Stochastic Components of Individual Consumption:

A Time Series Analysis of Grouped Data[♣]

Orazio P. Attanasio[*] & Margherita Borella[**]

This Draft: June 2008

Abstract

In this paper we propose a method to characterize the time series properties of individual consumption, income and interest rates using micro data, as studies in labour economics have characterized the time series properties of hours and earnings. Our approach, however, does not remove aggregate shocks. Having estimated the parameters of a flexible multivariate MA representation we relate the coefficients of our statistical model to structural parameters of theoretical models of consumption behaviour. Our approach offers a unifying framework that encompasses the Euler equation approach to the study of consumption and the studies that relate innovations to income to innovations to consumption, such as those that have found the so-called excess smoothness of consumption. Using a long time series of cross sections to construct synthetic panel data for the UK, we estimate our model and find that the restrictions of Euler equations are typically not rejected. We also fail to reject the hypothesis of no "excess smoothness".

JEL: E2, D1, C3
Keywords: Consumption, Euler Equations, Life Cycle Model

1.   Introduction

The time series properties of aggregate consumption (and some of its components) are well known in the time series literature. The papers by Sargent (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. Indeed, some of the properties uncovered by this literature have stimulated the development of different theoretical models. Very little is known, however, about the stochastic properties of consumption at the individual level.

In this paper, we propose a new methodology to analyze the time series properties of individual consumption expenditure, jointly with those of income and other (possibly aggregate) variables of interest using 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 several authors, including MaCurdy (1982), Abowd and Card (1989), Moffitt and Gottshalk (1995) and Meghir and Pistaferri (2002). These studies model only labour market variables. Moreover, they focus on the dynamic properties of purely idiosyncratic components and treat aggregate shocks as a nuisance parameter that is eliminated, together with deterministic life cycle effects, in preliminary regressions. Our contribution, instead, studies mainly consumption. 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 by time dummies.

Relative to many papers that have studied the time series properties of consumption expenditure, a novelty of our contribution lies in our use of micro level data. This is important for two reasons. First, we can, at least in principle, look at the dynamics of individual consumption data. Comparing and contrasting the time series properties of individual (or group level) and aggregate consumption can be informative about the nature of insurance and credit markets available to consumers or, more generally, on the smoothing mechanisms observed in the data. Second, even when we work with aggregates, 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). This second aspect is crucial if we want to relate findings to the parameters of structural models of consumption behaviour.

We represent the time series behaviour of 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 model. Given a set of identifying

restrictions, we will be able to identify both patterns of dynamic dependence and some contemporaneous correlations. Having characterized these properties of individual consumption, income and interest rates, we map the pattern of correlations that emerges 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.

In the last 25 years, many empirical studies of consumption behaviour have focused on some version of an Euler equation for intertemporal optimization and have estimated structural parameters and tested the model by exploiting the over-identifying restrictions implied by such an equation. These studies, too numerous to be listed here, have 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 consumption. In Section 5 of this paper, we show what are the restrictions that the Euler equation for consumption imposes on our MA representation of individual 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. We develop these implications of our approach in Section 5, where we provide an interpretation for the coefficients whose estimates are reported in Section 4.

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. And yet our approach can identify some of the contemporaneous correlation 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 these papers concluded that consumption was 'excessively smooth' in that consumption did not react enough to permanent innovation 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 (2006), 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 (2006) 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.

Most of the studies in the labour economics literature are based on large N asymptotics, as they exploit the cross sectional variability to identify the parameters of interest. Because of this, some of them, such as Abowd and Card (1989), allow the coefficient of interest to vary over time. Our study, instead, focuses on the time series properties of individual 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.[1]   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.

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 section 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. [2] We use surveys from 1974 to 2000 to form quarterly observations of consumption and income at the cohort level. For comparison with the existing literature, we also consider aggregate National Account data.

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 aggregate shocks become really the focus of our 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 transformation 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.

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 describe the data sources, while in section 4 we report the estimates of our statistical models. In section 5, we present a simple theoretical framework that can be used to interpret the results in Section 4. Within this framework, we show how to derive the so-called 'excess sensitivity' and 'excess smoothness'

---

[1] 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.

tests. In section 6, we present possible extensions to the basic theoretical framework described in Section 5. Section 7 concludes the paper.

2.   The Methodology

As stressed in the introduction, the main aims of this study are two. First, we would like to identify innovations to consumption and other variables using both micro and aggregate data 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. It will be important to stress the differences obtained with aggregate and micro data, as they will be informative about the relevance of aggregation issues, as discussed in Attanasio and Weber (1993).

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.

As we focus on business cycle fluctuations, we remove from our data all age or cohort effects. Effectively, our methodology removes all deterministic trends from the data and interprets them as arising from a combination of cohort and age effects. This leaves us with group specific shocks whose covariance structure is the focus of this study. It should be stressed, however, that, unlike many studies in the literature on the micro dynamics of earnings, we do not remove time effects. Indeed, the business cycle frequency is both important for our identification strategy and at the centre of our focus.

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.

As the nature of our data forces us to work with grouped data, effectively our model is 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, our sample averages are affected by measurement error arising from the limited size of our sample. The latter source of variability constitutes a nuisance which can, however, be controlled for given the information on the sample structure and given the

---

[2] Cohorts, however, are not the only interesting group that can be formed. One can consider education or occupation groups. Differences in the variability, persistence and covariance structures across these groups can be quite interesting.

information on within cells variability (see Deaton, 1985). We discuss these issues in the second part of this section.

2.1.   The statistical model

Let us consider a generic variable z_{ct}^h, where the index h denotes the individual household, the index t the time period and the index c the group to which household h belongs. Without loss of generality, such a variable can be written as:

(1) $z_{ct}^h = \bar{z}_{ct} + \eta_{ct}^h$

where $\bar{z}_{ct}$ denotes the mean of the variable x for group c at time t. Given a sample in which group membership is observable, the mean in equation (1) can be easily estimated by the sample means. In the next subsection we discuss the problems that arise because we do not observe $\bar{z}_{ct}$ but are forced to estimate it on the basis of samples of limited size. In this subsection, we treat $\bar{z}_{ct}$ as observable.

Notice that the variables in (1) can be non-linear transformations of the variables of interest. As we work with micro data, we can control the aggregation directly. This turns up to be important for at least three reasons. First, theoretical models often imply relationships among non linear functions of variables. Therefore, to give a structural interpretation to the results we present it will be crucial to control aggregation. Second, in the case of corners, the theoretical relationships one might want to consider at the aggregate level involve both means conditional on not being at a corner and overall means. Third, within cell heterogeneity, that is the variability of η_{ct}^h, is not only useful to correct for the measurement error in the estimation of $\bar{z}_{ct}$, but can also be informative to study the evolution of the inequality over time.

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 $\bar{z}_{ct}$. Any purely idiosyncratic persistence, embedded in η_{ct}^h, 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 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. Having said that, however, it should be stressed that even a relatively short genuine panel might be sufficient to

estimate a model which requires big T asymptotics if repeated panels are available: one can then group the cross moments and follow their dynamics over time.

The first step of our procedure consists in removing all deterministic trends from the variables of interest. Denoting as $\tilde{z}_{ct}^{h}$ the variables before de-trending, these are likely to be affected by time, group and age effects. As within the framework of a life cycle model groups are often formed on the basis of the year of birth of the household head, group effects are essentially cohort effects. This implies the impossibility of disentangling age, time and cohort effects. In what follows we label all deterministic trends in the data as 'age and cohort' effects. While this label is arbitrary, disentangling the various effects is not the aim of the study which focuses, instead, on modeling the innovations to income and consumption at the business cycle frequency.

To remove deterministic changes, therefore, we regress the variables of interest on a cohort specific second order polynomial in age with cohort specific intercepts. Because we use quarterly data, we also include seasonal effects in our first step regression.

$$(2) \qquad \tilde{z}_{ct}^{h} = \delta_c + \sum_{i=1}^{3} \alpha_{ci} q_{ti} + f^{c}(t-c) + z_{ct}^{h}$$

where $f^{c}(t-c)$ is the cohort specific polynomial in age (obtained as time - year of birth), q's are quarterly seasonal dummies and δ_c the cohort specific intercepts. Equation (2) is estimated by OLS. The cohort average of the residual, z_{ct}, reflects both genuine time variation in group averages and measurement error arising from the limited sample sizes in computing the group averages.[3]

It should be stressed again that equation (2) does *not* contain 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 cohort 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 an equation like (2). Having estimated the parameters of such a regression, we interpret the cohort averages of the estimated residuals z_{ct} as deviations of the average cohort data from the deterministic trends present in the data. It is the cohort averages of these residuals that we model and study.

---

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

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 z_{ct} and study its time series properties. Consider three vectors of observable variables, *x, y and r*, of dimension *m, n* and *p*, respectively. For each of the cohorts we consider, the model we estimate is the following:

$$(3) \qquad \Delta x_t = \sum_{j=0}^{q_{xx}} \alpha_j^{xx} u_{t-j}^x + \sum_{j=0}^{q_{xy}} \alpha_j^{xy} u_{t-j}^y + \sum_{j=0}^{q_{xr}} \alpha_j^{xr} u_{t-j}^r$$

$$\Delta y_t = \qquad\qquad \sum_{j=0}^{q_{yy}} \alpha_j^{yy} u_{t-j}^y + \sum_{j=0}^{q_{yr}} \alpha_j^{yr} u_{t-j}^r$$

$$r_t = \qquad\qquad\qquad\qquad \sum_{j=0}^{q_{rr}} \alpha_j^{rr} u_{t-j}^r$$

where we dropped the cohort subscript *c* for ease of notation. The shocks u_t^x, u_t^y and u_t^r are of dimensions *m, n* and *p*, respectively and the matrices of coefficients of the appropriate dimensions. The variables *x* are choices for the individual households: while in the basic specification we estimate *x* is non durable consumption (and *m*=1), in general *x* can include several variables such as, for instance, different components of consumption, hours of work and participation rates by adding more equations to the system. These variables are affected by all the shocks present in the system. The vector *y* includes variables that are household (or cohort) specific but are assumed not to be determined by individual choice, at least at the frequency we are considering. These could include, for instance, wage rates for male and females, and they can be group specific. In the specific model we estimate we consider *y* to be income. 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). The vector of *r* variables are also not determined by individual choices and, moreover, do not vary across individuals. We think of these 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.

The vector of unobservable random variables *u* is assumed to be stationary, uncorrelated over time and with a diagonal variance-covariance matrix. A more detailed exposition of the type of model we estimate is given in the Appendix.

The MA structure in equation (3) is quite general, but imposes 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 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 all the components of the diagonal of the matrix α^{xx}_0 are equal to one.

In system (3) the triangular restriction is imposed also on the lags of the relevant shock, but it is done only for ease of notation. As we discuss in Section 5, theory imposes several (over-identification) restrictions on the lag coefficients which we will exploit to give a structural interpretation to our time series model.

2.2.   Estimation

There are several ways in which one can estimate the model (3). By making assumptions on the distribution of the shocks that enter the system (3) it is possible to compute the likelihood function associated with a given sample and estimate the parameters of interest by maximizing such a function. It would be also 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, especially when trying to correct for the presence of measurement error, which we discuss below. For this reason we adopted the former method. The estimates we present assume that the residuals in system (3) 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.

So far, for ease of exposition, we have assumed that the group means of the variables of interest are observed. As, instead, they are estimated using samples of limited size, we have to consider the measurement error problem that this induces. That is, as the sample used in estimating the cohort averages is refreshed in each period, the error induced by sample variation implies an

MA(1) structure in the *changes* in the variable. When considering the empirical counterpart of system (3) we will need to take this structure into account explicitly. Given that we know the cell size and we can estimate the within cell variance, we have a substantial amount of information on the measurement error that we can use to correct the estimated sample moments.

More specifically, the changes in the variables we model can be decomposed into two components,

(4) $\qquad \Delta z_{ct} = \Delta \bar{z}_{ct} + \Delta \eta_{ct}$

where η_{ct} is the error induced by the sample variability and that we will call 'measurement error'. We can estimate its variance consistently from the within cell variability and the cell size. We will assume that is independent of $\Delta \bar{z}_{ct}$, the changes in the innovations we are interested in modelling. The time series variance of $\Delta \bar{z}_{ct}$ can then be consistently estimated by the sample equivalent of the following expression:

(5) $\qquad Var(\Delta \bar{z}_{ct}) = Var(\Delta z_{ct}) - E\left[\dfrac{2}{N_{ct}} Var_i(\eta_{ct}^h)\right]$

where the second variance has a subscript '*i*' to stress that is a cross sectional covariance.[4]

The within cell variance can be estimated from the micro data. If we assume that the variance within cells is constant, one can estimate such a variance very efficiently. If one assumes that the cross-sectional variances in equation (5) are known, one can considerably simplify the maximization of the likelihood function. We estimate such variances under the assumption that they are constant across cell and can be estimated using the entire sample. As we have about 12,000 observations for each cohort, we treat such variances as known.

2.3. Comparison with the existing literature

One of the first papers to use panel data to infer the time series properties of individual level variables is the widely cited paper by MaCurdy (1982) in which the author models deviations of wages and earnings from a regression equation including several variables (such as age and

---

[4] The latter variance will be different in each cell and at each point in time, if nothing else because cell sizes can be quite different. These variances can be averaged out to get a mean of the unconditional variance needed for the correction.

demographics) and a set of period dummies. These deviations are found to be well represented by an IMA(2) model, where the second coefficient of the MA component, while significantly different from zero, is estimated to be quite small. MaCurdy (1982) uses data from the PSID. The same survey was later used by Abowd and Card (1989) who, generalizing MaCurdy's approach, allow for non stationarity of the processes for earnings and wages. While they also found that the autocovariances of order higher than two are not significantly different from zero (and small in magnitude), they reject the hypothesis of stationarity. Abowd and Card (1989) also related their findings about correlations and autocorrelations in terms of alternative structural models, performing a variance decomposition exercise[5].

Altonji et al. (2002) probably constitutes the study closest to the present one. They consider the covariances and autocovariances of wages, hours and food consumption in the PSID and use them to estimate several factor models which are then given a structural interpretation.

The main conceptual difference between the existing studies in the literature and our approach is the fact that we do not remove time effects and, indeed, focus on the modelling of business cycle frequency shocks. For this reason we need to rely on T-asymptotics and for this reason our level of flexibility in analyzing non-stationarity is limited.[6] However, it is important to stress that the difference is not merely technical but of focus: we aim at identifying the nature of business cycle shocks to income and how these are absorbed by consumption and, possibly, its components and other variables. In other words, ours is an attempt to model the time effects that are removed in other studies. Our approach might be particularly informative in characterizing smoothing mechanisms and to interpret the shocks we identify in terms of underlying structural models.

The necessity of large T-asymptotics is also consistent with the same requirement in the estimation of Euler equations in the absence of complete markets.[7] In Section 5, we show how our estimates can be used to derive some of the same parameters typically estimated in the Euler equation approach. In particular, we will see what restrictions the Euler equation imposes on the

---

[5] Moffitt and Gottshalk (1995) have used the evolution in the cross sectional variances in earnings to identify permanent and transitory components.

[6] Throughout the study we assume stationarity, that is that the first and second moments we will be considering do not vary over time. This might seem unfortunate as there is some evidence of non stationarity in micro studies such as Abowd and Card (1989), Altonji et al (2002) and Meghir and Pistaferri (2004). However, it should be stressed, once more, that we are modeling time series rather than cross sectional variability and, therefore, the moments we are considering are conceptually different from those identified by the papers cited. Given that our estimators are based on T-asymptotics, considering the possibility of time-varying moments would involve the parametrization of their changes. We have not pursued this line of research.

parameters of our model and how, when these are satisfied, some of the parameters of our model have a direct structural interpretation. We will also show how our approach goes beyond what can be learned from an Euler equation. Identification of the contemporaneous correlation between income and consumption allows us to assess whether consumption reacts to income shocks in a way that is consistent with specific models.

The second main difference between our procedure and those in the studies cited above is that we focus on grouped data, while the others use individual data. The main reason of our exclusive attention to grouped data is the lack of a longitudinal dimension in our data. This forces us to give up the possibility of identifying any purely idiosyncratic dynamics. However, the focus on (a fixed number of) groups and on the time series variability gives a greater flexibility in modelling cross sectional heterogeneity (i.e. across groups, in our case) and a better chance to identify genuine time series uncertainty (as distinct from cross sectional heterogeneity). Increasing the number of groups would allow one to use N-asymptotics and give greater flexibility in the analysis of non-stationary time series processes. In what follows we do not make any use of N-asymptotics arguments.

The approach we propose can attempt at modelling the dynamics of several components of consumption and other relevant variables simultaneously. While other studies were limited by the availability of consumption measures (the PSID contains only information on food expenditure), the FES data set allows us to explore a much richer set up.

Finally, we only use the information on within group heterogeneity to correct the estimated sample moments for the small sample variability and measurement error. As pointed out by several authors, the evolution of within cell heterogeneity over time and age can be quite informative for a number of reasons. Deaton and Paxson (1994), for instance, relate the evolution of the within cell variance in consumption and income and check that the implications of the life cycle models for these cross sectional moments are satisfied. Blundell and Preston (1998), instead, use the differences between changes in the cross sectional variance of consumption and income to identify the variance of the permanent and transitory component of income.

A number of studies in the macro time series literature are slightly related to what we are doing. In particular, several studies papers have tried to identify the number of common factors in disaggregated business cycle models (see Watson and Engle (1983), Quah and Sargent (1994),

---

[7] Notice that, at least in principle, we can allow for group specific fixed effects in all our equations without much

Forni and Reichlin (1996)). Of more direct relevance, instead, are those studies that estimate time series representations for consumption and income processes using aggregate data. These studies include Sargent (1978), Flavin (1981), Campbell (1987), West (1988) Campbell and Deaton (1989). Of particular relevance is the paper by Hansen, Roberds and Sargent (1991).

3. Data

The data used in the estimation are drawn from the UK Family Expenditure Survey (FES) from 1978(1) to 2005(1). 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 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, whose head is an employee. We further selected 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. Table 1 reports more details.

---

difficulty. Situations in which this is appropriate are discussed in section 5.

Table 1 – Cohort definition

Cohort	Year of Birth	Observation period	Median age in first year	Median age in last year	Mean Cell Size
1	1947-53	1978-2005	28	55	122
2	1954-60	1978-2005	21	52	102

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

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.

In order to remove deterministic trends and seasonal effects, we regress (the logarithm of) the variables of interest on a second order polynomial in age and on quarterly dummies, as well as on the logarithm of family size. These estimates, which correspond to equation 2 above, were carried out by OLS, for each of the two cohorts separately.

Having obtained a cohort 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 augmented Dickey-Fuller, the GLS Dickey-Fuller, and the Kwiatkowski, Phillips, Schmidt, Shin (KPSS) tests. They all give indication of the existence of a unit root in all the series considered, i.e. non-durable expenditure and disposable income for both cohorts. This justifies our decision to model the first differences of these series.

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 deterministically de-trended, with a regression on a polynomial in time[8].

---

[8] We also estimate our models using a first-step Hodrick-Prescott filter on the interest rate: this procedure does not affect our results.

In addition to the micro data, as we detail below, we also use (de-trended) aggregate non-durable expenditure and disposable income series to estimate aggregate versions of our models: these data are drawn from the National Account Statistics quarterly data from 1978(1) to 2005(1).

4. 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, it is worthwhile to present briefly some of the time series properties of the data we use in estimation. In table 2, we report, for each of the two cohorts considered, a measure of the time series volatility of the log changes in income and consumption. In particular, we report the standard deviation (times 100) of the changes in de-trended log consumption and income. In the first two columns we report the raw standard deviation of our pseudo-panel. In Columns (3) and (4) we correct the figures in the first two columns for our estimate of the variability induced by changes in sample composition. As it can be seen, given the sample size of our cohorts and the heterogeneity within cell, the time series volatility of both consumption and income is more than halved once we correct for 'sampling variation'.[9] We compute this correction as shown in equation 5. After taking into account the correction for measurement error, non durable consumption appears less volatile than disposable income, as with aggregate data. The figures for aggregate data are reported in the last column of the table. The volatility of both consumption and income is much lower than for either cohort considered, even relative to the figures in which we attempt to control at the sampling errors induced by relative small sample sizes.

Table 2 – Volatility

	Total		After correction		Aggregate
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	data
Non-durable exp.	6.05	5.45	2.54	1.76	0.59
Income	5.39	5.95	3.55	2.94	1.64

[9] In the table, and in what follows, we refer to the correction, somewhat improperly, as a correction for measurement error. The measurement error refers to the measurement of the population mean of cohort consumption and income and it is induced by the finite size of samples used in estimating them.

Note: Individual data drawn from FES, sample 1978(1) – 2005(1). Aggregate data are drawn from the National accounts statistics, 1978(1) – 2005(1).

4.1. Univariate models

Before considering estimation of systems of equations as in (3), we report, in tables 3 to 5, estimates of separate MA models for consumption, income, and the interest rate. For the first two variables we use both micro and macro data. The models estimated on micro data are specified as univariate MA processes as the variables for the two cohorts are considered separately. In particular, table 3 shows estimates of univariate parsimonious MA representations for non durable expenditure for each cohort, both without and with correction for measurement error, as well as the estimates with aggregate data. The model estimated is therefore:

$$\Delta x_{ct} = \sum_{j=0}^{q} \alpha_{c,j}^{xx} u_{c,t-j}^{x} \qquad\qquad c=1, 2$$

where x_{ct} is the logarithm of non durable expenditure for cohort c at time t, and q is the number of lags. The measurement error term has been suppressed for ease of notation, and it is either disregarded (estimates without correction) or treated as shown in the appendix (estimates with correction for measurement error). The coefficient on the lag zero shock (α_0^{xx}) is constrained to be one, and it is not shown in the tables. The tables also report estimates of the square root of the variance of the disturbance.[10]

---

[10] All variables have been multiplied by 100 in estimation.

Table 3 – Parsimonious estimates: change in non-durable consumption.

	Cohort 1 Without Correction	Cohort 1 With Correction	Cohort 2 Without Correction	Cohort 2 With Correction	Aggregate Data
u^x_{t-1}	-0.743 (0.080)	-0.426 (0.193)	-0.658 (0.075)	-0.352 (0.178)	-0.051 (0.116)
u^x_{t-2}	-0.048 (0.119)				0.202 (0.113)
u^x_{t-3}	-0.097 (0.105)				0.107 (0.114)
u^x_{t-4}	0.224 (0.089)				0.046 (0.077)
u^x_{t-5}					-0.259 (0.085)
Standard Dev. (u^x_t) (s.e.)	4.775 (0.313)	2.266 (0.706)	4.584 (0.350)	2.389 (0.761)	0.750 (0.034)
Log L	-222.8	-224.8	-218.4	-219.7	-229
P-value 1 restriction (a)	0.060	0.000	0.000	0.000	0.047

Note: parsimonious univariate MA estimates of the change in log non-durable consumption, using cohort data – without/with correction for measurement error – and aggregate data. Standard errors in parentheses. (a) P-value for the Likelihood ratio test when testing the zero restriction on the maximum lag coefficient for each model.

The first two columns in table 3 report results for the first cohort estimated without/with correction for measurement error. Analogously, the third and fourth columns report the estimates for the second cohort without/with measurement error correction. The comparison between estimates based on micro data with or without correction reveals how measurement error influences estimates of the parameters. The last column reports estimates based on aggregate National Account data.

For the eldest cohort we find that, without correcting for measurement error, the first difference of the logarithm of non-durable expenditure follows an MA(4) process: when performing a likelihood ratio test, as reported in the first column in table 3, the restriction on the lag four coefficient to be zero is rejected with a P-value of 0.06. When we include the correction for measurement error, the non-durable series seems to be best captured by an MA(1).

For the second (younger) cohort we find that the dynamic of the first difference of non-durable expenditure is in this case much less prolonged, with an MA(1) being the preferred specification

in both cases. In addition, the p-values for the restriction to zero of the first lag coefficient are close to zero in both models.

The results obtained on aggregate data indicate that the first difference of non-durable expenditure follows an MA(5) process; the p-value of the zero restriction on the coefficient is 4.7%.

The evidence from this Table clearly indicates that the synthetic cohort data present very different dynamics, especially with respect to aggregate data. The possible explanations for these differences are many, ranging from aggregation issues to measurement error in micro consumption data or to the possibility of group level taste shocks that are washed away in aggregate data.

Table 4 reports results analogous to those in Table 3, but for (the first difference of the logarithm of) disposable income. In this case the cohort data always indicate that disposable income is an MA(1) with a negative coefficient, with the exception of the model for the youngest cohort in column 3, where no correction for measurement error is done. For aggregate data, the disposable income series displays a prolonged dynamics, as we are unable to restrict further the MA(5) specification. In this case the first coefficient is negative, but the sum of the five coefficients is positive and close to zero (equal to 0.09).

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) and with micro data (MaCurdy, 1982, Abowd and Card, 1989). Even after correcting for measurement error, our coefficient is larger in absolute value than what reported in these studies. Other studies, such as Blinder and Deaton (1985) using US aggregate data, find a positive autocorrelation in the differentiated series of income.

Table 4 – Parsimonious estimates: change in disposable income.

	Cohort 1 Without Correction	Cohort 1 With Correction	Cohort 2 Without Correction	Cohort 2 With Correction	Aggregate Data
u^Y_{t-1}	-0.680 (0.056)	-0.506 (0.104)	-0.898 (0.090)	-0.400 (0.192)	-0.454 (0.085)
u^Y_{t-2}			0.213 (0.100)		0.194 (0.108)
u^Y_{t-3}					0.004 (0.104)
u^Y_{t-4}					0.004 (0.110)
u^Y_{t-5}					0.339 (0.100)
Standard Dev. (u^Y_t)	4.328 (0.224)	2.816 (0.363)	4.385 (0.260)	1.966 (0.747)	1.124 (0.076)
Log L	-210.0	-212.4	-213.6	-216.1	-66.6
P-value 1 restriction (a)	0.000	0.000	0.027	0.000	0.001

Note: parsimonious univariate MA estimates of the change in log disposable income, using cohort data – without/with correction for measurement error – and aggregate data. Standard errors in parentheses.
(a) P-value for the Likelihood ratio test when testing the zero restriction on the maximum lag coefficient for each model.

Table 5 – Parsimonious estimates: interest rate

	Coefficient	Standard error
u^R_{t-1}	0.595	(0.075)
u^R_{t-2}	0.309	(0.077)
u^R_{t-3}	-0.119	(0.082)
u^R_{t-4}	-0.431	(0.090)
u^R_{t-5}	-0.501	(0.100)
u^R_{t-6}	-0.374	(0.099)
Standard Dev. (u^R_t)	0.827	(0.041)
Log L	-33.48	
P-value 1 restriction (a)	0.011	

Note: parsimonious univariate MA estimates of the real interest rate.
(a) P-value for the Likelihood ratio test when testing the zero restriction on the maximum lag coefficient for each model.

Table 5 reports estimates for a parsimonious univariate MA model for the interest rate. The interest rate dynamics is best captured by a six-lag specification; the p-value of the zero restriction on the sixth-lag coefficient is 1.1%.

4.2.   Multivariate models

Having estimated the univariate models, we now move on to the estimation of the main model discussed in Section 2, equation (3). 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. We thus perform estimates of the three-equation system for each cohort separately; these systems all include correction for the measurement error. 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 tables 6 and 7 for cohort 1 and 2 respectively. Given the interpretation of the results we give in the next section, we paid particular attention to the coefficients on lag 'income shocks' in the specification for consumption. Hence for both cohorts we report on the left panel of the table a specification which includes the one-lag shock to income in the consumption equation; on the right panel of the table we report the specification constraining the coefficient of that shock to zero. 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 analysed in subsequent sections.

For both cohorts we fail to 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 0.123 and 0.53 for the first and second cohort respectively. In the tables we only report the estimates for the first lag, but we also tried with additional lags obtaining 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 both tables we report, in the right hand side panel, the results of a specification that does not contain lagged income and consumption shocks.

As for the interest rate shocks, the results for the two cohorts are also very similar. Essentially, we only identify two significant lags for the interest rate shock in the consumption equation: lag one for the elder cohort and lag six for the younger cohort. This evidence is in line with the

univariate model 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. However, for both cohorts, the coefficient on the contemporaneous interest rate shock is insignificantly different from zero.

As for the income equations, consistently with the evidence on the univariate models, we do not find a very long dynamics. 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 presented by other authors for the US.

The only variable that exhibits long persistence is the real interest rate. While we estimate the process twice, once for each of the two cohorts, we do obtain very similar patterns of coefficients in the two cases. This is not surprising as the other equations (for cohort specific variables) do not impose any restrictions on the interest rate evolution.

As the two cohorts present very similar dynamics, in table 8 we present the estimates of a constrained model, in which the coefficients capturing the impact of the cohort-specific shocks are constrained to be the same across the two cohorts[11]. The results essentially confirm the cohort-specific estimates, and in particular we fail to reject the hypothesis that lagged 'income' shock is equal to zero in the consumption equation with a p-value of 0.65.

In Table 9, we report the estimates we obtain estimating system (3) on aggregate data. The results are dramatically different from those obtained on the micro data. First, the five lag consumption shock is significant in the consumption equation. Second, still in the consumption equation, the first and second lag income shocks are also strongly significant. Unlike with the micro data, the hypothesis that the coefficient on the lagged income shocks in the consumption equation is equal to zero is rejected at any standard level of significance. Finally, we do not find any significant role for interest rates shocks in the consumption equation.

---

[11] It is worth noticing that this "constrained" model is not nested to the "unconstrained" one, that is, to the sum of the models presented in tables 6 and 7, as an equation for the interest rate is added to each single-cohort model. In other words, the "unconstrained" model is estimated inefficiently (systems A1 and A3 in the appendix).

Table 6 – Consumption, income and interest rate, cohort 1, parsimonious

Equation:	Non durable	Income	Interest rate	Non durable	Income	Interest rate
u_t^x	1			1		
(se)	-			-		
u_t^Y	0.721			0.457	1	
(se)	(0.272)			(0.132)	-	
u_{t-1}^Y	-0.374	1		-	-0.391	
(se)	(0.253)	-			(0.177)	
u_t^R	-0.849	-0.513	1	-0.752	-	1
(se)	(0.695)	(0.137)	-	(0.700)		-
u_{t-1}^R	1.756	-	0.581	1.647	-	0.578
(se)	(0.962)		(0.083)	(0.984)		(0.083)
u_{t-2}^R	0.127	-	0.297	0.107	-	0.294
(se)	(0.872)		(0.079)	(0.894)		(0.080)
u_{t-3}^R	-0.391	-	-0.116	-0.331	-	-0.117
(se)	(0.857)		(0.089)	(0.834)		(0.090)
u_{t-4}^R	-0.263	-	-0.399	-0.330	-	-0.393
(se)	(0.733)		(0.094)	(0.730)		(0.095)
u_{t-5}^R	-0.214	-	-0.490	-0.140	-	-0.484
(se)	(0.847)		(0.117)	(0.847)		(0.117)
u_{t-6}^R	-0.109	-	-0.391	-0.128	-	-0.392
(se)	(0.614)		(0.100)	(0.627)		(0.103)
Standard Dev. ($u_t^{\cdot}$)	0.734	2.831	0.827	0.732	2.431	0.828
(se)	(0.201)	(0.401)	(0.042)	(0.196)	(0.433)	(0.042)
Log L	-438.5			-439.7		
P-value				0.123		

Note: Maximum likelihood estimates of the system (3) in the text, with correction for measurement error. Sample: 1978(1)-2001(1). The P-value reported is the P-value of the likelihood ratio test on the zero restriction for the coefficient u_{t-1}^Y on the consumption equation (left panel: unrestricted estimates; right panel: restricted estimates).

Table 7 – Consumption, income and interest rate, cohort 2, parsimonious

	Non durable	Income	Interest rate	Non durable	Income	Interest rate
u_t^x	1			1		
(se)						
u_t^Y	0.127	1		0.350	1	
(se)	(0.482)			(0.276)		
u_{t-1}^Y	0.262	-0.373		-	-0.433	
(se)	(0.555)	(0.341)			(0.237)	
u_t^R	-0.977	-	1	-0.993	-	1
(se)	(0.699)			(0.664)		
u_{t-1}^R	0.618	-	0.592	0.632	-	0.593
(se)	(0.893)		(0.086)	(0.847)		(0.085)
u_{t-2}^R	-0.056	-	0.303	-0.031	-	0.303
(se)	(0.958)		(0.089)	(0.921)		(0.089)
u_{t-3}^R	0.595	-	-0.107	0.563	-	-0.108
(se)	(0.811)		(0.091)	(0.779)		(0.091)
u_{t-4}^R	0.370	-	-0.448	0.385	-	-0.451
(se)	(0.878)		(0.103)	(0.854)		(0.102)
u_{t-5}^R	-0.057	-	-0.494	-0.092	-	-0.497
(se)	(1.233)		(0.112)	(1.169)		(0.111)
u_{t-6}^R	-1.073	-	-0.337	-1.036	-	-0.337
(se)	(0.752)		(0.104)	(0.724)		(0.103)
Standard Dev. $(u_t^\cdot)$	1.241	1.970	0.828	1.263	2.084	0.828
(se)	(0.311)	(0.847)	(0.044)	(0.303)	(0.665)	(0.045)
Log L	-451.3			-451.5		
P-value				0.527		

Note: Maximum likelihood estimates of the system (3) in the text, with correction for measurement error. Sample: 1978(1)-2005(1). The P-value reported is the P-value of the likelihood ratio test on the zero restriction for the coefficient u_{t-1}^Y on the consumption equation (left panel: unrestricted estimates; right panel: restricted estimates).

Table 8 – Consumption, income and interest rate, cohort 1 and 2, parsimonious

	Non durable	Income	Interest rate	Non durable	Income	Interest rate
u_t^x	1			1		
u_t^Y (se)	0.638 (0.356)	1		0.540 (0.159)	1	
u_{t-1}^Y (se)	-0.154 (0.335)	-0.387 (0.172)			-0.316 (0.206)	
u_t^R (se)	-0.567 (0.658)	-	1	-0.551 (0.641)	-	1
u_{t-1}^R (se)	1.198 (0.753)	-	0.592 (0.090)	1.168 (0.712)	-	0.592 (0.090)
u_{t-2}^R (se)	0.064 (0.635)	-	0.303 (0.091)	0.053 (0.635)	-	0.304 (0.092)
u_{t-3}^R (se)	0.073 (0.659)	-	-0.114 (0.091)	0.100 (0.621)	-	-0.115 (0.091)
u_{t-4}^R (se)	-0.005 (0.587)	-	-0.427 (0.103)	-0.018 (0.580)	-	-0.428 (0.103)
u_{t-5}^R (se)	-0.180 (0.877)	-	-0.498 (0.105)	-0.159 (0.869)	-	-0.498 (0.105)
u_{t-6}^R (se)	-0.395 (0.703)	-	-0.368 (0.106)	-0.414 (0.689)	-	-0.368 (0.106)
u_t^{11} (se)	0.882 (0.249)	2.292 (0.402)	0.827 (0.059)	0.872 (0.244)	2.098 (0.420)	0.827 (0.044)
u_t^{21} (se)	1.400 (0.420)	1.181 (0.469)		1.375 (0.424)	1.166 (0.440)	
u_t^{22} (se)	1.006 (0.468)	1.200 (0.418)		1.013 (0.460)	1.056 (0.401)	
Log L	-855.7			-855.8		
P-value				0.655		

Note: Maximum likelihood estimates of the system (3) in the text, with correction for measurement error. Sample: 1978(1)-2005(1). The P-value reported is the P-value of the likelihood ratio test on the zero restriction for the coefficient u_{t-1}^Y on the consumption equation (left panel: unrestricted estimates; right panel: restricted estimates).

u_t^{ij} is the ij element of the Cholesky decomposition of the variance covariance matrix of the shocks.

Table 9 – Consumption, income and interest rate, aggregate data, parsimonious specification

	Non durable	(se)	Income	(se)	Interest rate	(se)
u_t^x	1	-				
u_{t-1}^x	-0.072	(0.130)				
u_{t-2}^x	0.604	(0.026)				
u_{t-3}^x	0.494	(0.032)				
u_{t-4}^x	0.084	(0.160)				
u_{t-5}^x	-0.582	(0.032)				
u_t^Y	0.027	(0.021)	1	-		
u_{t-1}^Y	0.077	(0.019)	-0.466	(0.089)		
u_{t-2}^Y	-0.172	(0.018)	0.184	(0.126)		
u_{t-3}^Y	-	-	0.030	(0.116)		
u_{t-4}^Y	-	-	-0.007	(0.126)		
u_{t-5}^Y	-	-	0.274	(0.114)		
u_t^R	0.207	(0.068)	0.211	(0.149)	1	-
u_{t-1}^R	-0.074	(0.104)	-	-	0.539	(0.079)
u_{t-2}^R	0.153	(0.083)	-	-	0.209	(0.089)
u_{t-3}^R	-0.109	(0.088)	-	-	-0.167	(0.096)
$u_{c,t-4}^R$	-0.001	(0.060)	-	-	-0.375	(0.119)
u_{t-5}^R	0.002	(0.084)	-	-	-0.454	(0.127)
u_{t-6}^R	0.017	(0.112)	-	-	-0.357	(0.124)
Standard Dev. (u_t)	0.552	(0.055)	1.110	(0.091)	0.831	(0.055)
Log L	-84.8					

Note: Maximum likelihood estimates of the system (3) in the text, without correction for measurement error.
Sample: 1978(1)-2005(1).

5. A structural interpretation of the results

Having estimated the parameters in model (3), one could discuss the implications of these estimates for alternative theoretical frameworks. Alternatively, one could start from a theoretical framework and think of the implications that it has for the parameters of model (3). We follow this latter approach.

The starting point for a structural interpretation of the parameters in system (3) 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 (3). 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 considered as deterministic. As we focus on business cycle frequencies, we do not think that this assumption is particularly strong.[12]

5.1. *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:

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

(7) $\log(\lambda_t) = \log(U_x(x_t, z_t))$

where the variable k 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 z. 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 is the interest rate. The specification in equations (6) and (7) 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 (6) and (7) 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 z. We do not report results that allow for these effects because our groups are defined in terms of year of birth and we have removed cohort and age effects. In

---

[12] 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.

a larger data set, where one could form groups based on, say, education level of the household head, one could easily generalize model (3) to allow for them.

If we assume that the interest rate and k are constant over time and that the observable component of the vector z contains only deterministic variables that can be captured by the deterministic trends removed in our first step, equations (6) and (7) have very simple and strong implications for the model in (3).

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 (6) and (7) 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'[13] 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 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. Notice that in Tables 6, 7 and 8 we do not find any evidence of 'excess sensitivity' of consumption to income, consistent with the evidence in Attanasio and Weber (1993). The picture that emerges from aggregate data in Table 9 is, of course, very different, and consistent with previous evidence such as that presented, for instance, by Campbell and Mankiw (1991).[14]

The approach followed so far is quite similar to that of Altonji et al. (2002), with an important difference: the fact that we do not remove time effects.[15] A first and very simple generalization of the model which stresses the differences between our methodology and that of Altonji et al. is to consider time varying interest rates. This extension could be of particular interest as allows one to estimate the elasticity of intertemporal substitution.

Allowing for a time variable interest rate involves adding an additional equation to the model, so that one can measure the correlation between innovations to interest rates and consumption (and

---

[13] Or martingale, to be precise.

[14] Hall and Mishkin (1982) using the US micro data PSID find that consumption changes are related to lagged levels of income, contradicting the prediction of the permanent income model.

[15] Altonji et al. (2002) follow two different strategies that could be, in principle be pursued here. The first consists in parametrizing the innovation to marginal utility of wealth as a function of the innovations of wages, non labor income and possibly other variables deemed to be relevant for the problem. The other is to use explicitly the Euler equation (8) to difference out λ_t.

other variables). 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.

If we consider an asset whose rate of return is the same across groups and that is widely held, than equations (6) and (7) induce a set of additional restrictions on system (3). Let's define

$$A^{xx}(L) = \sum_{j=1}^{q} \alpha_j^{xx} L^j \,,$$ and analogously for $A^{xr}(L)$ and $A^{rr}(L)$. It is easy to show that an

isoelastic utility function with a coefficient of relative risk aversion γ implies that:

(8) $$\gamma(A^{xx}(L)u_{c,t+1}^{x} + A^{rx}(L)u_{c,t+1}^{r}) = A^{rr}(L)u_{c,t+1}^{r}.$$

As this has to hold for every possible realization of the residuals, the restrictions on the coefficients of system (3) are that:

(9) $$A^{xx}(L) = 0 \text{ and}$$

(10) $$\gamma A^{xr}(L) = A^{rr}(L).$$

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 z. 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 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.

In Table 10, we report the estimated elasticity of intertemporal substitution implied by the restrictions in equation (10) and by the estimates of the coefficients we have fitted both on cohort and aggregate data.

As the first model we estimated does not constrain the coefficients of the system (3) to be the same across cohorts, they imply different estimates of the elasticity of intertemporal substitution. For the first cohort, we find that the elasticity is estimated at 1.03 and is statistically different from zero. For the second cohort, we find an elasticity equal to 0.85, although imprecisely estimated. When we constrain the coefficients for the two cohorts to be the same, the estimates we obtain imply an elasticity equal to 0.95 and statistically different from zero. These estimates are consistent with those reported by Attanasio and Weber (1993) and other studies that used the FES. Given the evidence in Table 6, 7 and 8, it is clear that the restrictions of the same coefficients for the two cohorts is not rejected. The estimate for aggregate data is negative and not statistically different from zero.

It also worth stressing the fact that for neither of the two cohorts we reject the null that lagged income shocks determine consumption growth. This implies that standard 'excess sensitivity' tests do not reject the null in micro data. As mentioned above, the same test on aggregate data does reject the null of no excess sensitivity. These differences between micro and aggregate data results is in line with Attanasio and Weber (1993), who show the importance of correct aggregation in estimating Euler equations for consumption.

Table 10 – Elasticity of intertemporal substitution

	Cohort 1	Cohort 2	Cohort 1+2	Aggregate Data
EIS	1.03	0.849	0.946	-0.030
(se)	*(0.473)*	*(0.576)*	(0.409)	*(0.125)*
Log L	-442.30	-454.80	*-858.23*	-109.14
P-value (a)	0.39	0.40	0.43	0.00

Note: estimates of the elasticity of intertemporal substitution obtained estimating by maximum likelihood the restricted versions of the models in tables 6, 7, 8 and 9: the restrictions imposed are given by equation (10) in the text.
(a) P-value of the likelihood ratio test of the restriction imposed by (10) on the models in table 6, 7 and 8 (right panel) and 9, respectively for cohort 1, cohort 2, cohort 1 and 2, and aggregate data.

5.2. *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 (6) and (7). 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. In situations in which 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 (3) 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 (2005) and Attanasio and Pavoni (2006).

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:

$$\Delta y_t = w_t^y + \beta_{111} w_{t-1}^y + \beta_{120} w_t^c + \beta_{121} w_{t-1}^c$$
$$\Delta c_t = \qquad\qquad w_t^c + \beta_{221} w_{t-1}^c$$

$$E[w_t^y] = E[w_t^y w_{t-1}^y] = E[w_t^c] = E[w_t^c w_{t-1}^c] E[w_t^y w_t^c] = 0$$
$$Var(w_t^y) = \sigma_{wy}^2; \ Var(w_t^c) = \sigma_{wc}^2$$

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 (3) 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 (2006) 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 (2006) 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:[16]

$$(11) \qquad \alpha_0^{xy} = \alpha^{yy}(z) = \sum_{j=0}^{q} \alpha_j^{yy} z^j$$

where $z = (1/1+r)$ and r is the interest rate. Campbell (1987), West (1988) and Campbell and Deaton (1989), 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.

Using the results we obtained estimating various versions of system (3), and following Attanasio and Pavoni (2006), we can test the restriction in equation (11) against the alternative of excess smoothness. In particular we test the hypothesis that $\alpha_0^{xy} / \alpha^{yy}(z) = 1$ against the alternative that the same ratio is less than 1. We evaluate the term $\alpha^{yy}(z)$ at a quarterly interest rate equal to 1 per cent. We report these results in Table 11. Of particular interest are the results derived from

---

[16] In our model, 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

the estimates in the right panel in table 6 and 7 (cohort 1 and 2 respectively) and 8 (the two cohorts together) 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 0.75 with an estimated standard error of 0.20. For the second (and younger) cohort the excess smoothness test is equal to 0.61, with an estimated standard error equal to 0.45. When estimating the two cohorts together, the excess sensitivity test is equal to 0.91, with a standard error greater than 1. In all cases, therefore, we fail to reject the hypothesis of no excess smoothness.

Turning to the estimates based on aggregate data, we find a rejection of the hypothesis that the excess smoothness parameters equal 1. These results are consistent with those presented in Campbell (1987) and Campbell and Deaton (1989).

Table 11 – Excess smoothness

Model	$\alpha_0^{xy} / \alpha^{yy}(z)$	S.E.
Cohort Data		
Cohort 1 (Table 6, right panel)	0.745	*(0.201)*
Cohort 2 (Table 7, right panel)	0.613	*(0.446)*
Cohort 1 and 2 (Table 8, right panel)	0.914	*(1.043)*
Aggregate Data		
(Table 9)	0.027	*(0.022)*

Note: Estimates of excess smoothness obtained from the maximum likelihood estimates of the models in table 6, 7, 8 and 9. The discounted sum of the income coefficients has been computed at a quarterly interest rate equal to 1%.

equation for consumption the intertemporal budget constraint does not impose restrictions on the time series properties of savings and consumption.

6.   Extensions

The simple versions of the model considered above can be extended in a variety of ways. While we leave the empirical investigations of these extensions to future research, here we sketch the implications of some generalizations of the simple model.

6.1.   *Non separability with labour supply and other components of consumption*

Implicit in the formulation of the model above is the assumption that non durable consumption is separable from other components of consumption excluded from the analysis (such as durables and housing) as well as from leisure. The latter might be particularly important as deviations from this assumption could explain the observed correlation between expected income and consumption. Indeed, in many empirical analysis of Euler equations based on micro data (such as those of Attanasio and Weber (1993), Blundell, Browning and Meghir (1994), Banks, Blundell and Preston (1994), Attanasio and Browning (1995), Attanasio and Weber (1995)) labour supply, and in particular, female labour force participation seems to play an important role.[17] This evidence is not entirely surprising, as many components of consumption expenditure are accounted by job related expenses or, in the case of female labour force participation, might substitute for home production services.

The generalization of the simple model proposed in the previous section is straightforward: equations (6) and (7) still hold, except that the marginal utility of consumption has to depend on the excluded commodities and on labour supply. It should be stressed that equations (6) and (7) are robust to the presence of various kinds of complications in the determination of durables and/or labour supply, such as fixed adjustment costs and the like. The marginal utility of non durable consumption is defined as a function of non durable consumption and the optimal level of the other relevant variables, regardless of how they are determined.

The difficulty, in the case of durable consumption, concerns the observability of the existing stock of durables at each point in time.[18] In the case of labour supply and in particular female labour supply, for which corner solutions are important, one has to allow the marginal utility depend explicitly on participation at the individual level. Given the nature of the data, this does not constitute an important problem. If the (log of) marginal utility of consumption depends additively on a participation indicator, aggregating equation (7) one has a model in which average

---

[17] Browning and Meghir (1991) test explicitly for the dependence of a demand system on labour supply behaviour.
[18] One might try to construct group level estimates of the existing stock of durables by cumulating the observed expenditures. We have not yet attempted this procedure.

changes in log non durable consumption depend, among other things, on the (changes) in participation rates at a point in time for a given cohort. It is therefore necessary to model female participation rates in a way analogous to the way in which we model non durable consumption, wages, or income. Such an equation can be easily added to the system of equations (3). It should be stressed that this procedure, while allowing the study of the properties of non durable consumption is silent about the determinants of labour force participation.

6.2. Multiple commodities

So far we have worked with the assumption of a single and homogeneous non durable good. More precisely, we have considered total non durable consumption and studied its allocation over time as a function of a single price index. Of course, this approach is only justified under stringent conditions on preferences (see Gorman, 1953).[19] It might therefore be important to model simultaneously the allocation of resources over time and, at a point in time, among several commodities. Furthermore, even when the Gorman aggregation conditions are satisfied, the study of a demand system can be of interest. Finally, the consideration of several Euler equations simultaneously might give more powerful tests of the model considered.

Let's then assume that $\mathbf{q}$ is a vector of m commodities, with prices $\mathbf{p}$. Instead of equation (7) we will then have m equations relating the marginal utility of each commodity to its price and to the marginal utility of wealth λ :

(7') $\log(p_t^i) + \log(\lambda_t) = \log(U_{q^i}(q_t^i, z_t)); \quad i = 1, \ldots, m$

where the index i refers to the commodity. As we are writing (7') as an equality, we are implicitly ruling out the possibility of corners in any of the m commodities. From equation (7') it is also clear why it is important to have an unobserved component in preferences. Without it, one could consider equation (7') for two different commodities to eliminate the marginal utility of wealth and obtain an equation that has no error!

From our perspective, to map a system of equations such as (7') in anything like (3) involves considering the innovations in each commodity (and possibly in total consumption) and modelling the vector of relative prices. As far as the latter are concerned, they can be treated in a

---

[19] Blundell, Browning and Meghir (1994) and Attanasio and Weber (1995) address this issue. In the first paper, that uses UK data, the authors find that while the restrictions that would grant the use of a single price index are formally rejected, the use of a Stone price index constitute a good approximation of the 'true' price index that should be used.

fashion similar to the interest rate: they can be assumed to be constant across consumers. Three different possibilities are open for the treatment of commodities. On the one hand one can consider m-1 equations obtained by using a specific commodity as a benchmark and eliminate therefore the marginal utility of wealth from the system. This approach involves therefore to consider static relationships, possibly expressed in ratios of marginal utilities and relative prices. Second, one can use a function of total non durable expenditure as an approximation of the marginal utility of wealth and use it in each of the m equations. The last alternative is to use equation (6) in each of the expressions in (7') and therefore derive an Euler equation for each of the commodities considered.

Several considerations are in order. First, regardless of the approach used, the discussion above about the possibility that the vector z includes some choice variables is relevant here. The demand system that one obtains eliminating the marginal utility of wealth is effectively the *conditional* demand system discussed in detail by Browning and Meghir (1991).

Second, the first two approaches are essentially static and can be expressed in terms of the levels of the variables of interest. The residuals of these equations arise from unobserved hetereogeneity across consumers (groups) and measurement error. Indeed, the first two approches give rise to equations that can in principle be estimated using cross sectional data (except that one has to have enough price variability, that can only be observed over time). On the other hand, these equations, within the framework of a life cycle model with intertemporally separable preferences, are uninformative about the way in which households react to shocks.[20] Third, the last approach, of deriving m Euler equations, is intrinsically dynamic and is the natural extension of what we do considering a single commodity. This is the preferable line as it delivers some interesting restrictions of the system of equations (7') (extended for changes in prices). In particular, one can see that the innovations to that system of equations, once one controls for changes in prices, should be driven by a single factor: the innovations to the marginal utility of wealth.

6.3.   *Absorption of shocks by different commodities*

As mentioned above, the existence of complete contingent markets implies strong restrictions to the variance covariance structure of the residuals of system (3). In particular, the

Attanasio and Weber (1995), using a slightly different parametrization of preferences and US data, find a more important role for a second price index.

[20] Meghir and Weber (1996) interpret any evidence of dynamics in a system of demand equations as an indication of intertemporal non-separability. Using US data they are unable to identify any dynamic effect, once they condition on durable and semi-durable consumption.

contemporaneous innovations to the marginal utility of wealth should be perfectly correlated across groups. This is the idea exploited in a number of empirical applications and tests of the perfect insurance hypothesis, such as Hayashi, Altonji and Kotlikoff (1996) and Attanasio and Davis (1996). While Hayashi et al. focus on individual US data and reject strongly the hypothesis of risk sharing, even among related families, Attanasio and Davis use data grouped by education and cohort and reject the hypothesis of complete markets only when they consider relatively low frequency changes. This last result might be either an indication that high frequency are effectively insured against or that they are effectively self insured, so that Attanasio and Davis' test lack the necessarily power to reject the null.

Regardless of the interpretation of the result, if group shocks to wages are not reflected in consumption at high frequency, it might be interesting to study whether other variables (such as durable expenditure, labour supply) absorb such shocks or whether smoothing mechanisms such as income support or borrowing are capable of absorbing completely transitory shocks. A similar question has recently been asked by Gruber (1997) and Dynarski and Gruber (1997) by considering how consumption, several of its components and other variables vary with shocks to income.

Our framework is suitable to analyze the correlation between several components of consumption (and labour supply) with innovations to wages or other variables. Deviations from the complete market benchmark, or even from the self-insurance paradigm implied by the life cycle model might be reflected in particular pattern of correlations in systems analogous to (7'). Browning and Crossley (2003), for instance, have suggested that Canadian unemployed households might react to income shocks by delaying the replacement of small durables. They identify these effects by estimating a small demand system and verifying that income support variables have an effect on the expenditure on small durables over and above that accounted for by total expenditure in any particular period. Our approach suggests to consider the variability of various components of consumption for different groups and to measure the correlation between the innovations to the components that are most variable and that to income or wages. However, as the estimation of such a system is particularly cumbersome, we leave this topic for future research.

The problems discussed in this section would be best addressed by considering finer groups than those defined by year of birth cohorts. Attanasio and Davis (1996), for instance, find that most of the variability in relative wages is explained by variations across education groups (rather than across cohorts). Furthermore, on a priori basis, some groups are more likely to be able to smooth

shocks by borrowing or by decumulating assets, while alternative smoothing mechanism might be used by those groups that do not have access to them. Attanasio (1997), for instance, has found that in the US, not only that the expenditure of households headed by individual with low education is relatively more volatile, but that most of the difference is accounted for by differences in the volatility of durable expenditure.

7.   Conclusions

In this paper we have analyzed the time series properties of individual 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 explicitly take into account the measurement error present in the individual measures of consumption and income. Data are drawn from the UK Family Expenditure Survey.

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 et al. (1994) and Attanasio and Browning (1995). The results are not induced by the method we use: when using aggregate data we obtain that lagged income shocks do predict consumption changes, as in Campbell and Mankiw (1990). When we use the estimated parameters to identify the elasticity of intertemporal substitution using micro data, we estimates in the range of 0.9, which is consistent with the estimates reported by Attanasio and Weber (1993). Notice that our identification of persistence in consumption growth induced by consumption shocks of the type we identified might be explained by taste shocks or movements in second moments.

Our approach can be also used to assess the extent to which innovation 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 not find evidence of 'excess smoothness' of consumption. That is, we find that consumption reacts to permanent innovations to income in a way that is consistent with the life cycle model. This contrasts with the findings based on aggregate data for which we find (like Campbell and Deaton, 1989) evidence of excess smoothness. Our evidence

contrasts, instead with that in Attanasio and Pavoni (2008) 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 (2008) use a different normalization and do not model variation in the interest rate. They also use a slightly different time period.

In the last section of the paper, we have proposed a number of extensions to the simple theoretical framework we presented in Section 5. These features would lead to the estimation of a model richer than the one presented in Section 4. Such a model could include the consideration of several components of consumption as well as additional variables, such as labour supply. We have left this empirical work for future research.

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 section 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 the dynamics of our data. 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.

References

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

Altonji, J.G., A.P. Martins and A. Siow (2002), "Dynamic Factors Models of Consumption, Hours and Income", *Research in Economics*, 56, 3-59.

Attanasio, O.P. (1997) "Consumption and saving behaviour: modelling recent trends", *Fiscal Studies,* 18(1), 23-47.

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

Attanasio, O.P. and S.J. Davis (1996) "Relative Wage Movements and the Distribution of Consumption", *Journal of Political Economy*, 104(6), 1227-62.

Attanasio, O.P. and N. Pavoni (2006): "Risk Sharing in Private Information Models with Asset Accumulation: Explaining the Excess Smoothness of Consumption", UCL Mimeo.

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

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

Banks, J., R. Blundell and I. Preston (1994), "Life-Cycle Expenditure Allocations and the Consumption Cost of Children", *European Economic Review*, 38(7), 1391-1410.

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

Blundell, R., L. Pistaferri and I. Preston (2004) "Consumption inequality and partial insurance", IFS Working Papers W04/28, Institute for Fiscal Studies.

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

Blundell, R. and I. Preston (1998), "Consumption Inequality and Income Uncertainty", *Quarterly Journal of Economics*, 113, 603-40.

Brewer, M., A. Goodman, J. Shaw and L. Sibieta (2006), *Poverty and Inequality in Britain: 2006*, IFS Commentary 101 (http://www.ifs.org.uk/comms/comm101.pdf), London: IFS

Browning, M. and C. Meghir (1991), "The Effect of Male and Female Labour Supply on Commodity Demands", *Econometrica*, 53, 503-44.

Browning, M. and T. Crossley (2003) "Shocks, Stocks and Socks", Department of Economics Working Papers 2003-07, McMaster University.

Caballero, R.J. (1990), "Expenditure on Durable Goods: A Case for Slow Adjustment", *Quarterly Journal of Economics*, 105, 727-43.

Campbell, J.Y. (1987) "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis", *Econometrica*, 55(6), 1249-73.

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

Campbell, J. Y. and A.S. Deaton (1989): "Why Is Consumption So Smooth?", *Review of Economic Studies* 56, pp. 357–374.

Campbell, J.Y. and N.G. Mankiw (1991) : ""The Response of Consumption to Income: A Cross-Country Investigation", *European Economic Review* 35, pp. 723–767.

Cunha, F., Heckman, J., and Navarro, S. (2004). "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. (2005). "Separating Uncertainty from Heterogeneity in Life Cycle Earnings," NBER Working Papers 11024, National Bureau of Economic Research, Inc.

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

Deaton, A. and C. Paxon (1994), "Intertemporal choice and inequality", *Journal of Political Economy*, 102, 437-467.

Dynarski, S. and J. Gruber (1997) "Can Families Smooth Variable Earnings?", *Brookings Papers on Economic Activity*, 0(1), 229-84.

Flavin, M.A. (1981) "The Adjustment of Consumption to Changing Expectations about Future Income", *Journal of Political Economy*, 89(5), 974-1009.

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

Gorman, W.M. (1959), "Separable Utility and Aggregation", *Econometrica*, 27, 469-481.

Gruber, J. (1997) "The Consumption Smoothing Benefits of Unemployment Insurance" *American Economic Review*, 87(1), 192-205.

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

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

Hayashi, F., J. Altonji and L. Kotlikoff (1996) "Risk-Sharing between and within Families", *Econometrica*, 64(2), 261-94.

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

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

Meghir, C. and G. Weber (1996), "Intertemporal Non-Separability or Liquidity Constraints? A Disaggregate Analysis on US Panel Data", Econometrica, 64(5), 1151-81 .

Moffitt, R.A. and P. Gottschalk (1995), "Trends in the Covariance Structure of Earnings in the U.S.: 1969-1987", mimeo, Johns Hopkins University.
.
Quah, D. (1990), "Permanent and Transitory Movements in Labor Income: An Explanation for "Excess Smoothness" in Consumption", *Journal of Political Economy*. 98, pp. 449-475

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

Sargent, T.J. (1978) "Rational Expectations, Econometric Exogeneity and Consumption", *Journal of Political Economy*, 86(4), 673-700.

Tanner, S. (1998), 'How Much Do Consumers Spend? Comparing the FES and National Accounts' in Banks, J. and P. Johnson (eds), *How Reliable is the Family Expenditure Survey? Trends in Incomes and Expenditures Over Time*, IFS Report 57 (http://www.ifs.org.uk/comms/r57.pdf), London: IFS.

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

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

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

Appendix

A1. The MA structure

For ease of notation, we consider a multivariate MA(1) model in three variables: the first two are cohort aggregates from individual data, and therefore need to be corrected for measurement error, while the third is an aggregate variable (e.g. we consider the first difference of log consumption, of log disposable income, and the interest rate).

The system for each cohort (c=1, 2) is then:

$$
\begin{aligned}
\Delta x_t^c &= u_t^{xc} + \alpha_1^{xx-c} u_{t-1}^{xc} + \alpha_0^{xy-c} u_t^{yc} + \alpha_1^{xy-c} u_{t-1}^{yc} + \alpha_0^{xr-c} u_t^r + \alpha_1^{xr-c} u_{t-1}^r + \eta_t^{xc} - \eta_{t-1}^{xc} \\
\Delta y_t^c &= \hphantom{u_t^{xc} + \alpha_1^{xx-c} u_{t-1}^{xc} + \alpha_0^{xy-c}} u_t^{yc} + \alpha_1^{yy-c} u_{t-1}^{yc} + \alpha_0^{yr-c} u_t^r + \alpha_1^{yr-c} u_{t-1}^r + \eta_t^{yc} - \eta_{t-1}^{yc} \\
r_t &= \hphantom{u_t^{xc} + \alpha_1^{xx-c} u_{t-1}^{xc} + \alpha_0^{xy-c} u_t^{yc}} u_t^r + \alpha_1^{rr} u_{t-1}^r
\end{aligned}
\tag{A1}
$$

The shock vector $\upsilon_t = \begin{bmatrix} u_t^{xc} & u_t^{yc} & u_t^r & \eta_t^{xc} & \eta_t^{yc} \end{bmatrix}'$ has mean zero and variance-covariance matrix:

$$
\begin{bmatrix}
\sigma_{ux,c}^2 & 0 & 0 & 0 & 0 \\
0 & \sigma_{uy,c}^2 & 0 & 0 & 0 \\
0 & 0 & \sigma_{ur}^2 & 0 & 0 \\
\hline
0 & 0 & 0 & \sigma_{\eta x,c}^2 & \sigma_{\eta xy,c}^2 \\
0 & 0 & 0 & \sigma_{\eta xy,c}^2 & \sigma_{\eta y,c}^2
\end{bmatrix}
$$

When estimating the elasticity of intertemporal substitution (γ) we constrain the coefficient(s) of the lagged interest rate shock in the consumption equation to be proportional to the coefficient(s) of lagged interest rate shock in its own equation:

$$
\begin{aligned}
\Delta x_t^c &= u_t^{xc} + \alpha_1^{xx-c} u_{t-1}^{xc} + \alpha_0^{xy-c} u_t^{yc} + \alpha_1^{xy-c} u_{t-1}^{yc} + \alpha_0^{xr-c} u_t^r + \gamma^c \alpha_1^{rr} u_{t-1}^r + \eta_t^{xc} - \eta_{t-1}^{xc} \\
\Delta y_t^c &= \hphantom{u_t^{xc} + \alpha_1^{xx-c} u_{t-1}^{xc} + \alpha_0^{xy-c}} u_t^{yc} + \alpha_1^{yy-c} u_{t-1}^{yc} + \alpha_0^{yr-c} u_t^r + \alpha_1^{yr-c} u_{t-1}^r + \eta_t^{yc} - \eta_{t-1}^{yc} \\
r_t &= \hphantom{u_t^{xc} + \alpha_1^{xx-c} u_{t-1}^{xc} + \alpha_0^{xy-c} u_t^{yc}} u_t^r + \alpha_1^{rr} u_{t-1}^r
\end{aligned}
\tag{A2}
$$

When estimating the restricted model for the two cohorts (1 and 2) the system becomes:

$$\Delta x_t^1 = u_t^{x1} + \alpha_1^{xx} u_{t-1}^{x1} + \alpha_0^{xy} u_t^{y1} + \alpha_1^{xy} u_{t-1}^{y1} + \alpha_0^{xr} u_t^r + \alpha_1^{xr} u_{t-1}^r + \eta_t^{x1} - \eta_{t-1}^{x1}$$

$$\Delta x_t^2 = u_t^{x2} + \alpha_1^{xx} u_{t-1}^{x2} + \alpha_0^{xy} u_t^{y2} + \alpha_1^{xy} u_{t-1}^{y2} + \alpha_0^{xr} u_t^r + \alpha_1^{xr} u_{t-1}^r + \eta_t^{x2} - \eta_{t-1}^{x2}$$

$$\Delta y_t^1 = \qquad\qquad u_t^{y1} + \alpha_1^{yy} u_{t-1}^{y1} + \alpha_0^{yr} u_t^r + \alpha_1^{yr} u_{t-1}^r + \eta_t^{y1} - \eta_{t-1}^{y1} \qquad\qquad\text{(A3)}$$

$$\Delta y_t^2 = \qquad\qquad u_t^{y2} + \alpha_1^{yy} u_{t-1}^{y2} + \alpha_0^{yr} u_t^r + \alpha_1^{yr} u_{t-1}^r + \eta_t^{y2} - \eta_{t-1}^{y2}$$

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

and the zero-mean shock vector :

$$\upsilon_t = \begin{bmatrix} u_t^{x1} & u_t^{x2} & u_t^{y1} & u_t^{y2} & u_t^r & \eta_t^{x1} & \eta_t^{x2} & \eta_t^{y1} & \eta_t^{y2} \end{bmatrix}'$$

has variance-covariance matrix:

$$
\left[
\begin{array}{ccccc|cccc}
\sigma_{ux1}^2 & \sigma_{ux1x2}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\sigma_{ux1x2}^2 & \sigma_{ux2}^2 & 0 & 0 & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \sigma_{uy1}^2 & \sigma_{uy1y2}^2 & & & & & \\
\vdots & 0 & \sigma_{uy1y2}^2 & \sigma_{uy2}^2 & 0 & & & & \\
 & \vdots & 0 & 0 & \sigma_{ur}^2 & 0 & 0 & 0 & 0 \\
\hline
 & \vdots & 0 & 0 & \sigma_{\eta x1}^2 & 0 & \sigma_{\eta x1y1}^2 & 0 \\
 & & \vdots & 0 & 0 & 0 & \sigma_{\eta x2}^2 & 0 & \sigma_{\eta x2y2}^2 \\
 & & & 0 & \sigma_{\eta x1y1}^2 & 0 & \sigma_{\eta y1}^2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \sigma_{\eta x2y2}^2 & 0 & \sigma_{\eta y2}^2
\end{array}
\right]
$$

A2. The state space representation of the model

Here we show how we write the state-space representation of model (A1). To estimate the system without correction for measurement error – as it is done in the case of aggregate data for consumption and disposable income – or to estimate a single equation MA, the model is simplified in an obvious way.

Recall the system for each cohort may be written as:

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

$$\Delta y_t = \qquad\qquad u_t^y + \alpha_1^{yy} u_{t-1}^y + \alpha_0^{yr} u_t^r + \alpha_1^{yr} u_{t-1}^r + \eta_t^y - \eta_{t-1}^y$$

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

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

Define:

$M = (n+z)$ i.e. the number of dependent variables in the model (5 in the example);

k the dimension of the space vector;

$k_1 = M+n$ 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:

$$\xi_{t+1} = F\xi_t + v_t$$
$$X_t = H\xi_t$$

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 of v is a diagonal matrix apart from the entries in which there is the covariance among the measurement error terms η. The measurement equation has no noise, so its variance-covariance matrix, R, is equal to zero. All these matrices are described in the last section.

The log likelihood function of the model is given by:

$$\log L = -\frac{NT}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log|G_t| - \frac{1}{2}\sum_{t=1}^{T}\varepsilon_t{}'G_t^{-1}\varepsilon_t$$

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

$$\varepsilon_t = X_t - H\hat{\xi}_{t/t-1}$$

with associated MSE:

$$G_t = HP_{t/t-1}H' + R$$

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:

$$\hat{\xi}_{t/t-1} = F\hat{\xi}_t$$

$$P_{t/t-1} = FP_{t-1}F' + Q$$

$$\hat{\xi}_t = \hat{\xi}_{t/t-1} + P_{t/t-1}H'(HP_{t/t-1}H' + R)^{-1}(X_t - H\hat{\xi}_{t/t-1})$$

$$P_t = P_{t/t-1} - P_{t/t-1}H'(HP_{t/t-1}H' + R)^{-1}HP_{t/t-1}$$

Definition of the matrices

Matrix *H*, in the measurement equation, is given by:

$$\begin{bmatrix} 1 & \alpha_0^{xy} & \alpha_0^{xr} & 1 & 0 & \alpha_1^{xx} & \alpha_1^{xy} & \alpha_1^{xr} & -1 & 0 \\ 0 & 1 & \alpha_0^{xr} & 0 & 1 & 0 & \alpha_1^{yy} & \alpha_1^{yr} & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & \alpha_1^{rr} & 0 & 0 \end{bmatrix}$$

The matrix F, in the transition equation, is given by:

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}$$

where the three zero blocks are of dimension ($k_1 \times k_1$), and *I* is an identity matrix of dimension k_1.

The error term in the transition equation is:

$$\upsilon_{t+1} = \begin{bmatrix} u_{t+1}^x & u_{t+1}^y & u_{t+1}^r & \eta_{t+1}^x & \eta_{t+1}^y & \mathbf{0} \end{bmatrix}'$$

where the zero block is a vector of dimension $k-k_1$.

The ($k_1 \times k_1$) upper-left block of the variance-covariance matrix of **v**, *Q*, is:

$$\begin{bmatrix} \sigma_{ux}^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{uy}^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{ur}^2 & 0 & 0 \\ \hline 0 & 0 & 0 & \sigma_{\eta x}^2 & \sigma_{\eta xy}^2 \\ 0 & 0 & 0 & \sigma_{\eta xy}^2 & \sigma_{\eta y}^2 \end{bmatrix}$$

while the other three ($k-k_1 \times k-k_1$) blocks are zero.

The variances of the measurement error terms are treated as known in the estimation.