log L(\theta) = -\frac{MT}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log|G_t| - \frac{1}{2}\sum_{t=1}^{T}\varepsilon_t' G_t^{-1}\varepsilon_t$$

which is the prediction error decomposition form of the likelihood. The prediction errors are given by:

$$(A.4) \qquad \varepsilon_t = X_t - H\hat{\xi}_{t/t-1}$$

with associated MSE:

$$(A.5) \qquad G_t = HP_{t/t-1}H' + R$$

where the matrix R is the variance-covariance matrix of the disturbance term in the measurement equation, and in this case is equal to zero. The prediction errors and their MSE's can be calculated using the Kalman filter recursions. If the system is non-Gaussian, it is still possible to maximize equation A3: this procedure yields consistent and asymptotically normal estimates, with variance-covariance matrix given by (Watson, 1989):

(A6) $$Cov(\hat{\theta}) = T^{-1}[I_{2D}I_{OP}^{-1}I_{2D}]^{-1}$$

where I_{2D} is the matrix of second derivatives:

(A7) $$I_{2D} = -\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \log L(\theta)}{\partial\theta\partial\theta'}\Big|_{\theta=\hat{\theta}}$$

And I_{OP} is the outer product estimate of the information matrix:

(A8) $$I_{OP} = \frac{1}{T}\sum_{t=1}^{T}\left[\frac{\partial \log L(\theta)}{\partial\theta}\Big|_{\theta=\hat{\theta}}\right]\left[\frac{\partial \log L(\theta)}{\partial\theta}\Big|_{\theta=\hat{\theta}}\right]'$$

Diagnostic tables. We perform unit root tests on (the residuals of) our variables of interest, and report results in table A1. We perform both the modified Dickey-Fuller (proposed by Elliott, Rothenberg, and Stock, 1996), which is based on the null hypothesis of a unit root, and the Kwiatkowski, Phillips, Schmidt, Shin (KPSS, 1992) which, on the opposite, is based on the null hypothesis of stationarity. In the table we report the value of the two statistics at the lag selected with the modified Akaike information criterion (MAIC) for each variable of interest. In the last two rows we report the critical values at the 5 and 10 per cent level. Results for the dfgls test indicate that we do not reject the null of a unit root for all our series, with the exception of the interest rate. Results for the kpss indicate that the null of trend stationarity is not rejected at the 10 per cent level for the interest rate.

Table A1 – Unit root tests

Test:	dfgls	kpss
H0:	unit root	trend stationarity
$\log(x)$ cohort 1	-1.358	0.166
$\log(x)$ cohort 2	-1.806	0.113
$\log(y)$ cohort 1	-2.052	0.128
$\log(y)$ cohort 2	-2.132	0.129
r	-5.655	0.0832
critical values:		
10%	-2.6	0.11
5%	-2.8	0.14

In tables A2-A7 we report, for all the model estimates presented in the main text, diagnostic tests based on the residuals of such models. In particular, we compute the Ljung-Box (Q) statistic for white noise (i.e. no autocorrelation) using both the residuals and the squared residuals, to detect ARCH effects. We also report a D'Agostino-Pearson test of linearity based on skewness and kurtosis in order to detect, although indirectly, departures from linearity.

Table A2 – Q- and normality test on the residuals of univariate consumption models

Cohort 1	with correction			no correction		
	78(2)-93(1)	93(2)-05(1)	78(2)-05(1)	78(2)-93(1)	93(2)-05(1)	78(2)-05(1)
Column in Table 2	(iii)	(iii)	(iv)	(vii)	(vii)	(viii)
Residuals						
Q(12)	0.674	0.822	0.976	0.760	0.828	0.982
Q(24)	0.168	0.968	0.985	0.182	0.958	0.987
Q(36)	0.309	0.955	0.985	0.358	0.973	0.987
Squared residuals						
Q(12)	0.443	0.350	0.906	0.505	0.458	0.940
Q(24)	0.595	0.561	0.939	0.642	0.719	0.965
Q(36)	0.551	0.762	0.986	0.627	0.860	0.995
Normality	0.597	0.900	0.714	0.567	0.574	0.590
Cohort 2	with correction			no correction		
	80(2)-93(1)	93(2)-05(1)	80(2)-05(1)	80(2)-93(1)	93(2)-05(1)	80(2)-05(1)
Column in Table 2	(iii)	(iii)	(iv)	(vii)	(vii)	(viii)
Residuals						
Q(12)	0.607	0.486	0.830	0.611	0.489	0.868
Q(24)	0.981	0.266	0.931	0.980	0.270	0.947
Q(36)	0.848	0.225	0.989	0.868	0.232	0.992
Squared residuals						
Q(12)	0.331	0.841	0.608	0.348	0.823	0.583
Q(24)	0.075	0.641	0.752	0.070	0.613	0.798
Q(36)	0.071	0.462	0.869	0.077	0.426	0.894
Normality	0.434	0.470	0.404	0.445	0.449	0.396

Note: Residuals from models in Table 2 in the text.

Table A3 – Q- and normality test on the residuals of univariate income models

Cohort 1	with correction			no correction		
	78(2)-93(1)	93(2)-05(1)	78(2)-05(1)	78(2)-93(1)	93(2)-05(1)	78(2)-05(1)
Column in Table 3	(iii)	(iii)	(iv)	(vii)	(vii)	(viii)
Residuals						
Q(12)	0.143	0.686	0.919	0.140	0.672	0.941
Q(24)	0.067	0.912	0.896	0.023	0.939	0.819
Q(36)	0.045	0.746	0.919	0.019	0.805	0.880
Squared residuals						
Q(12)	0.834	0.295	0.107	0.842	0.504	0.151
Q(24)	0.907	0.178	0.016	0.924	0.422	0.039
Q(36)	0.904	0.099	0.003	0.918	0.247	0.010
Normality	0.275	0.463	0.159	0.302	0.242	0.153
Cohort 2	with correction			no correction		
	80(2)-93(1)	93(2)-05(1)	80(2)-05(1)	80(2)-93(1)	93(2)-05(1)	80(2)-05(1)
Column in Table 3	(iii)	(iii)	(iv)	(vii)	(vii)	(viii)
Residuals						
Q(12)	0.524	0.187	0.347	0.224	0.086	0.304
Q(24)	0.733	0.241	0.264	0.422	0.132	0.225
Q(36)	0.542	0.226	0.370	0.233	0.123	0.316
Squared residuals						
Q(12)	0.398	1.000	0.200	0.344	1.000	0.178
Q(24)	0.806	0.983	0.534	0.794	0.985	0.505
Q(36)	0.911	0.997	0.545	0.915	0.996	0.525
Normality	0.136	0.370	0.187	0.115	0.711	0.182

Note: Residuals from models in Table 3 in the text.

Table A4 – Q- and normality test on the residuals of univariate interest rate models

	Model (v)	Model (iv)	Model (iv)	Model (iv)
Time period	78(2)-05(1)	78(2)-85(1)	85(2)-93(1)	93(2)-05(1)
Residuals				
Q(12)	0.593	0.807	0.820	0.878
Q(24)	0.732	0.780	0.447	0.937
Squared residuals				
Q(12)	0.0001	0.493	0.856	0.669
Q(24)	0.0004	0.851	0.819	0.581
Normality	0.0005	0.332	0.838	0.196

Note: Residuals from models in Table 4 in the text.

Table A5 – Q- and normality tests on residuals from cohort 1 and 2, total

	Cohort 1					Cohort 2				
	x	y	r78-81	r81-93	r93-05	x	y	r78-81	r81-93	r93-05
Residuals										
Q(12)	0.911	0.885	0.541	0.6094	0.7164	0.7133	0.2948	0.8859	0.8047	0.8568
Q(24)	0.997	0.932		0.5843	0.7026	0.8748	0.2195		0.4606	0.9201
Q(36)	0.990	0.946				0.9644	0.3503			
Squared residuals										
Q(12)	0.913	0.133	0.773	0.927	0.226	0.6214	0.1755	0.9996	0.5667	0.6353
Q(24)	0.723	0.027		0.972	0.174	0.5747	0.4981		0.5566	0.4917
Q(36)	0.947	0.003				0.6969	0.4811			
Normality	0.619	0.112	0.725	0.466	0.343	0.1724	0.3011	0.0081	0.7061	0.2312

Note: Residuals from model in Table 5 in the text.

Table A6 – Q- and normality tests on residuals from cohort 1, by education

	Lower education					Higher education				
	x	y	r78-81	r81-93	r93-05	x	y	r78-81	r81-93	r93-05
Residuals										
Q(12)	0.657	0.971	0.504	0.722	0.916	0.759	0.870	0.633	0.580	0.764
Q(24)	0.666	0.613		0.377	0.953	0.604	0.884		0.577	0.716
Q(36)	0.734	0.822				0.790	0.980			
Squared residuals										
Q(12)	0.004	0.605	0.565	0.823	0.609	0.757	0.016	0.804	0.944	0.221
Q(24)	0.004	0.512		0.702	0.532	0.900	0.061		0.981	0.174
Q(36)	0.055	0.050				0.683	0.213			
Normality	0.696	0.802	0.760	0.830	0.241	0.598	0.057	0.731	0.455	0.357

Note: Residuals from model in Table 6 in the text.

Table A7 – Q- and normality tests on residuals from cohort 2, by education

	Lower education					Higher education				
	x	y	r78-81	r81-93	r93-05	x	y	r78-81	r81-93	r93-05
Residuals										
Q(12)	0.4535	0.3996	0.8952	0.8142	0.8628	0.9465	0.2571	0.877	0.803	0.8208
Q(24)	0.5743	0.6655		0.4738	0.9252	0.9994	0.6327		0.4448	0.9047
Q(36)	0.1783	0.3471				0.9924	0.3023			
Squared residuals										
Q(12)	0.5318	0.8675	0.9997	0.6097	0.6765	0.9908	0.4113	0.9996	0.5147	0.6051
Q(24)	0.7707	0.9483		0.6093	0.542	0.9388	0.4617		0.4978	0.4342
Q(36)	0.8696	0.9174				0.9431	0.2729			
Normality	0.4633	0.1961	0.0082	0.7476	0.207	0.2689	0.5187	0.0077	0.6452	0.2583

Note: Residuals from model in Table 7 in the text.