Dynamic Econometric Models for Cohort and Panel Data: Methods and Applications to Life-Cycle Consumption

by

Maria Dolores Collado-Vindel

Department of Economics

The London School of Economics and Political Science

July 1994

Thesis submitted to the University of London for the Degree of Doctor of Philosophy

ABSTRACT

The purpose of this research is to analyze dynamic models for cohort and panel data, with special emphasis in the applications to life-cycle consumption.

In the second chapter of the thesis we analyze the estimation of dynamic models from time-series of independent cross-sections. The population is divided in groups with fixed membership (cohorts) and the cohort sample means are used as a panel subject to measurement errors. We propose measurement error corrected estimators and we analyze their asymptotic properties. We also calculate the asymptotic biases of the non-corrected estimators to check up to what extent the measurement error correction is needed. Finally, we carry out Monte Carlo simulations to get an idea of the performance of our estimators in finite samples.

The purpose of the second part is to test the life-cycle permanent income hypothesis using an unbalanced panel from the Spanish family expenditure survey. The model accounts for aggregate shocks and within period non-separability in the Euler equation among consumption goods, contrary to most of the literature in this area. The results do not indicate excess sensitivity of consumption growth to income.

In the last chapter, we specify a system of nonlinear intertemporal (or Frisch) demands. Our choice of specification is based on seven criteria for such systems. These criteria are in terms of consistency with the theory, flexibility and econometric tractability. Our specification allows us to estimate a system of exact Euler equations in contrast to the usual practice in the literature. We then estimate the system on Spanish panel data. This is the first time that a Frisch demand system has been estimated on panel data. We do not reject any of the restrictions derived from theory. Our results suggest strongly that the intertemporal substitution elasticity is well determined.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I am very grateful to my supervisor Manuel Arellano. He has shared with me many of the ideas behind this thesis, and his suggestions and comments have contributed to improve this research.

I wish to thank Víctor Aguirregabiria, César Alonso, Orazio Attanasio, Olympia Bover, Martin Browning and Jorge Padilla for helpful comments.

I also acknowledge the financial support from the Bank of Spain, and the research facilities provides by the Centre for Economic Performance at the LSE, and CEMFI (Madrid).

A mis padres

CHAPTER 1. INTRODUCTION

Economic theory can answer many questions related to the economic behaviour of agents. However, in most cases, economic theory does not provide quantitative answers. For example, the life-cycle permanent income hypothesis explains why people do not just expend their current income. According to the life-cycle model of consumption, consumers take into account their expectations about their future income, and substitute consumption over time so as to keep their marginal utility of consumption constant (see Hall (1978)). However, the theory does not say anything about how "willing" are households to substitute consumption over time, or whether the willingness to substitute depends on the particular characteristics of the family. If we want to give an answer to these questions, we have to go further and use the theory to derive an empirical model that can be estimated using the data available. On the basis of the estimated model we can obtain the elasticity of intertemporal substitution, which is a measure of the willingness to substitute consumption over time.

The purpose of econometrics is not only to provide quantitative answers to economic problems, but also to test the theory and to discriminate among alternative theories using the empirical evidence. Following the example above, we could ask ourselves whether people actually behave according to the life-cycle model of consumption, or whether, due to imperfections in the credit markets, consumers cannot borrow as much as they would like to. It is possible to establish a theoretical model that explains the behaviour of consumers in the presence of borrowing constraints (see Zeldes (1989)). However, if we want to see up to what extent borrowing constraints affect consumer's behaviour, we have to use econometric techniques to discriminate among the alternative models.

If we want to give coherent qualitative answers to economic problems, it is crucial that our empirical model is based on the theory. Furthermore, we need to use the appropriate data set and the appropriate econometric techniques for estimation and testing. For example, it would be very difficult to estimate a dynamic model of individual behaviour, in the absence of observations for several periods of time. Under certain assumptions, some of these models can be estimated using aggregate time series data. However, as it has been argued in the literature (see Deaton (1992)), these assumptions are sometimes quite unrealistic, and the results based on macro data might not be very reliable. On the other hand, measurement errors at the micro level are often a very serious problem. On balance, it seems to be more appropriate to use a data set that contains individual observations for several periods of time, i.e, a micro panel, when the objective is to test models of individual behaviour.

There are many authors that have contributed to develop econometric techniques to handle panel data (see Chamberlain (1984) and Hsiao (1986) for surveys of the literature), and those techniques have been widely used in applied work. One of the main advantages of panel data is that we can control for unobservable individual effects that are correlated with the explanatory variables. This sort of models arise from economic theory. For example, in a certain class of life-cycle models, the individual effects represent the marginal utility of wealth (see MaCurdy (1981), Browning, Deaton and Irish (1985)).

At the firm level, there are good data sets for several countries, and these data sets have been used to estimate models of investment, labour demand, etc. However, for many countries there is no panel data on households. For example, in the U.K., there is no panel data on household consumption and labour supply, and even for the US, the PSID contains information on food consumption, labour supply, and family characteristics, but it does not provide information on expenditures in other goods. The lack of panel data on households for many countries was the main reason why most of the empirical work on the life-cycle model of consumption during the eighties was referred to the US, and was based on the PSID (Hall and Mishkin (1982), Zeldes (1989), Runkle (1991), etc). Therefore, there is very little evidence based on panel data for other countries (Hayashi (1985), Deaton (1991)).

The pioneer work of Deaton (1985) opens an alternative possibility to estimate models of individual behaviour using micro data. If we have time series of independent cross-sections, that is, if we observe independent samples of individuals for different periods of time, we can divide the population in groups (cohorts) so that each group contains the same individuals over time. Then, we can calculate the sample means for each group on each time period, and we can use the sample means as a panel subject to measurement errors.

The cohort population means have a genuine panel structure given that at the population level the groups contain the same individuals over time. However, the sample means are only consistent estimators of the true cohort population means, and therefore, when we work with the sample means we will have a measurement error problem. The advantage in this case is that we can estimate the variance of the measurement errors using the survey data. Then, we can use these estimates to correct the classical estimators for panel data. Deaton (1985) proposes a measurement error corrected within groups estimator for the static model with individual effects, which is consistent for a fixed number of observations per cohort. He analyzes the asymptotic properties of this estimator. Verbeek and Nijman (1993) modify Deaton's estimator to achieve consistency for a fixed number of time periods and a fixed number of individuals per cohort.

It is obvious that the larger the number of observations per cohort the less severe the measurement error problem will be. However, in practice, the cross-section dimension of our data set will be finite and therefore, a large number of observations per group will imply a small number of groups[1]. In applied studies with cohort data (see Browning Deaton and Irish (1985), Attanasio and Weber (1993), Blundell Browning and Meghir (1994)), the population is normally divided in a small number of groups with quite a large number of observations in each, and the

[1] The cross-section sizes of the most widely used data set are around 2000-6000 observations. If the we want to have groups of about 200 observations we will only have 10-30 groups.

sample means are treated as a genuine panel. Verbeek and Nijman (1992) study under which conditions this approach would be valid. They also consider the impact of the cohort sizes on the bias of the classical within groups estimator for the static model with individual effects.

In Chapter 2, we deal with the estimation of a dynamic model using time series of cross-sections. We propose a generalized method of moments (GMM) estimator corrected for measurement errors. This estimator is consistent as the number of cohorts tends to infinity, for a fixed number of time periods and a fixed number of individuals per cohort. We derive the asymptotic distribution of our estimator. We also consider a measurement error corrected within groups estimator (WG), which is consistent as the number of time periods tends to infinity. Moffitt (1993) also analyzes the estimation of a dynamic model from time series of independent cross-sections. However, his approach is completely different, and the estimator that he proposes is only consistent if the number of observations per cohort tends to infinity.

In the second part of this chapter, we obtain the asymptotic biases of the non-corrected GMM and WG estimators, and we analyze the size of the biases for different values of the parameters of the model. In the last section we carry out Monte Carlo simulations, and we discuss the performance in finite samples of the estimators proposed.

In the last two chapters of the thesis, we study two different questions related to the Euler equations for consumption. In Chapter 3, we consider the problem of within period separability. As we mention above, most of the empirical research on the life-cycle model of consumption using panel data is based on the PSID. This data set only contains information on food consumption. Therefore, additive within period separability between food and non-food consumption has to be assumed. In Chapter 3 we relax this assumption by considering a parameterization of the utility function which implies that the Euler equation for food consumption depends on consumption of other goods. This approach was proposed by Attanasio and Weber (1992). They estimate a very similar model using cohort data from the American Consumer Expenditure Survey. The advantage that we have is the panel structure of our data set (the Spanish Family Expenditure Survey). We do not need to group our data, and therefore we have a much larger cross-section dimension for our data set. Furthermore, we also consider the problem of aggregate shocks, which can invalidate the results based on cross-section asymptotics. If the aggregate shocks affect all the individuals in the same way, this problem can be easily solved by introducing time dummies in the model. However, if the effect of aggregate shocks is different for different families, the time dummies would not be of much help. If this were the case, we would expect to obtain different estimates of the parameters of the model for different periods of time. We consider two data sets for two periods of time (1978-83 and 1985-89) which are need to test the stability of our results.

In the last chapter, we discuss the specification and estimation of a system of intertemporal demands i.e. a Frisch demand system. The advantage of this approach is that we can obtain at the same time estimates of the parameters involved in the intra-temporal and inter-temporal allocation of consumption. This framework was introduced by Browning, Deaton and Irish (1985). In their paper they estimate a system of Frisch demands for consumption and labour supply. We start by discussing the different methods that have been used in the literature to estimate models of inter-temporal allocation of expenditure. We establish a set of criteria which should ideally be satisfied by a Frisch demand system in terms of consistency with the theory, flexibility and econometric tractability. Guided by these criteria, we chose a functional form for the Frisch system that allows us to estimate a system of <u>exact</u> Euler equations. We then estimate our model using panel data for Spain.

CHAPTER 2. ESTIMATING DYNAMIC MODELS FROM TIME SERIES OF CROSS-SECTIONS

1.- INTRODUCTION

As it has been stressed in the literature, the use of panel data is sometimes crucial to identify models of individual behaviour. For example, dynamic models can not be estimated using a single cross-section. Another example of the advantage of panel data is that we can take into account unobservable individual characteristics which may influence individual decisions. If the individual effects are correlated with the explanatory variables, the model can not be identified from a single cross-section. However, if these effects are constant over time, the model can be properly estimated using panel data.

The problem that arises at this level is that for many countries there is no panel data available with the information required to estimate some models of individual behaviour. For example, in the U.K. there is no panel data on household consumption and labour supply. However, a large survey on consumer expenditure and labour supply (The Family Expenditure Survey) is carried out with a regular periodicity. This type of data can not be treated as a real panel since the individuals in the sample are different from period to period. Nevertheless, the population can be divided in cohorts (groups with fixed membership over time) according to a certain characteristic (eg. year of birth), and the data on the sample means of the observations for each cohort in each time period can be used as a panel.

Furthermore, using cohort data we can avoid the attrition problem that often appears in true panels. There are several applied papers in the literature using this kind of data (Browning, Deaton and Irish (1985), Attanasio and Weber (1993), Blundell, Browning and Meghir (1994)).

It is important to notice that the cohort sample mean is only an estimator of the true cohort population mean. Therefore, the estimators of the parameters of the model based on the sample means will be biased, and these biases will only be negligible if the cohort sample sizes are sufficiently large.

The purpose of this chapter is to develop estimators for dynamic models using time series of cross-section data, which are consistent as the number of cohorts tends to infinity, for a fixed number of observations in the time series dimension, and a fixed number of members per cohort. The estimation of a static regression model was first consider by Deaton (1985). He proposed a corrected within groups estimator for the static regression model with unobservable individual effects, which is consistent for a fixed number of cohort members. Verbeek and Nijman (1993) analyze an alternative estimator which is consistent as the number of cohorts tends to infinity (for a fixed number of time periods and a fixed number of individuals per cohort). The correction arises naturally as a consequence of the errors in variable structure of the data. Verbeek and Nijman (1992) propose analytic formulas for the asymptotic bias of the classical within groups estimator (without correcting for measurement errors). The estimation of dynamic models from cohort data has been considered by

Moffitt (1993). He derives a class of cohort estimators as instrumental-variable estimators based on the micro data. However no attempt is made to correct for measurement errors.

For the dynamic model the within groups estimator is not even consistent using genuine panel data, unless the number of time periods tends to infinity (see Nickell (1981)). In applied work, we do not usually have data available for a large number of time periods. Therefore generalized method of moments (GMM) estimators, which are consistent for finite T, normally lead to less biased estimates (eg. Holtz-Eakin, Newey and Rosen (1988), and Arellano and Bond (1991)).

In this chapter we consider GMM estimators for a dynamic regression model with unobservable individual effects. Taking into account the errors in variable structure of cohort data, we can find instruments for the lagged dependent variable in the model in first differences, which are correlated with the disturbance term only trough the measurement errors. These instruments can be validly used if we correct the GMM estimator using the measurement errors variances.

The performance of the different estimators proposed depends mainly on two different things. One is the size of the sample we have available. If we have individual observations for a very short number of time periods, within groups will generally lead to poor estimates. The other thing to be considered, is the level of aggregation. Given a certain sample, there is a trade-off between improving the performance of the measurement error corrected estimators by choosing a large number of cohorts, versus diminishing the effect of the measurement errors by choosing a small number of cohorts each consisting of a large number of individuals.

The chapter is organized as follows. In section 2 we construct a measurement error corrected GMM estimator for the static regression model, which sets the framework that we are going to use in the rest of the chapter. In section 3, we extend this procedure to deal with dynamic models. We also consider a within groups estimator corrected for measurement errors which is consistent when the number of time periods tend to infinity and the cohort sizes are fixed. In section 4, we calculate analytic formulae for the asymptotic bias of different estimators for the AR(1) model, and we analyze their behaviour for different values of the parameters of the model. In section 5, we present and analyze the results from the Monte Carlo experiments. Section 6 concludes.

2.- ERRORS-IN-VARIABLE ESTIMATOR FOR THE STATIC MODEL

Consider the static regression model with individual effects

$$y_{it}=x_{it}'\beta+\theta_i+v_{it} \qquad \begin{aligned} v_{it} &\sim iid(0,\sigma^2) \\ \theta_i &\sim iid(0,\sigma_\theta^2) \end{aligned} \tag{2.1}$$

y_{it} is the dependent variable for individual i at time t, x_{it} is a vector of explanatory variables, θ_i is the individual effect, and v_{it} is the disturbance term. We will assume that $E(x_{it}v_{is})=0$ $\forall t,s$. If the individual effects are correlated with the explanatory variables, the model can not be identified with a single cross-section[2]. If the data available are time series of cross-sections the model can not be directly estimated using panel data techniques, since the individuals are different from period to period. However, the population can be divided in groups with fixed membership over time (cohorts) according to a certain characteristic. Let g be a random variable which determines the cohort membership for each individual (i.e. for any individual $i \in c$ if and only if $g_i \in I_c$). Taking expectations conditional on g_i in model (2.1), we have

$$E(y_{it}/g_i \in I_c) = E(x_{it}/g_i \in I_c)'\beta + E(\theta_i/g_i \in I_c) + E(v_{it}/g_i \in I_c) \qquad (2.2)$$

A necessary condition for identification is that the cohort population means vary across cohorts and over time. This is a sensible assumption, for example, if we are modelling consumption we can divide the population according to the year of birth, and we can expect consumption to vary with age. Hence, average consumption will be different for different cohorts and will vary over time.

We can rewrite (2.2) using a simple notation as

$$\begin{aligned} y^*_{ct} &= x^{*\prime}_{ct}\beta + \theta^*_c + v^*_{ct} \qquad c=1,..,C \\ &\qquad\qquad\qquad\qquad\quad t=1,..,T \\ v^*_{ct} &\sim iid(0,\sigma^2_{v^*}) \\ \theta^*_c &\sim iid(0,\sigma^2_{\theta^*}) \end{aligned} \qquad (2.3)$$

where $y^*_{ct} = E(y_{it}/g_i \in I_c)$, $x^*_{ct} = E(x_{it}/g_i \in I_c)$, etc. The v^*_{ct}'s are uncorrelated with the explanatory variables, while the cohort effects are potentially

---

[2] Unless we have external instruments available.

correlated.

The problem estimating this model is that we do not observe the cohort population means. However, we do observe a certain number of individuals on each group for each time period. We can assume that for any individual in a given cohort c

$$\begin{aligned} y_{it} &= y^*_{ct}+\zeta_{it} \\ x_{it} &= x^*_{ct}+\eta_{it} \end{aligned} \qquad \begin{pmatrix}\zeta_{it}\\ \eta_{it}\end{pmatrix} / g_i \in I_c \sim \mathrm{iid}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma^2_\zeta & \sigma'_{\zeta\eta}\\ \sigma_{\zeta\eta} & \Sigma_\eta\end{pmatrix}\right) \tag{2.4}$$

Then, we can consider the sample mean of the observations for each cohort in each time period

$$\begin{aligned} \frac{1}{nc}\sum_{i\in c} y_{it} &= y^*_{ct}+\frac{1}{nc}\sum_{i\in c}\zeta_{it} \qquad & y_{ct} &= y^*_{ct}+\zeta_{ct} \\ \frac{1}{nc}\sum_{i\in c} x_{it} &= x^*_{ct}+\frac{1}{nc}\sum_{i\in c}\eta_{it} \qquad & x_{ct} &= x^*_{ct}+\eta_{ct} \end{aligned}$$

$$\begin{pmatrix}\zeta_{ct}\\ \eta_{ct}\end{pmatrix} \sim \mathrm{iid}\left(\begin{pmatrix}0\\0\end{pmatrix}, \frac{1}{nc}\begin{pmatrix}\sigma^2_\zeta & \sigma'_{\zeta\eta}\\ \sigma_{\zeta\eta} & \Sigma_\eta\end{pmatrix}\right) \tag{2.5}$$

where nc is the number of individuals per cohort[3]. The sample means can be used as a panel subject to measurement errors, where the measurement error covariance matrix will in general be unknown, but it can be estimated using the micro survey data.

The covariance matrix of $(\zeta_{it},\eta_{it})'/g_i\in I_c$ and the variance of the v^*'s can and will normally depend on the choice of the cohorts. The only assumption we are making here is that the cohorts are chosen so that, the covariance matrix of $(\zeta_{it},\eta_{it})'/g_i\in I_c$ does not depend on the particular cohort

---

[3] nc is assumed to be constant across cohorts and over time to simplify notation. This assumption can be easily relaxed.

to which the individual belongs. The validity of this assumption will rely on the homogeneity of the cohorts chosen and can be relaxed without changing the major findings of this chapter[4].

Using (2.5), we can rewrite (2.3) in terms of the observables

$$y_{ct}=x_{ct}'\beta+\theta_c^*+u_{ct} \qquad \begin{matrix} c=1,..,C \\ t=1,..,T \end{matrix} \qquad (2.6)$$
$$u_{ct}=V_{ct}^*+\zeta_{ct}-\eta_{ct}'\beta$$

As we mention earlier, the cohort effects in (2.6) are potentially correlated with the explanatory variables in the model. However, as it happens when we work with genuine panel data, we can easily eliminate those effects using any operator orthogonal to the unity vector (e.g. first differences, deviations from time means, etc.).

Let us consider the model in first differences for the observables as a system of equations

$$\begin{matrix} \Delta y_{c2}=\Delta x_{c2}'\beta+\Delta u_{c2} \\ \dots\dots\dots\dots\dots\dots \\ \Delta y_{cT}=\Delta x_{cT}'\beta+\Delta u_{cT} \end{matrix} \qquad c=1,..,C \qquad (2.7)$$

Contrary to what happens when we work with panel data, the explanatory variables in model (2.7) are correlated with the error terms through the measurement errors. However, the matrix

[4] If the covariance matrix in (2.4) is different for different cohorts, we will use the observations on a particular cohort to estimate its covariance matrix.

$$Z_c = \begin{bmatrix} x'_{c1}..x'_{cT} & & \\ & \cdot & \\ & & \cdot \; x'_{c1}..x'_{cT} \end{bmatrix}$$

can be used as a matrix of instruments for the system. Rearranging the columns of Z_c in a convenient way, the moment restrictions are given by

$$E(Z_c'\Delta u_c) = \frac{1}{nc}\begin{bmatrix} \Sigma_\eta \\ \cdot \\ \cdot \\ \Sigma_\eta \\ -- \\ -\Sigma_\eta \\ \cdot \\ \cdot \\ -\Sigma_\eta \\ -- \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}\beta + \frac{1}{nc}\begin{bmatrix} -\sigma_{\zeta\eta} \\ \cdot \\ \cdot \\ -\sigma_{\zeta\eta} \\ -- \\ \sigma_{\zeta\eta} \\ \cdot \\ \cdot \\ \sigma_{\zeta\eta} \\ -- \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} = \Lambda\beta + \lambda \tag{2.8}$$

where $\Delta u_c = (\Delta u_{c2}, .., \Delta u_{cT})'$. The measurement error corrected GMM estimator of β (GMMC) is obtained by minimizing

$$\sum_{c=1}^{C}(Z_c'\Delta u_c - \Lambda\beta - \lambda)' A_c \sum_{c=1}^{C}(Z_c'\Delta u_c - \Lambda\beta - \lambda)$$

where the optimal choice of A_c is any consistent estimator of the inverse of the covariance matrix of $Z_c'\Delta u_c$ (cf. Hansen (1982)). The GMMC estimator is consistent for fixed T when C goes to infinity and is given by

$$\begin{aligned} \tilde{\beta} = {} & [\sum_{c=1}^{C}(\Delta X_c' Z_c + \Lambda') A_c \sum_{c=1}^{C}(Z_c'\Delta X_c + \Lambda)]^{-1} * \\ & [\sum_{c=1}^{C}(\Delta X_c' Z_c + \Lambda') A_c \sum_{c=1}^{C}(Z_c'\Delta y_c - \lambda)] \end{aligned} \tag{2.9}$$

where $\Delta y_c = (\Delta y_{c2}, .., \Delta y_{cT})'$, and $\Delta X_c = (\Delta x_{c2}, .., \Delta x_{cT})'$. If the covariance matrix of the measurement errors is unknown, the GMMC estimator in (2.9) is unfeasible. However, we can replace Λ and λ by consistent estimators based on the survey data[5].

Notice that the corresponding GMM estimator for the true panel coincides with the within groups estimator (see Arellano and Bover (1994)). However this is no longer the case when we are using cohort data, and this estimator does not coincide with the measurement error corrected within groups estimator proposed by Deaton (1985) which is consistent using T asymptotics. The GMM estimator in (2.9) does not coincide with Deaton's first difference estimator either. Deaton's estimator is based on a linear combination of the moment restrictions we use in (2.8).

3.- ESTIMATION OF A DYNAMIC MODEL USING COHORT DATA

Consider the following dynamic model

$$y_{it}=\alpha y_{it-1}+x_{it}'\beta+\theta_i+v_{it} \qquad \begin{array}{l} v_{it} \sim \text{iid}(0,\sigma^2) \\ \theta_i \sim \text{iid}(0,\sigma_\theta^2) \end{array} \tag{3.1}$$

We will assume that the individual effects are potentially correlated with all the explanatory variables in the model, and that $E(x_{it}v_{is})=0\ \forall t,s$. If the data available are repeated cross-sections containing different individuals over time, as we said earlier, we can divide the population in C cohorts

[5] In the appendix we obtain the asymptotic distribution of the GMM estimator for the dynamic model when the covariance matrix of the measurement errors is unknown. For the static model this distribution can be derived analogously.

containing the same individuals over time. Taking expectations conditional on the cohort the individual belongs to, and proceeding in a similar way as we did above for the static model, we can write the cohort population version of model (3.1) as

$$y^*_{ct}=\alpha y^*_{ct-1}+x^{*\prime}_{ct}\beta+\theta^*_c+v^*_{ct} \qquad \begin{array}{l} c=1,..,C \\ t=2,..,T \end{array}$$

$$\begin{aligned} \theta^*_c &\sim iid(0,\sigma^2_{\theta^*}) \\ v^*_{ct} &\sim iid(0,\sigma^2_{v^*}) \end{aligned} \tag{3.2}$$

The unobservable cohort population means can be estimated by their sample counterparts.

The autoregressive model for a true panel can not be consistently estimated (as $N\rightarrow\infty$ for fixed T) using dummy variables for the fixed effects, or equivalently using the within groups estimator because in the deviations from time means model the lagged dependent variable is correlated with the error term. Notice that this type of convergence is relevant because panel data are in most of the cases available for very few periods of time. In the genuine panel case, we can estimate the model in first differences using lagged, present, and future values of the x's as instruments for the lagged y's provided that the x's are strictly exogenous with respect to the v's. However, this procedure is no longer possible in the pure autoregressive model without exogenous explanatory variables, or in a model with only predetermined variables, unless we have external instruments available. Furthermore, even if some of the regressors are strictly exogenous, we can find more efficient estimators adding lagged values of the predetermined variables to the instrument set (see Arellano and Bond (1991)).

When we are working with time series of cross-sections it is very important to establish the type of asymptotics for different estimators. Depending on the data available and on the size of the cohorts that we have chosen, some estimators will lead to better estimates than others. We are going to present different types of estimators for the first order autoregressive model with explanatory variables, and we will discuss their applicability depending on the sample size.

The variables in model (3.2) are unobservable but they can be estimated by the cohort sample means. Using (2.5), we can rewrite the model in terms of the observables as

$$\begin{aligned} y_{ct} &= \alpha y_{ct-1} + x_{ct}'\beta + \theta_c^* + u_{ct} \qquad \begin{matrix} c=1,..,C \\ t=2,..,T \end{matrix} \\ u_{ct} &= v_{ct}^* + \zeta_{ct} - \alpha\zeta_{ct-1} - \eta_{ct}'\beta \end{aligned} \tag{3.3}$$

We will first consider a measurement error corrected within groups estimator. As we said above, even for the true panel, the within groups estimator for the dynamic model is only consistent when T tends to infinity. When we are dealing with cohort data we can achieve consistency (as $T\to\infty$) correcting the within groups estimator by the measurement error variances.

The model in deviations from time means is

$$\tilde{y}_{ct} = \alpha\tilde{y}_{ct-1} + \tilde{x}_{ct}'\beta + \tilde{u}_{ct} \tag{3.4}$$

where $\tilde{y}_{ct} = y_{ct} - \frac{1}{T-1}\sum_{s=2}^{T} y_{cs}$, $\tilde{y}_{ct-1} = y_{ct-1} - \frac{1}{T-1}\sum_{s=1}^{T-1} y_{cs}$, $\tilde{x}_{ct} = x_{ct} - \frac{1}{T-1}\sum_{s=2}^{T} x_{cs}$

and $\tilde{u}_{ct}=u_{ct}-\frac{1}{T-1}\sum_{s=2}^{T}u_{cs}$. When T goes to infinity all the variables on the right hand size of (3.4) are correlated with the error term only trough the measurement errors in the following way

$$E\left[\begin{pmatrix}\tilde{y}_{ct-1}\\ \tilde{x}_{ct}\end{pmatrix}\tilde{u}_{ct}\right]=E\left[\begin{pmatrix}\tilde{y}^*_{ct-1}\\ \tilde{x}^*_{ct}\end{pmatrix}\tilde{u}_{ct}\right]+E\left[\begin{pmatrix}\tilde{\zeta}_{ct-1}\\ \tilde{\eta}_{ct}\end{pmatrix}\tilde{u}_{ct}\right]$$

$$E\left[\begin{pmatrix}\tilde{y}^*_{ct-1}\\ \tilde{x}^*_{ct}\end{pmatrix}\tilde{u}_{ct}\right]\to 0 \quad \text{as } T\to\infty$$

$$E\left[\begin{pmatrix}\tilde{\zeta}_{ct-1}\\ \tilde{\eta}_{ct}\end{pmatrix}\tilde{u}_{ct}\right]\to\frac{1}{nc}\begin{bmatrix}-\alpha\sigma^2_\zeta\\ \sigma_{\zeta\eta}-\Sigma_\eta\beta\end{bmatrix} \quad \text{as } T\to\infty$$

and the measurement error corrected within groups estimator (WGC) is given by

$$\begin{pmatrix}\hat{\alpha}_{wgc}\\ \hat{\beta}_{wgc}\end{pmatrix}=\left(\frac{1}{C(T-1)}\sum_{c=1}^{C}\sum_{t=2}^{T}\begin{pmatrix}\tilde{y}^2_{ct-1} & \tilde{y}_{ct-1}\tilde{x}'_{ct}\\ \tilde{x}_{ct}\tilde{y}_{ct-1} & \tilde{x}_{ct}\tilde{x}'_{ct}\end{pmatrix}-\frac{1}{nc}\begin{pmatrix}\sigma^2_\zeta & 0\\ 0 & \Sigma_\eta\end{pmatrix}\right)^{-1}* \tag{3.5}$$
$$\left(\frac{1}{C(T-1)}\sum_{c=1}^{C}\sum_{t=2}^{T}\begin{pmatrix}\tilde{y}_{ct-1}\tilde{y}_{ct}\\ \tilde{x}_{ct}\tilde{y}_{ct}\end{pmatrix}-\frac{1}{nc}\begin{pmatrix}0\\ \sigma_{\zeta\eta}\end{pmatrix}\right)$$

In applied work we do not always have data available for a large number of time periods, and in this case within groups is not an appropriate technique. However we can think of a measurement error corrected within groups estimator for fixed T, which will not be consistent but might lead to an improvement by eliminating completely the asymptotic bias due to measurement errors in the following way

$$E\left[\begin{pmatrix}\tilde{\zeta}_{ct-1}\\ \tilde{\eta}_{ct}\end{pmatrix}\tilde{u}_{ct}\right] = \frac{1}{nc}\begin{bmatrix} -\left(\frac{T}{(T-1)^2}+\alpha\frac{T-2}{T-1}\right)\sigma_{\zeta}^2+\frac{T}{(T-1)^2}\sigma_{\zeta\eta}'\beta \\ \left(\frac{T-2}{T-1}+\alpha\frac{T}{(T-1)^2}\right)\sigma_{\zeta\eta}-\frac{T-2}{T-1}\Sigma_{\eta}\beta \end{bmatrix}$$

and the measurement errors corrected within groups estimator for fixed T (WGCT) is given by

$$\begin{pmatrix}\hat{\alpha}_{wgct}\\ \hat{\beta}_{wgct}\end{pmatrix}=\left(\frac{1}{C(T-1)}\sum_{c=1}^{C}\sum_{t=2}^{T}\begin{pmatrix}\tilde{y}_{ct-1}^2 & \tilde{y}_{ct-1}\tilde{x}_{ct}'\\ \tilde{x}_{ct}\tilde{y}_{ct-1} & \tilde{x}_{ct}\tilde{x}_{ct}'\end{pmatrix} - \frac{1}{nc}\begin{pmatrix}\frac{T-2}{T-1}\sigma_{\zeta}^2 & -\frac{T}{(T-1)^2}\sigma_{\zeta\eta}'\\ -\frac{T}{(T-1)^2}\sigma_{\zeta\eta} & \frac{T-2}{T-1}\Sigma_{\eta}\end{pmatrix}\right)^{-1}$$

$$* \left(\frac{1}{C(T-1)}\sum_{c=1}^{C}\sum_{t=2}^{T}\begin{pmatrix}\tilde{y}_{ct-1}\tilde{y}_{ct}\\ \tilde{x}_{ct}\tilde{y}_{ct}\end{pmatrix}-\frac{1}{nc}\begin{pmatrix}-\frac{T}{(T-1)^2}\sigma_{\zeta}^2\\ \frac{T-2}{T-1}\sigma_{\zeta\eta}\end{pmatrix}\right) \qquad (3.6)$$

This estimator will not be consistent for fixed T, however, it might lead to better estimates than the classical within groups estimator for small values of T[6].

As we said above within groups estimation is not an appropriate technique when the time series dimension of the sample is small. Hence, we are going to study an alternative estimator which is consistent for finite T when C goes to infinity.

Taking first differences in model (3.3) yields to

---

[6] In section 4 we obtain the asymptotic bias of the WG and WGCT estimators (For fixed T as $C\rightarrow\infty$) for the AR(1) model without explanatory variables. At least in this case the asymptotic bias of the WGCT estimator is smaller that the bias of the WG.

$$\begin{array}{ll} \Delta y_{c3}=\alpha\Delta y_{c2}+\Delta x_{c3}'\beta+\Delta u_{c3} & \\ \cdots\cdots\cdots\cdots\cdots\cdots & c=1,..,C \\ \Delta y_{cT}=\alpha\Delta y_{cT-1}+\Delta x_{cT}'\beta+\Delta u_{cT} & \end{array}$$

If the x^*'s are strictly exogenous, i.e. $E(x_{ct}^* v_{cs}^*)=0$ $t,s=1,..,T$, the matrix

$$Z_c=\begin{bmatrix} y_{c1}x_{c1}'..x_{cT}' & & \\ & \cdot & \\ & & \cdot \\ & & y_{c1}..y_{cT-2}x_{c1}'..x_{cT}' \end{bmatrix}$$

can be used as a matrix of instruments for the system (see Arellano and Bond (1991) for an analysis of this model in the genuine panel case). Rearranging the columns of Z_c in a convenient way, the moment restrictions are given by

$$E(Z_c'\Delta u_c)=\frac{1}{nc}\begin{bmatrix} 0 & 0 \\ \cdots & \cdots \\ 0 & 0 \\ --- & --- \\ \sigma_\zeta^2 & 0 \\ \cdots & \cdots \\ \sigma_\zeta^2 & 0 \\ --- & --- \\ \sigma_{\zeta\eta} & 0 \\ \cdots & \cdots \\ \sigma_{\zeta\eta} & 0 \\ --- & --- \\ -\sigma_{\zeta\eta} & \Sigma_\eta \\ \cdots & \cdots \\ -\sigma_{\zeta\eta} & \Sigma_\eta \\ --- & --- \\ 0 & -\Sigma_\eta \\ \cdots & \cdots \\ 0 & -\Sigma_\eta \end{bmatrix}\begin{pmatrix}\alpha \\ \beta\end{pmatrix}+\frac{1}{nc}\begin{bmatrix} 0 \\ .. \\ 0 \\ -- \\ 0 \\ .. \\ 0 \\ -- \\ 0 \\ .. \\ 0 \\ -- \\ -\sigma_{\zeta\eta} \\ .. \\ -\sigma_{\zeta\eta} \\ -- \\ \sigma_{\zeta\eta} \\ .. \\ \sigma_{\zeta\eta} \end{bmatrix}=\Lambda\delta+\lambda$$

where $\Delta u_c=(\Delta u_{c3},..,\Delta u_{cT})'$. The measurement error corrected GMM estimator of δ (GMMC) is obtained by minimizing

$$\sum_{c=1}^{C}(Z_c'\Delta u_c-\Lambda\delta-\lambda)'A_c\sum_{c=1}^{C}(Z_c'\Delta u_c-\Lambda\delta-\lambda)$$

and is given by

$$\begin{aligned}\tilde{\delta}=&[\sum(\Delta W_c'Z_c+\Lambda')A_c\sum(Z_c'\Delta W_c+\Lambda)]^{-1}*\\&[\sum(\Delta W_c'Z_c+\Lambda')A_c\sum(Z_c'\Delta y_c-\lambda)]\end{aligned}\tag{3.7}$$

where $W_c=(y_{c(-1)},X_c)$, $y_{c(-1)}=(y_{c2},..,y_{cT-1})'$, $X_c=(x_{c3},..,x_{cT})'$.

The asymptotic distribution of the GMMC estimator when Λ and λ are assumed to be known can be derived straightforwardly using a standard central limit theorem

$$\sqrt{C}(\tilde{\delta}-\delta)\to_d N(0,(D_0'A_0D_0)^{-1}D_0'A_0V_0A_0D_0(D_0'A_0D_0)^{-1})\tag{3.8}$$

where

$$\begin{aligned}A_0=&\ \text{plim}\ A_c\\D_0=&\ \text{plim}\frac{\partial}{\partial\delta'}\sum_c(Z_c'\Delta u_c-\Lambda\delta-\lambda)=-(E(Z_c'\Delta W_c)+\Lambda)\\V_0=&\ E[(Z_c'\Delta u_c-\Lambda\delta-\lambda)(Z_c'\Delta u_c-\Lambda\delta-\lambda)']\end{aligned}$$

We can obtain a consistent estimator of the asymptotic covariance matrix in (3.8) by replacing A_0 by A_c and D_0, V_0 by

$$\begin{aligned}\hat{D}=&-\frac{1}{C}\sum_{c=1}^{C}Z_c'\Delta W_c-\Lambda\\\hat{V}=&\frac{1}{C}\sum_{c=1}^{C}(Z_c'\Delta\tilde{u}_c-\Lambda\tilde{\delta}-\lambda)(Z_c\Delta\tilde{u}_c-\Lambda\tilde{\delta}-\lambda)'\end{aligned}$$

where $\Delta\tilde{u}_c$ is the vector of residuals.

The GMMC estimator in (3.7) is consistent as $C\to\infty$ for any choice of the weighting matrix. For example, we could use

$$A_c=\left[\frac{1}{C}\sum_{c=1}^{C} Z_c' H Z_c\right]^{-1} \qquad H=\begin{bmatrix} 2 & -1 & \cdots & 0 \\ -1 & 2 & \cdots & 0 \\ & \cdots\cdots & & \\ 0 & \cdots & -1 & 2 \end{bmatrix}$$

Notice that in the genuine panel case, this matrix is the natural choice for the one step GMM and it is optimal if the disturbances in model (3.1) are uncorrelated and homoscedastic.

We can obtain a more efficient estimator in a second step using, as weighting matrix, any consistent estimator of V_0^{-1}. Once we have a preliminary consistent estimate $\tilde{\delta}$, the two step GMMC is obtained using

$$A_c=\left[\frac{1}{C}\sum_{c=1}^{C} (Z_c'\Delta\tilde{u}_c-\Lambda\tilde{\delta}-\lambda)(Z_c'\Delta\tilde{u}_c-\Lambda\tilde{\delta}-\lambda)'\right]^{-1}$$

where $\Delta\tilde{u}_c$ is the vector of residuals based on the one step estimator.

When the covariance matrix of the measurement errors is unknown, it has to be estimated using the micro survey data. Λ and λ are replaced by consistent estimates, and the asymptotic distribution of the GMMC estimator changes (see appendix a).

4.- <u>ASYMPTOTIC BIASES IN THE AR(1) MODEL</u>

In section 3 we have considered different types of measurement error corrected estimators for a dynamic model with individual effects, that are consistent under different kinds of asymptotics. The corresponding non-corrected estimators are not consistent under the same type of asymptotics, and it is interesting to know the size of their asymptotic biases, in order to have an idea on whether it is worth or not to use the measurement error correction. Furthermore, the within groups estimators are not consistent for fixed T, however, in order to compare their asymptotic behaviour with the GMM type estimators that rely on C asymptotics, we will calculate the asymptotic biases of the within groups estimators for fixed T.

The asymptotic biases of these estimators for the dynamic model with explanatory variables will depend on the particular mechanism generating the x's. For this reason, we will concentrate in this section on the pure AR(1) model.

Consider the AR(1) model with individual effects

$$y_{it}=\alpha y_{it-1}+\theta_i+v_{it} \tag{4.1}$$

Proceeding analogously as we did earlier for the model with exogenous variables, we can write the cohort population version of (4.1) as

$$y^*_{ct}=\alpha y^*_{ct-1}+\theta^*_c+v^*_{ct} \quad \begin{matrix} c=1,..,C \\ t=2,..,T \end{matrix} \tag{4.2}$$

The sample means for each cohort are consistent estimators of the cohort population means. Following the same notation as in section 3, we can rewrite (4.2) in terms of the observables

$$\begin{aligned} y_{ct}&=\alpha y_{ct-1}+\theta^*_c+u_{ct} \\ u_{ct}&=v^*_{ct}+\zeta_{ct}-\alpha\zeta_{ct-1} \end{aligned} \tag{4.3}$$

where

$$y_{ct}=y^{*}_{ct}+\zeta_{ct} \qquad \zeta_{ct}\sim iid\left(0,\frac{1}{nc}\sigma^{2}_{\zeta}\right)$$

a) Within Groups Estimators

We can write (4.3) in deviations from time means as

$$\tilde{y}_{ct}=\alpha\tilde{y}_{ct-1}+\tilde{u}_{ct} \tag{4.4}$$

The OLS estimator of α in the model above is the within groups estimator. Its asymptotic bias for fixed T is (see appendix b)

$$\operatorname{plim}_{C\to\infty}(\hat{\alpha}_{wg}-\alpha)=-\frac{h_T\phi+\left[\frac{T}{T-1}+(T-2)\alpha\right]}{\frac{T-2}{1-\alpha^2}\phi-\frac{2\alpha}{1-\alpha^2}h_T\phi+T-2} \tag{4.5}$$

where

$$h_T=\frac{1}{1-\alpha}\left[1-\frac{1}{T-1}\frac{1-\alpha^{T-1}}{1-\alpha}\right] \quad \text{and} \quad \phi=\frac{nc\sigma^{2}_{v^*}}{\sigma^{2}_{\zeta}} \tag{4.6}$$

The asymptotic bias (4.5) is negative if $0\leq\alpha<1$, and its absolute value decreases as ϕ or T increase. If $-1<\alpha<0$ the asymptotic bias can be positive or negative, and it will increase or decrease with T and ϕ depending on the particular values of α, T and ϕ.

Next, using T asymptotics

$$\operatorname{plim}_{T\to\infty}(\hat{\alpha}_{wg}-\alpha) = -\frac{\alpha}{\frac{1}{1-\alpha^2}\phi+1} \tag{4.7}$$

In section 3 we obtained the measurement error corrected within groups estimator (3.5), which is consistent using T asymptotics (WGC). For the pure AR(1) model (4.4), this estimator is given by

$$\hat{\alpha}_{wgc} = \frac{\frac{1}{C(T-1)}\sum_{t=2}^{T}\sum_{c=1}^{C}\tilde{y}_{ct}\tilde{y}_{ct-1}}{\frac{1}{C(T-1)}\sum_{t=2}^{T}\sum_{c=1}^{C}\tilde{y}_{ct-1}^{2}-\frac{\sigma_{\zeta}^{2}}{nc}}$$

and its asymptotic bias for fixed T is

$$\operatorname{plim}_{C\to\infty}(\hat{\alpha}_{wgc}-\alpha) = -\frac{h_T\phi+\left[\frac{T}{T-1}-\alpha\right]}{\frac{T-2}{1-\alpha^2}\phi-\frac{2\alpha}{1-\alpha^2}h_T\phi-1} \tag{4.8}$$

The asymptotic bias of this estimator can be bigger or smaller than the non-corrected WG depending on the parameters of the model. The bias is negative if $\phi>2$[7]. If the bias is negative its absolute value decreases as T or ϕ increase.

For the dynamic model with explanatory variables we also derived another measurement error corrected within groups estimator (WGCT). This estimator was obtained using the appropriate correction to eliminate completely the asymptotic bias due to measurement errors (for fixed T). For

---

[7] Notice that $\phi<2$ means that if for instance nc=25 then $\sigma_{v^*}^2/\sigma_\zeta^2<0.08$, i.e. the measurement error variance is at least 12.5 times bigger than the variance of v^*_c. This seems to be quite unreasonable from our point of view.

the AR(1) model it is given by

$$\hat{\alpha}_{wgct}=\frac{\frac{1}{C(T-1)}\sum_{t=2}^{T}\sum_{c=1}^{C}\tilde{y}_{ct}\tilde{y}_{ct-1}+\frac{T}{(T-1)^2}\frac{\sigma_\zeta^2}{nc}}{\frac{1}{C(T-1)}\sum_{t=2}^{T}\sum_{c=1}^{C}\tilde{y}_{ct-1}^2-\frac{T-2}{T-1}\frac{\sigma_\zeta^2}{nc}}$$

and

$$\text{plim}_{C\to\infty}(\hat{\alpha}_{wgct}-\alpha)=-\frac{h_T}{\frac{T-2}{1-\alpha^2}-\frac{2\alpha}{1-\alpha^2}h_T} \tag{4.9}$$

This bias coincides with the asymptotic bias of the within groups estimator for the true panel derived by Nickell (1981), and it is smaller than the asymptotic bias of the WG estimator (at least for α positive). Notice that this bias is always negative and it decreases as T increases (it approaches zero as $T\to\infty$)

b) <u>GMM Estimator</u>

Let us consider model (4.3) in first differences as a system of equations

$$\begin{array}{ll} \Delta y_{c3}=\alpha\Delta y_{c2}+\Delta u_{c3} & \\ \cdots\cdots\cdots\cdots & c=1,..,C \\ \Delta y_{cT}=\alpha\Delta y_{cT-1}+\Delta u_{cT} & \end{array}$$

We can use the matrix

$$Z_c = \begin{bmatrix} y_{c1} & & & \\ & y_{c1}y_{c2} & & \\ & & \cdot & \\ & & & \cdot \\ & & & y_{c1}\cdot\cdot y_{cT-2} \end{bmatrix}$$

as a matrix of instruments for the system. The GMM estimator without correcting for measurement errors is given by

$$\tilde{\alpha} = \frac{\left(\sum_{c=1}^{C} Z_c' \Delta y_{c(-1)}\right)' A_c \sum_{c=1}^{C} Z_c' \Delta y_c}{\left(\sum_{c=1}^{C} Z_c' \Delta y_{c(-1)}\right)' A_c \sum_{c=1}^{C} Z_c' \Delta y_{c(-1)}}$$

and the asymptotic bias is (see appendix c)

$$\text{plim}_{c\to\infty}(\tilde{\alpha}-\alpha) = - \frac{\left[v\frac{1}{1+\alpha}\phi+\mu\right]' A_0 \mu \alpha}{\left[v\frac{1}{1+\alpha}\phi+\mu\right]' A_0 \left[v\frac{1}{1+\alpha}\phi+\mu\right]} \tag{4.10}$$

where

$$\begin{aligned} A_0 &= \text{plim } A_c \\ \mu &= (1,0,1,0,0,1,..)' \\ v &= (1,\alpha,1,\alpha^2,\alpha,1,...)' \end{aligned}$$

The asymptotic bias in (4.10) is negative if α is positive, and it decreases as ϕ increases.

In the figures at the end of this section, we have represented the absolute value of the asymptotic biases derived above as a function of α[8]. We have considered different values of T (T=5,7,10,15), and different values for the variances $\sigma_{v^*}^2$, σ_ζ^2 and σ_θ^2. We can consider that σ_ζ^2/nc is the

[8] We have considered positive values of α, and values of $\phi>2$. Therefore, as we explained earlier, the asymptotic biases are all negative.

variance of the within cohorts component of the model, and σ_{v*}^2, and σ_θ^2 are the variances of the between cohorts components of the model (the time varying and the time invariant components respectively). Notice that the asymptotic bias (for fixed T) of the within groups estimators depends on T, α, and ϕ (the ratio of the time varying component of the between cohorts variance to the within cohort variance); and the asymptotic bias of the GMM estimator depends also on $\rho = nc\sigma_\theta^2/\sigma_\zeta^2$ (the ratio of the time invariant component of the between cohorts variance to the within cohort variance) through the matrix A_0[9].

The asymptotic bias of the GMM estimator (4.10) is presented in figure 1. As α increases, the asymptotic bias increases. For values of α not very close to one the bias is small (it is zero for $\alpha=0$) and increases very slowly. However, for values of α close to one the bias increases rather quickly, and it can be very large for values of α around 0.9 (the actual value of the bias depends on the other parameters: the ratios of the variances (ϕ and ρ) and T).

Looking at the top of figure 1, we can see that the bigger ϕ is, the smaller is the bias. The reason is that an increase in this ratio means, on the one hand, an increase in the proportion of the between cohort variance due to the time varying component and on the other hand, an increase in the between cohorts component of the total variance, and both have a positive effect on the bias (positive here means that the bias decreases).

---

[9] We are considering the two step GMM estimator and hence A_0 is the inverse of the covariance matrix of $Z'_c\Delta u_c$ which depends on ρ.

The influence of ρ is weaker and has the opposite sign (see bottom of fig.1). An increase in this ratio would mean an increase in the between cohorts component of the total variance, and would lead to a reduction of the bias, but at the same time, the proportion of the between cohorts variance due to the time invariant component increases and this has a negative effect on the bias. This two opposite effects together produce a total negative effect (the bias increases).

The asymptotic bias of the GMM estimator depends also on the time series dimension. As T increases the asymptotic bias becomes smaller (see fig.1).

The asymptotic bias of the WG estimator (4.5) is presented in figure 2. This bias depends on α, ϕ and T. As α increases the asymptotic bias of the WG estimator increases at a low rate for values of α not very close to one and at a high rate for values of α close to one. The asymptotic bias is not zero for $\alpha = 0$.

The asymptotic bias of the WG estimator is smaller the bigger ϕ is. When the within cohort variance approaches zero, the asymptotic bias of the WG estimator approaches the asymptotic bias for the true panel which is not zero for finite T. The asymptotic bias of the WG estimator depends also on the time series dimension of the data set. As T increases the asymptotic bias decreases, however, for $T=10$ or $T=15$ it is still non negligible.

In figure 3 we present the asymptotic bias of the different estimators proposed in this section, in order to compare their behaviour. As we mention above, the asymptotic bias of the WGC estimator (for fixed T) can be bigger or smaller than the asymptotic bias of the WG estimator depending on the parameters of the model. However when we use the appropriate measurement error correction for finite T (WGCT estimator), we eliminate completely the bias due to the measurement error problem, and the asymptotic bias coincides with the asymptotic bias of the within groups estimator for the true panel, which is always smaller than the asymptotic bias of the non-corrected estimator (WG).

If we compare the asymptotic bias of the WG and GMM estimators, we can see that in most of the cases, the asymptotic bias of the GMM estimator is smaller than the asymptotic bias of the WG. Only when α is close to one and T is not very small the WG estimator will lead to better asymptotic results.

figure 1

ASYMPTOTIC BIAS GMM ESTIMATOR

within cohort variance = 0.5

nc=25

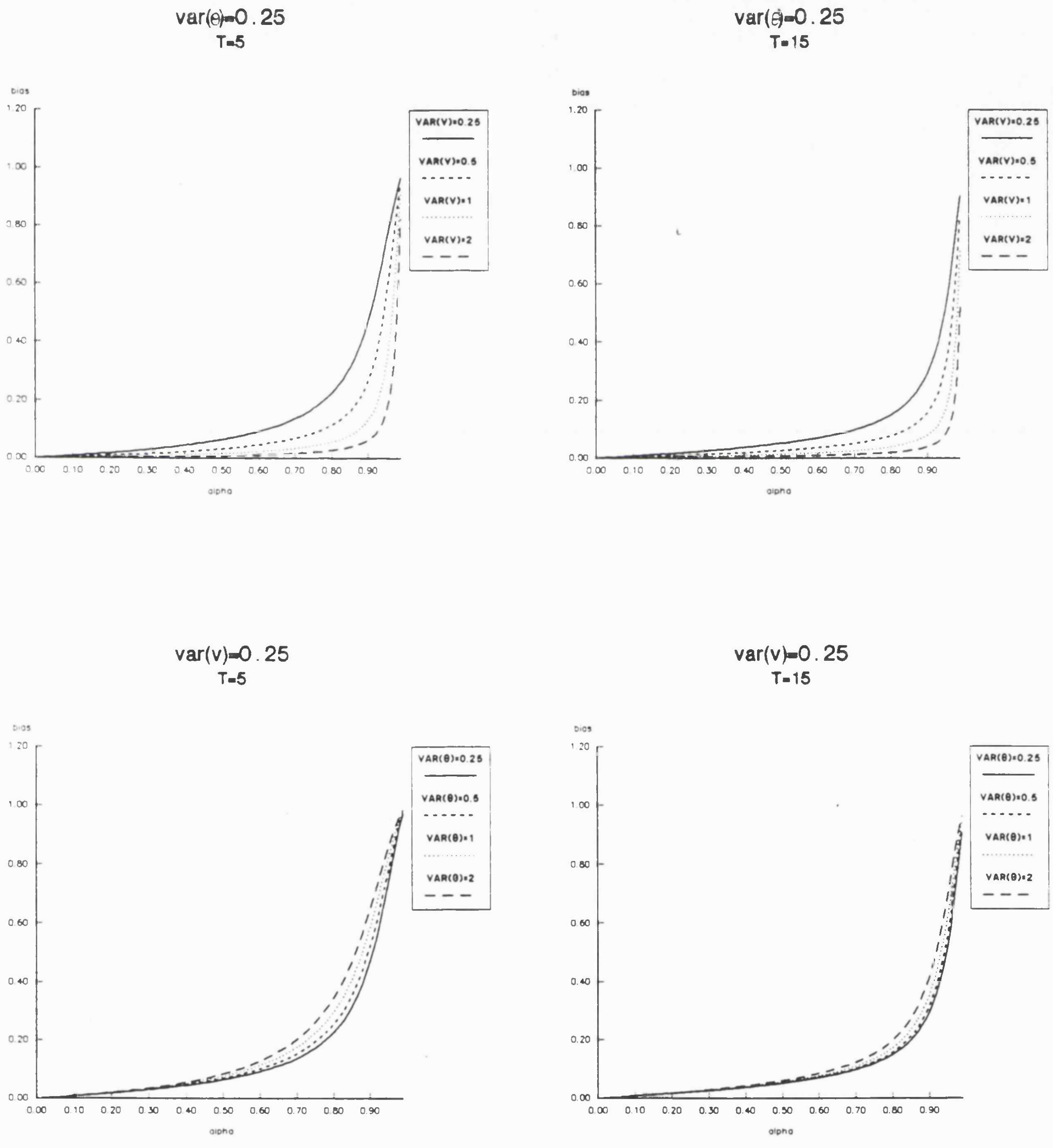


figure 2
ASYMPTOTIC BIAS WG ESTIMATOR
within cohort variance = 0.5
nc=25

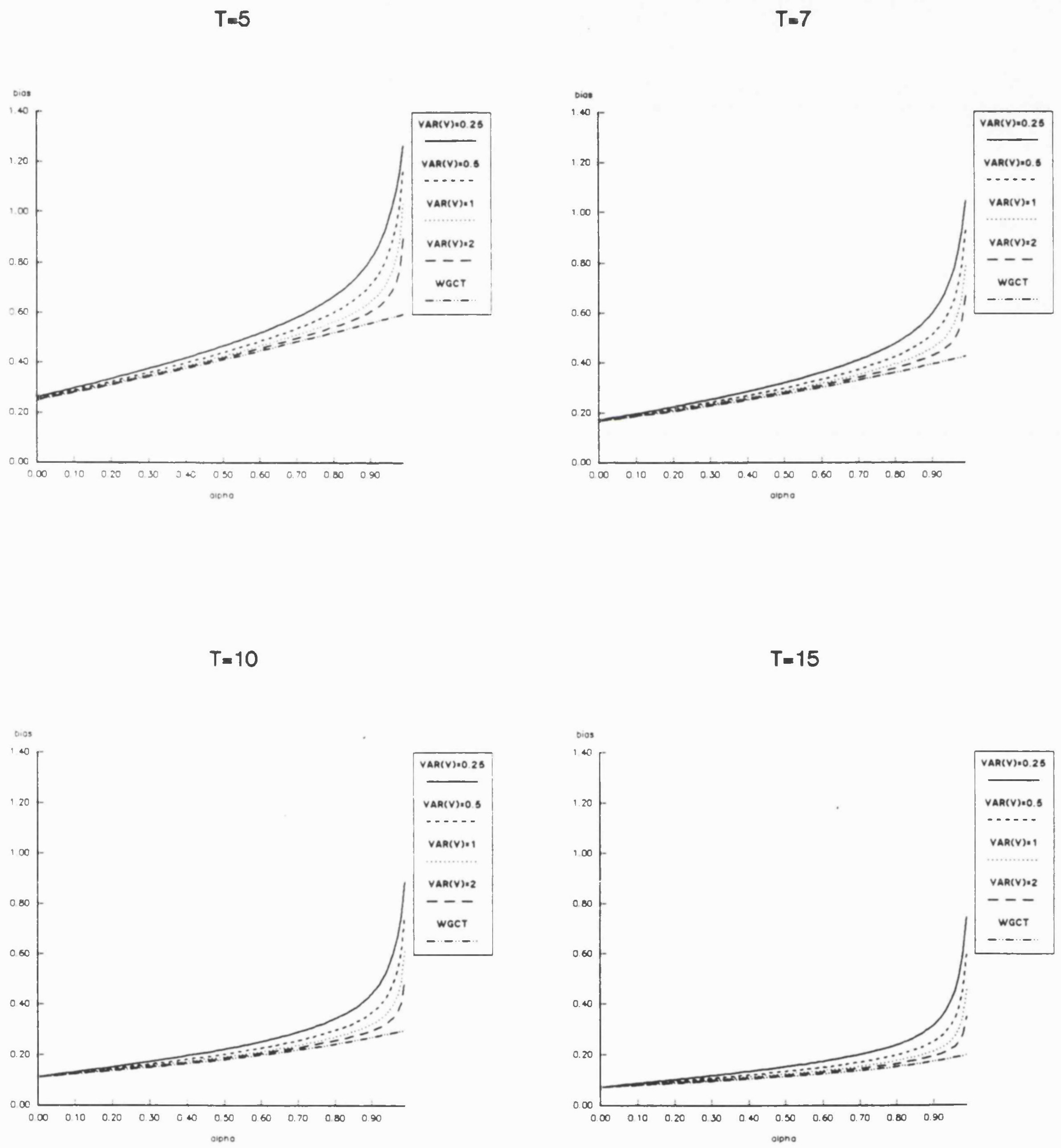
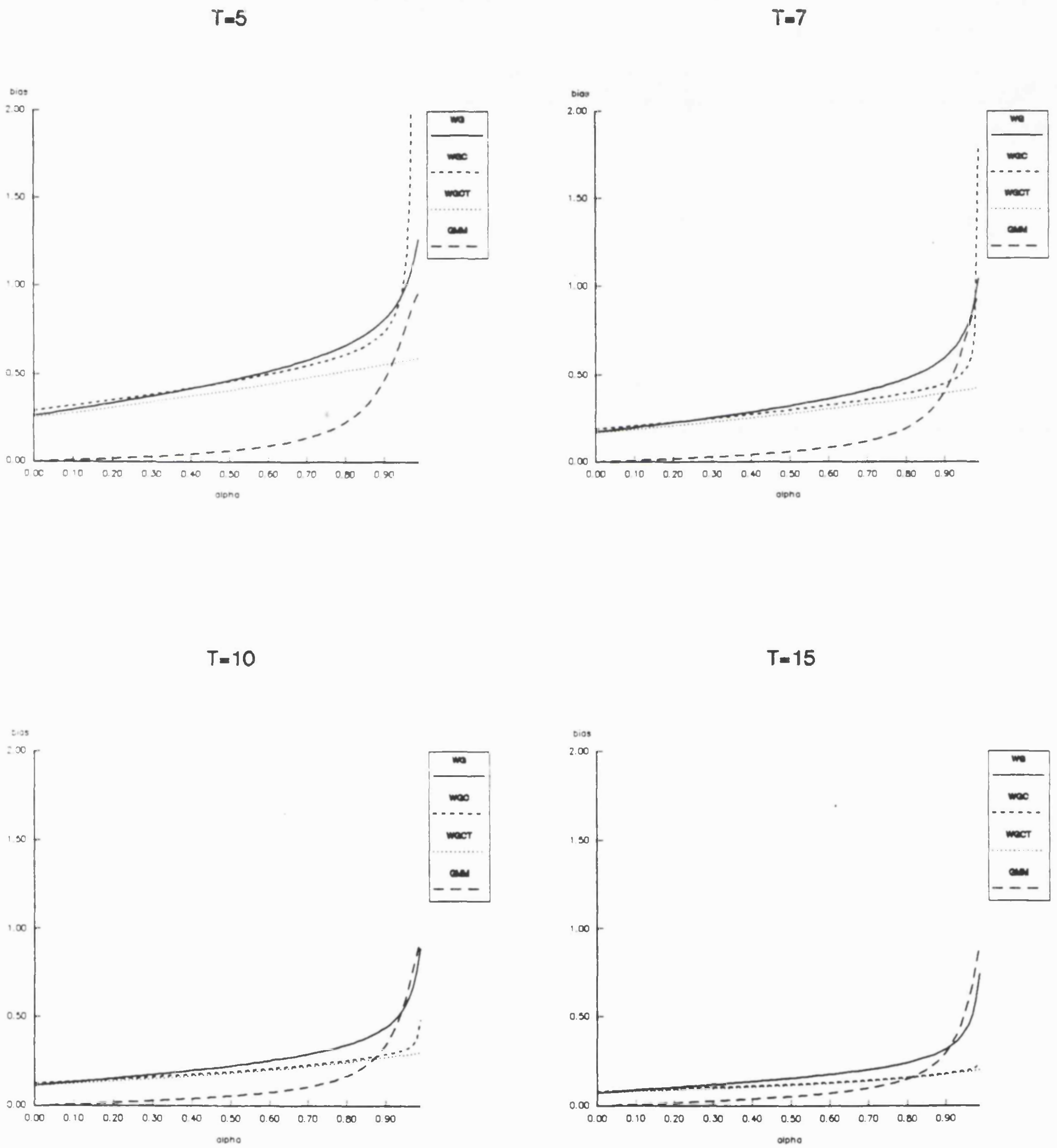


figure 3

ASYMPTOTIC BIASES

within cohort variance = 0.5

nc=25, var(v)=0.25, var(θ)=0.25



5.- MONTE CARLO SIMULATIONS

In section 4 we obtained analytic formulae for the asymptotic biases of different estimators for the AR(1) model without explanatory variables, and we presented some figures that help us to analyze their behaviour for different values of the parameters of the model. However, it is interesting to know up to what extent the asymptotic behaviour approximates the actual behaviour observed in finite samples.

To get an idea of the finite sample performance of the different estimators for the AR(1) model, we have carried out Monte Carlo simulations for different values of the parameters of the model.

The data were generated using the following model. First, the cohort population means were constructed using an AR(1) model

$$y^*_{ct} = \alpha y^*_{ct-1} + \theta^*_c + v^*_{ct}$$

$$v^*_{ct} \sim \text{iid } N(0,\sigma^2_{v^*}) \qquad \theta^*_c \sim \text{iid } N(0,\sigma^2_{\theta^*})$$

Then, the individual observations on each cohort were generated as follows

$$y_{it} \sim \text{iid } N(y^*_{ct},\sigma^2_{\zeta})$$

After generating the sample, the cohort sample means and variances were calculated.

Using the model above, the variance of y can be decomposed in the following way

$$\sigma_y^2 = \sigma_\zeta^2 + \frac{\sigma_{v^*}^2}{1-\alpha^2} + \frac{\sigma_{\theta^*}^2}{(1-\alpha)^2}$$

That is

$$\begin{matrix}\text{total} \\ \text{variance}\end{matrix} = \begin{matrix}\text{within} \\ \text{cohort} \\ \text{variance}\end{matrix} + \begin{matrix}\text{time varying + time invarying} \\ \text{between cohorts} \\ \text{variance}\end{matrix}$$

It is clear that the performance of the estimators will depend on the proportion of the variance due to the different components. In order to compare the behaviour of the estimators for different values of the autoregressive parameter, without having the additional effect of a change in the composition of the variance, we have chosen, for each α, the appropriate values of $\sigma^2_{v^*}$ and $\sigma^2_{\theta^*}$ to keep this proportion constant.

We have performed experiments for different time dimensions T=5,7,10,15, different values of the autoregressive parameter α=0.1, 0.5, 0.9, and different proportions on the composition of the variance of y. The size of the cross-sections is 2000 divided in 80 cohorts of 25 individuals each. The results from the simulations are summarized in tables 1-4[10].

Let us consider first the non-corrected estimators. It is clear from the tables that the biases of the WG and GMM estimators increase as α increases, for any composition of the variance of y, and any time series dimension. For example, if we look at table 1, for T=5 the absolute bias

---

[10] We have generated 100 samples for each experiment. We have also performed some simulations with 1000 replications and the results do not change (see appendix D).

of the WG estimator is 0.28 for $\alpha=0.1$, 0.44 for $\alpha=0.5$ and 0.75 for $\alpha=0.9$, and for the GMM, the bias is 0.03 for $\alpha=0.1$, 0.11 for $\alpha=0.5$, and 0.53 for $\alpha=0.9$. Furthermore, the sizes of the biases are quite close to the asymptotic biases (compare the results on tables 1-4 with the correspondent figures in section 4).

With regard to the influence of the composition of the variance of y on the behaviour of the non-corrected estimators, the conclusion from the tables is that the biases of the WG and GMM estimators are smaller the bigger is the ratio of the time varying component of the between cohort variance to the within cohort variance. The behaviour of these estimators depends also on the time series dimension of the sample. The absolute value of the bias decreases as T increases for both estimators but the influence is stronger for the WG.

If we compare the WG and GMM estimators, we can see that in general the GMM estimator leads to better results than the WG. Only when T and α are big and the proportion of the within cohort variance is high, the bias of the WG is smaller than the corresponding GMM (see table 1, $\alpha=0.9$, $T=15$). This result also coincides with the asymptotic behaviour (see fig.3 in section 4).

If we look at the results for the measurement error corrected estimators, we can see that they lead to less biased estimates than the corresponding non-corrected estimators, with a small increase, in most of the cases, of the standard deviation. Notice that the WGCT estimator is not consistent as C tends to infinity and it leads to worse results than the

GMMC estimator specially for small values of T.

We finally comment on the performance of the WGC estimator. As we said earlier, the measurement error correction used to construct this estimator is the appropriate one in order to eliminate measurement error bias as T tends to infinity. Given that we are considering finite values of T this estimator leads to worse results than the WGCT estimator and it can lead to results even worse than the non-corrected WG for small values of T and α. This result coincides with the asymptotic behaviour (see fig.3).

Table 1
Mean and Standard Deviation of the Estimators

Within cohort variance 50%
Between cohorts variance (time invariant component) 25%
Between cohorts variance (time varying component) 25%

	WG	WGC	WGCT	GMM	GMMC
			$\alpha = 0.1$		
T=5					
mean	-0.1844	-0.2059	-0.1647	0.0730	0.0766
st. dev.	0.0629	0.0702	0.0682	0.1119	0.1211
T=7					
mean	-0.0866	-0.0954	-0.0744	0.0793	0.0844
st. dev.	0.0468	0.0517	0.0505	0.0742	0.0804
T=10					
mean	-0.0255	-0.0279	-0.0161	0.0825	0.0877
st. dev.	0.0381	0.0417	0.0413	0.0536	0.0580
T=15					
mean	0.0136	0.0147	0.0213	0.0817	0.0871
st. dev.	0.0333	0.0362	0.0359	0.0448	0.0485
			$\alpha = 0.5$		
T=5					
mean	0.0557	0.0651	0.1141	0.3917	0.4346
st. dev.	0.0683	0.0798	0.0758	0.1468	0.1619
T=7					
mean	0.1903	0.2156	0.2365	0.4141	0.4600
st. dev.	0.0488	0.0543	0.0524	0.0948	0.1043
T=10					
mean	0.2814	0.3133	0.3233	0.4231	0.4670
st. dev.	0.0384	0.0417	0.0409	0.0611	0.0658
T=15					
mean	0.3449	0.3793	0.3842	0.4290	0.4731
st. dev.	0.0317	0.0338	0.0334	0.0485	0.0521
			$\alpha = 0.9$		
T=5					
mean	0.1534	0.2792	0.4453	0.3694	0.5205
st. dev.	0.0736	0.1313	0.1133	0.2492	0.3647
T=7					
mean	0.3231	0.4822	0.5347	0.4199	0.5802
st. dev.	0.0539	0.0732	0.0666	0.1633	0.2215
T=10					
mean	0.4645	0.6198	0.6376	0.4979	0.6869
st. dev.	0.0426	0.0468	0.0441	0.1062	0.1290
T=15					
mean	0.5821	0.7185	0.7244	0.5607	0.7684
st. dev.	0.0296	0.0293	0.0286	0.0790	0.0864

Table 2
Mean and Standard Deviation of the Estimators

Within cohort variance 40%
Between cohorts variance (time invariant component) 20%
Between cohorts variance (time varying component) 40%

	WG	WGC	WGCT	GMM	GMMC
			$\alpha = 0.1$		
T=5					
mean	-0.1814	-0.1918	-0.1715	0.0792	0.0815
st. dev.	0.0631	0.0666	0.0657	0.1040	0.1083
T=7					
mean	-0.0839	-0.0881	-0.0777	0.0864	0.0894
st. dev.	0.0469	0.0493	0.0488	0.0708	0.0738
T=10					
mean	-0.0221	-0.0231	-0.0173	0.0888	0.0917
st. dev.	0.0384	0.0402	0.0400	0.0510	0.0531
T=15					
mean	0.0168	0.0176	0.0209	0.0872	0.0902
st. dev.	0.0335	0.0349	0.0348	0.0438	0.0456
			$\alpha = 0.5$		
T=5					
mean	0.0732	0.0793	0.1034	0.4265	0.4501
st. dev.	0.0683	0.0739	0.0720	0.1338	0.1400
T=7					
mean	0.2078	0.2215	0.2318	0.4479	0.4724
st. dev.	0.0484	0.0511	0.0501	0.0884	0.0926
T=10					
mean	0.3000	0.3169	0.3218	0.4534	0.4768
st. dev.	0.0383	0.0399	0.0395	0.0570	0.0591
T=15					
mean	0.3632	0.3812	0.3837	0.4563	0.4797
st. dev.	0.0313	0.0323	0.0321	0.0461	0.0477
			$\alpha = 0.9$		
T=5					
mean	0.2330	0.3185	0.3975	0.5372	0.6506
st. dev.	0.0734	0.0979	0.0895	0.2588	0.3106
T=7					
mean	0.3999	0.4944	0.5201	0.5867	0.7054
st. dev.	0.0526	0.0605	0.0572	0.1525	0.1648
T=10					
mean	0.5360	0.6240	0.6328	0.6533	0.7729
st. dev.	0.0401	0.0406	0.0393	0.0930	0.0952
T=15					
mean	0.6449	0.7199	0.7228	0.6964	0.8195
st. dev.	0.0273	0.0263	0.0260	0.0651	0.0635

Table 3
Mean and Standard Deviation of the Estimators

Within cohort variance 29%
Between cohorts variance (time invariant component) 14%
Between cohorts variance (time varying component) 57%

	WG	WGC	WGCT	GMM	GMMC
			α = 0.1		
T=5					
mean	-0.1798	-0.1849	-0.1748	0.0824	0.0837
st. dev.	0.0631	0.0649	0.0645	0.0979	0.0999
T=7					
mean	-0.0825	-0.0846	-0.0794	0.0905	0.0921
st. dev.	0.0470	0.0482	0.0479	0.0685	0.0699
T=10					
mean	-0.0203	-0.0207	-0.0178	0.0926	0.0941
st. dev.	0.0387	0.0396	0.0395	0.0496	0.0506
T=15					
mean	0.0186	0.0190	0.0206	0.0905	0.0921
st. dev.	0.0335	0.0342	0.0342	0.0432	0.0441
			α = 0.5		
T=5					
mean	0.0828	0.0861	0.0981	0.4446	0.4569
st. dev.	0.0682	0.0710	0.0701	0.1248	0.1276
T=7					
mean	0.2171	0.2242	0.2294	0.4654	0.4780
st. dev.	0.0482	0.0495	0.0490	0.0841	0.0860
T=10					
mean	0.3099	0.3186	0.3211	0.4697	0.4817
st. dev.	0.0383	0.0391	0.0389	0.0547	0.0557
T=15					
mean	0.3730	0.3823	0.3835	0.4712	0.4832
st. dev.	0.0310	0.0315	0.0314	0.0447	0.0454
			α = 0.9		
T=5					
mean	0.2857	0.3355	0.3740	0.6477	0.7110
st. dev.	0.0724	0.0834	0.0793	0.2471	0.2674
T=7					
mean	0.4482	0.5001	0.5129	0.6962	0.7631
st. dev.	0.0514	0.0547	0.0531	0.1372	0.1376
T=10					
mean	0.5790	0.6260	0.6304	0.7462	0.8096
st. dev.	0.0382	0.0380	0.0373	0.0825	0.0822
T=15					
mean	0.6812	0.7206	0.7220	0.7766	0.8413
st. dev.	0.0258	0.0251	0.0250	0.0567	0.0549

Table 4
Mean and Standard Deviation of the Estimators

Within cohort variance 18%
Between cohorts variance (time invariant component) 9%
Between cohorts variance (time varying component) 73%

	WG	WGC	WGCT	GMM	GMMC
			α = 0.1		
T=5					
mean	-0.1789	-0.1814	-0.1764	0.0838	0.0844
st. dev.	0.0631	0.0640	0.0638	0.0942	0.0951
T=7					
mean	-0.0818	-0.0829	-0.0803	0.0927	0.0935
st. dev.	0.0470	0.0476	0.0475	0.0671	0.0678
T=10					
mean	-0.0193	-0.0195	-0.0181	0.0947	0.0956
st. dev.	0.0389	0.0393	0.0393	0.0490	0.0495
T=15					
mean	0.0195	0.0197	0.0206	0.0925	0.0933
st. dev.	0.0335	0.0338	0.0338	0.0428	0.0433
			α = 0.5		
T=5					
mean	0.0877	0.0895	0.0955	0.4535	0.4597
st. dev.	0.0681	0.0695	0.0691	0.1196	0.1209
T=7					
mean	0.2218	0.2254	0.2280	0.4741	0.4804
st. dev.	0.0481	0.0487	0.0485	0.0815	0.0824
T=10					
mean	0.3151	0.3195	0.3208	0.4783	0.4844
st. dev.	0.0384	0.0388	0.0387	0.0536	0.0541
T=15					
mean	0.3781	0.3828	0.3834	0.4792	0.4853
st. dev.	0.0308	0.0310	0.0310	0.0439	0.0442
			α = 0.9		
T=5					
mean	0.3164	0.3432	0.3623	0.7061	0.7381
st. dev.	0.0714	0.0766	0.0745	0.2328	0.2412
T=7					
mean	0.4756	0.5029	0.5092	0.7549	0.7890
st. dev.	0.0505	0.0520	0.0512	0.1279	0.1269
T=10					
mean	0.6027	0.6270	0.6292	0.7940	0.8258
st. dev.	0.0371	0.0368	0.0365	0.0769	0.0767
T=15					
mean	0.7007	0.7210	0.7217	0.8187	0.8514
st. dev.	0.0250	0.0247	0.0246	0.0527	0.0516

NOTES TO TABLES

(i) 100 simulations.

(ii) Cross-sections of 2000 individuals.

(iii) AR(1) model with individual effects.

(iv) WG: within groups estimator. WGC: measurement error corrected WG (appropriate correction for $T \to \infty$). WGCT measurement error corrected WG (appropriate correction for finite T).

(v) GMM: Generalized method of moments estimator.

GMMC: measurement errors corrected GMM. Matrix of instruments used

$$Z_c = \begin{bmatrix} y_{c1} & & & & & \\ & y_{c1} & y_{c2} & & & \\ & & & \ddots & & \\ & & & y_{cT-4} & y_{cT-3} & y_{cT-2} \end{bmatrix}$$

6.- CONCLUSIONS

The problem analyzed in this chapter is how we can estimate dynamic models using time series of cross-section data. As it happens when we work with panel data, we can consider different type of estimators which are consistent for different types of asymptotics (cross-section or time series). We propose a measurement error corrected within groups estimator which is consistent as the number of time periods tends to infinity, and a measurement error corrected GMM estimator which is consistent as the number of cohorts tends to infinity. We also calculate the asymptotic biases of the non-corrected estimators and we analyze the size of the biases depending on the parameters of the model.

In the last section of the chapter we have carried out Monte Carlo simulations to study the small sample properties of the estimators proposed. The conclusions derived from the simulations reinforce the asymptotic results. The measurement error correction seems to be important, and the corrected estimators lead to less biased results. Furthermore, for small values of T, GMM estimators are better than within groups.

APPENDIX A. Asymptotic Distribution of the GMM Estimator when the Covariance Matrix of the Measurement Errors is Unknown.

Let $\hat{\Lambda}$ and $\hat{\lambda}$ be consistent estimators of Λ and λ respectively. The criterion function we have to minimize is in this case

$$s(\delta) = \frac{1}{C}\sum_{c=1}^{C}(Z_c'\Delta u_c-\hat{\Lambda}\delta-\hat{\lambda})'A_c\frac{1}{C}\sum_{c=1}^{C}(Z_c'\Delta u_c-\hat{\Lambda}\delta-\hat{\lambda})$$

Let

$$b_c(\delta) = \frac{1}{C}\sum_{c=1}^{C}(Z_c'\Delta u_c-\hat{\Lambda}\delta-\hat{\lambda}) = \frac{1}{C}\sum_{c=1}^{C}(Z_c'\Delta u_c-\Lambda\delta-\lambda)-[(\hat{\Lambda}\delta+\hat{\lambda})-(\Lambda\delta+\lambda)]$$

The elements of $\hat{\Lambda}$ and $\hat{\lambda}$ are either zero or else estimates of the variances and covariances of the measurement errors, and given that we are assuming that they are constant across cohorts and over time the elements of $\hat{\Lambda}$ and $\hat{\lambda}$ are calculated as follows

$$\hat{\sigma}_\zeta^2 = \frac{1}{C}\sum_{c=1}^{C}\frac{1}{(nc-1)T}\sum_{t=1}^{T}\sum_{i\in c,t}(y_{it}-y_{ct})^2 = \frac{1}{C}\sum_{c=1}^{C}\hat{\sigma}_\zeta^c$$

Analogously we construct $\hat{\sigma}_{\zeta\eta}$ and $\hat{\Sigma}_\eta$.

Using the definition of the vec operator and the Kronecker product

$$(\hat{\Lambda}\delta+\hat{\lambda})-(\Lambda\delta+\lambda) = (\delta'\otimes I)(\mathrm{vec}\hat{\Lambda}-\mathrm{vec}\Lambda)+(\hat{\lambda}-\lambda)$$

and it can be written as

$$(\hat{\Lambda}\delta+\hat{\lambda})-(\Lambda\delta+\lambda) = (\delta'\otimes I{:}I)\frac{1}{C}\sum_{c=1}^{C}\begin{pmatrix}\mathrm{vec}\hat{\Lambda}_c-\mathrm{vec}\Lambda\\ \hat{\lambda}_c-\lambda\end{pmatrix}$$

where $\hat{\Lambda}_c$ and $\hat{\lambda}_c$ are obtained using the individuals in cohort c.

Hence

$$b_c(\delta) = (I{:}-\delta'\otimes I{:}-I)\frac{1}{C}\sum_{c=1}^{C}\begin{pmatrix}Z_c'\Delta u_c-\Lambda\delta-\lambda\\ \mathrm{vec}\hat{\Lambda}_c-\mathrm{vec}\Lambda\\ \hat{\lambda}_c-\lambda\end{pmatrix} = H\frac{1}{C}\sum_{c=1}^{C}(\hat{\psi}_c-\psi)$$

The ψ_c's are iid with mean ψ and covariance matrix $V_\psi=E(\psi_c-\psi)(\psi_c-\psi)'$.

Then, using a CLT

$$\sqrt{C}\, b_c(\delta) \to_d N(0,V_0) \qquad \text{where} \qquad V_0 = HV_\psi H'$$

The asymptotic distribution of the GMMC estimator is given by

$$\sqrt{C}(\tilde{\delta}-\delta)\to_d N(0,(D_0'A_0D_0)^{-1}D_0'A_0V_0A_0D_0(D_0'A_0D_0)^{-1})$$

where

$$A_0 = \mathrm{plim}\, A_c$$
$$D_0 = \mathrm{plim}\frac{\partial}{\partial\delta'}\sum_c(Z_c'\Delta u_c-\hat{\Lambda}\delta-\hat{\lambda}) = -(E(Z_c'\Delta W_c)+\Lambda)$$

A consistent estimator of the asymptotic covariance matrix is obtained by replacing A_0 by A_c and D_0, V_0 by

$$\hat{D} = -\frac{1}{C}\sum_{c=1}^{C}Z_c'\Delta W_c-\hat{\Lambda}$$
$$\hat{V} = \hat{H}\hat{V}_\psi\hat{H}'$$

where

$$\hat{H} = (I:-\tilde{\delta}'\otimes I:-I)$$
$$\hat{V}_{\psi} = \frac{1}{C}\sum_{c=1}^{C}\hat{\psi}_c\hat{\psi}_c' - \left(\frac{1}{C}\sum_{c=1}^{C}\hat{\psi}_c\right)\left(\frac{1}{C}\sum_{c=1}^{C}\hat{\psi}_c'\right)$$

Δu_c has to be replaced by $\Delta\tilde{u}_c$ the vector of residuals.

APPENDIX B. Asymptotic Bias of the Within Groups Estimator.

Consider the within groups estimator obtained using the t-the cross-section

$$\hat{\alpha}_{wgt} = \frac{\frac{1}{C}\sum_{c=1}^{C}\tilde{y}_{ct-1}\tilde{y}_{ct}}{\frac{1}{C}\sum_{c=1}^{C}\tilde{y}_{ct-1}^2} = \alpha + \frac{\frac{1}{C}\sum_{c=1}^{C}\tilde{y}_{ct-1}\tilde{u}_{ct}}{\frac{1}{C}\sum_{c=1}^{C}\tilde{y}_{ct-1}^2}$$

and

$$\text{plim}_{C\to\infty}(\hat{\alpha}_{wgt}-\alpha) = \frac{\text{plim}\frac{1}{C}\sum_{c=1}^{C}\tilde{y}_{ct-1}\tilde{u}_{ct}}{\text{plim}\frac{1}{C}\sum_{c=1}^{C}\tilde{y}_{ct-1}^2} = \frac{A_t}{B_t}$$

We are going to calculate A_t and B_t

$$A_t = E\,[\tilde{y}_{ct-1}\tilde{u}_{ct}] = E\,[\tilde{y}_{ct-1}^{*}\tilde{u}_{ct}] + E\,[\tilde{\zeta}_{ct-1}\tilde{u}_{ct}]$$

The second term can be obtained as follows

$$E\,[\tilde{\zeta}_{ct-1}\tilde{u}_{ct}] = E\,[\tilde{\zeta}_{ct-1}(\tilde{v}_{ct}^{*}+\tilde{\zeta}_{ct}-\alpha\tilde{\zeta}_{ct-1})] =$$
$$= E(\tilde{\zeta}_{ct-1}\tilde{\zeta}_{ct}) - \alpha E(\tilde{\zeta}_{ct-1}^2) = -\frac{T}{(T-1)^2}\frac{\sigma_\zeta^2}{nc} - \alpha\frac{T-2}{T-1}\frac{\sigma_\zeta^2}{nc}$$

Following Nickell (1981), we can calculate the first term

$$E(\tilde{y}^*_{ct-1}\tilde{u}_{ct}) = E(\tilde{y}^*_{ct-1}\tilde{v}^*_{ct}) = -\frac{\sigma^2_{v^*}}{(T-1)(1-\alpha)}\left[1-\alpha^{t-2}-\alpha^{T-t}+\frac{1}{T-1}\frac{1-\alpha^{T-1}}{1-\alpha}\right]$$

and

$$A_t = -\frac{\sigma^2_{v^*}}{(T-1)(1-\alpha)}\left[1-\alpha^{t-2}-\alpha^{T-t}+\frac{1}{T-1}\frac{1-\alpha^{T-1}}{1-\alpha}\right]-\left[\frac{T}{(T-1)^2}+\frac{T-2}{T-1}\alpha\right]\frac{\sigma^2_\zeta}{nc}$$

analogously we can obtain B_t

$$B_t = E(\tilde{y}^2_{ct-1}) = E(\tilde{y}^{*2}_{ct-1})+E(\tilde{\zeta}^2_{ct-1}) =$$
$$= \frac{\sigma^2_{v^*}}{1-\alpha^2}\frac{T-2}{T-1}-\frac{2\alpha}{1-\alpha^2}\frac{\sigma^2_{v^*}}{(T-1)(1-\alpha)}\left[1-\alpha^{t-2}-\alpha^{T-t}+\frac{1}{T-1}\frac{1-\alpha^{T-1}}{1-\alpha}\right]+\frac{T-2}{T-1}\frac{\sigma^2_\zeta}{nc}$$

using the whole sample

$$\hat{\alpha}_{wg} = \frac{\frac{1}{C(T-1)}\sum_{t=2}^{T}\sum_{c=1}^{C}\tilde{y}_{ct-1}\tilde{y}_{ct}}{\frac{1}{C(T-1)}\sum_{t=2}^{T}\sum_{c=1}^{C}\tilde{y}^2_{ct-1}} = \alpha + \frac{\frac{1}{C(T-1)}\sum_{t=2}^{T}\sum_{c=1}^{C}\tilde{y}_{ct-1}\tilde{u}_{ct}}{\frac{1}{C(T-1)}\sum_{t=2}^{T}\sum_{c=1}^{C}\tilde{y}^2_{ct-1}}$$

and

$$\text{plim}_{\,T\to\infty}(\hat{\alpha}_{wg}-\alpha) = \frac{\frac{1}{T-1}\sum_{t=2}^{T}A_t}{\frac{1}{T-1}\sum_{t=2}^{T}B_t}$$

adding up the expressions for A_t and B_t and substituting we obtain

$$\text{plim}_{C\to\infty}(\hat{\alpha}_{wg}-\alpha)=$$

$$= -\frac{\frac{1}{(T-1)(1-\alpha)}\left[1-\frac{1}{T-1}\frac{1-\alpha^{T-1}}{1-\alpha}\right]\frac{nc\sigma^2_{v^*}}{\sigma^2_\zeta}+\left[\frac{T}{(T-1)^2}+\frac{T-2}{T-1}\alpha\right]}{\frac{1}{1-\alpha^2}\left[1-\frac{1}{T-1}\right]\frac{nc\sigma^2_{v^*}}{\sigma^2_\zeta}-\frac{2\alpha}{1-\alpha^2}\frac{1}{(T-1)(1-\alpha)}\left[1-\frac{1}{T-1}\frac{1-\alpha^{T-1}}{1-\alpha}\right]\frac{nc\sigma^2_{v^*}}{\sigma^2_\zeta}+\frac{T-2}{T-1}}$$

rearranging terms and defining

$$h_T = \frac{1}{1-\alpha}\left[1-\frac{1}{T-1}\frac{1-\alpha^{T-1}}{1-\alpha}\right] \quad \text{and} \quad \phi=\frac{nc\sigma^2_{v^*}}{\sigma^2_\zeta}$$

we obtain

$$\text{plim}_{C\to\infty}(\hat{\alpha}_{wg}-\alpha)= -\frac{h_T\phi+\left[\frac{T}{T-1}+(T-2)\alpha\right]}{\frac{T-2}{1-\alpha^2}\phi-\frac{2\alpha}{1-\alpha^2}h_T\phi+T-2}$$

Analogously we obtain the asymptotic bias of the measurement error corrected within groups estimators (WGC and WGCT).

APPENDIX C. Asymptotic bias of the GMM Estimator.

The non-corrected GMM estimator for the pure AR(1) model with individual effects is given by

$$\tilde{\alpha}=\frac{\left(\sum_{c=1}^{C} Z_c'\Delta y_{c(-1)}\right)' A_c \sum_{c=1}^{C} Z_c'\Delta y_c}{\left(\sum_{c=1}^{C} Z_c'\Delta y_{c(-1)}\right)' A_c \sum_{c=1}^{C} Z_c'\Delta y_{c(-1)}} = \alpha+\frac{\left(\sum_{c=1}^{C} Z_c'\Delta y_{c(-1)}\right)' A_c \sum_{c=1}^{C} Z_c'\Delta u_c}{\left(\sum_{c=1}^{C} Z_c'\Delta y_{c(-1)}\right)' A_c \sum_{c=1}^{C} Z_c'\Delta y_{c(-1)}}$$

$$\text{plim}_{C\to\infty}(\tilde{\alpha}-\alpha) = \frac{\text{plim}\ \frac{1}{C}\sum_{c=1}^{C}\Delta y_{c(-1)}Z_c\ \text{plim}\ A_c\ \text{plim}\ \frac{1}{C}\sum_{c=1}^{C}Z_c'\Delta u_c}{\text{plim}\ \frac{1}{C}\sum_{c=1}^{C}\Delta y_{c(-1)}Z_c\ \text{plim}\ A_c\ \text{plim}\ \frac{1}{C}\sum_{c=1}^{C}Z_c'\Delta y_{c(-1)}}$$

Let us now calculate the different probability limits in the expression above

$$\text{plim}\ \frac{1}{C}\sum_{c=1}^{C}Z_c'\Delta y_{c(-1)} = E(Z_c'\Delta y_{c(-1)}) = E(Z_c^{*\prime}\Delta y_{c(-1)}^{*}) + E\begin{bmatrix}\zeta_{c1} & & & \\ & \zeta_{c1} & & \\ & \zeta_{c2} & & \\ & & \cdot & \\ & & & \cdot \\ & & & \zeta_{c1} \\ & & & \cdot \\ & & & \cdot \\ & & & \zeta_{cT-2}\end{bmatrix}\begin{bmatrix}\Delta\zeta_{c2} \\ \cdot \\ \cdot \\ \Delta\zeta_{c(T-1)}\end{bmatrix} =$$

$$= E(Z_c^{*\prime}\Delta y_{c(-1)}^{*}) - \mu\frac{\sigma_\zeta^2}{nc}$$

$$\text{plim}\ \frac{1}{C}\sum_{c=1}^{C}Z_c'\Delta u_c = E(Z_c'\Delta u_c) = \alpha\frac{\sigma_\zeta^2}{nc}\mu$$

$$\text{plim}\ A_c = A_0$$

where μ = (1,0,1,0,0,1,..).

assuming that the process is stationary

$$y_{ct}^{*} = \frac{1}{1-\alpha}\theta_c^{*} + \sum_{j=0}^{\infty}\alpha^j v_{ct-j}^{*}$$

$$\Delta y_{ct}^{*} = \sum_{j=0}^{\infty}\alpha^j\Delta v_{ct-j}^{*}$$

$$E\left[Z_c^{*\prime}\Delta y_{c(-1)}^*\right] = E\begin{bmatrix} y_{c1}^* & & & \\ & y_{c1}^* & & \\ & y_{c2}^* & & \\ & & \cdot & \\ & & \cdot & \\ & & & y_{c1}^* \\ & & & \cdot \\ & & & \cdot \\ & & & y_{c(T-2)} \end{bmatrix}\begin{bmatrix} \Delta y_{c2}^* \\ \cdot \\ \cdot \\ \Delta y_{cT-1}^* \end{bmatrix} = -\begin{bmatrix} 1 \\ -- \\ \alpha \\ 1 \\ -- \\ \alpha^2 \\ \alpha \\ 1 \\ -- \\ \cdot \\ \cdot \end{bmatrix}\frac{\sigma_{v^*}^2}{1+\alpha} = -\nu\frac{\sigma_{v^*}^2}{1+\alpha}$$

Then

$$\operatorname{plim}_{C\to\infty}\frac{1}{C}\sum_{c=1}^{C} Z_c'\Delta y_{c(-1)} = -\nu\frac{\sigma_{v^*}^2}{1+\alpha} - \mu\frac{\sigma_\zeta^2}{nc}$$

and the asymptotic bias is

$$\operatorname{plim}_{c\to\infty}(\tilde{\alpha}-\alpha) = -\frac{\left[\nu\frac{1}{1+\alpha}\phi+\mu\right]' A_0\mu\alpha}{\left[\nu\frac{1}{1+\alpha}\phi+\mu\right]' A_0\left[\nu\frac{1}{1+\alpha}\phi+\mu\right]}$$

APPENDIX D. Monte Carlo simulations (1000 replications)

Table D
Mean and Standard Deviation of the Estimators

Within cohort variance 50%
Between cohorts variance (time invariant component) 25%
Between cohorts variance (time varying component) 25%

	WG	WGC	WGCT	GMM	GMMC
			α = 0.1		
T=5					
mean	-0.1858	-0.2076	-0.1661	0.0680	0.0718
st. dev.	0.0584	0.0655	0.0632	0.1070	0.1161
T=7					
mean	-0.0936	-0.1031	-0.0820	0.0723	0.0766
st. dev.	0.0461	0.0509	0.0499	0.0752	0.0813
T=10					
mean	-0.0303	-0.0331	-0.0214	0.0803	0.0854
st. dev.	0.0380	0.0416	0.0411	0.0552	0.0598
T=15					
mean	0.0150	0.0163	0.0229	0.0823	0.0877
st. dev.	0.0311	0.0339	0.0336	0.0422	0.0456
			α = 0.5		
T=5					
mean	0.0507	0.0592	0.1090	0.3873	0.4329
st. dev.	0.0644	0.0755	0.0712	0.1455	0.1648
T=7					
mean	0.1843	0.2089	0.2301	0.4057	0.4506
st. dev.	0.0481	0.0539	0.0521	0.0921	0.1016
T=10					
mean	0.2790	0.3105	0.3205	0.4249	0.4699
st. dev.	0.0395	0.0430	0.0421	0.0656	0.0712
T=15					
mean	0.3479	0.3827	0.3876	0.4323	0.4769
st. dev.	0.0297	0.0317	0.0313	0.0463	0.0497
			α = 0.9		
T=5					
mean	0.1425	0.2583	0.4294	0.3453	0.4963
st. dev.	0.0698	0.1239	0.1018	0.2391	0.3762
T=7					
mean	0.3207	0.4789	0.5316	0.4220	0.5989
st. dev.	0.0513	0.0704	0.0641	0.1601	0.2298
T=10					
mean	0.4637	0.6176	0.6353	0.5078	0.7077
st. dev.	0.0421	0.0475	0.0451	0.1159	0.1427
T=15					
mean	0.5853	0.7215	0.7272	0.5714	0.7819
st. dev.	0.0308	0.0302	0.0294	0.0778	0.0827

CHAPTER 3. SEPARABILITY AND AGGREGATE SHOCKS IN THE LIFE-CYCLE MODEL OF CONSUMPTION: EVIDENCE FOR SPAIN

1.- INTRODUCTION

Since 1978, when Hall published his paper on the life-cycle permanent income hypothesis, many authors have estimated revised versions of his model and have tested its implications using both aggregate and micro data.

At a macro level most of the empirical tests lead to the rejection of the permanent income hypothesis (e.g. Flavin (1981), Campbell and Mankiw (1991)). However, the fact that the life-cycle model of consumption is in general rejected using aggregate time series data does not necessarily invalidate the theory at the individual level. As pointed out by previous authors, the failure of the model with macro data can be due to the violation of the aggregation assumptions needed to justify the use of aggregate data (see Ch. 5 Deaton (1992), Attanasio and Weber (1993)).

At a micro level there is evidence in favour and against the permanent income hypothesis. Most of the research in this area (e.g. Hall and Mishkin (1982), Zeldes (1989) and Runkle (1991)) is based on the Panel Study of Income Dynamics (PSID). This data set only includes information on food consumption, and therefore, preferences have to be parameterized such that the Euler equation for food consumption does not depend on consumption of other goods[11]. As argued in Attanasio and Weber (1992), the violation of this assumption can be responsible for the rejection of the permanent income hypothesis when the data set used is the PSID. The reason why we can spuriously find evidence of excess sensitivity of consumption to income, when the consumption measure used is food consumption, is the following: if the utility function is not additive in food and non-food consumption, the Euler equation for food will depend on consumption of other goods. Hence, if no measure of non-food consumption is included, a spurious dependence of food consumption on income can be induced. This spurious dependence would lead to a rejection of the permanent income hypothesis. In this chapter, we consider groups of composite commodities, and we estimate the Euler equations derived from a life-cycle model of consumer behaviour.

The main data set that we use is a rotating panel from the Spanish family expenditure survey (Encuesta Continua de Presupuestos Familiares) corresponding to 1985-89. This data set has several advantages, in order to estimate the life-cycle model of consumption, compared with other data sets used in the literature. On the one hand, in this survey, very detailed information on expenditures is recorded. This fact makes this survey more appealing than the PSID. On the other hand, the structure of the Spanish survey is more convenient than the most widely used consumer surveys.

[11] The normalization of the utility function has to be chosen such that the utility function is additive separable between food and non-food consumption. Notice that within period allocation of expenditures is invariant to monotonic transformations of the utility function, but intertemporal allocation of consumption depends on the normalization of the utility function.

In the Spanish survey households are interviewed during eight consecutive quarters and a complete information on expenditure, income and family characteristics is recorded. The consumer surveys most widely used do not have this panel structure. The British Family Expenditure Survey has independent waves, and in the American Consumer Expenditure Survey, even though households are interviewed in four consecutive quarters, the information on income is only recorded in the 1st and 4th interview. The frequency of the data, quarterly as opposed to annual, is another advantage of the spanish survey relative to the PSID for the purpose of studying consumption decisions.

Another important issue that has recently attracted attention in the literature is the presence of aggregate shocks that could invalidate the instruments and hence the identification of the model when the time series dimension of the data set is small (see Deaton (1992)). If the aggregate shocks affect all the individuals in the same way the problem can be easily solved by introducing time dummies in the model. However, if the effect of the shocks is not the same for everybody, for example some people can obtain higher benefits in a recession period, the introduction of time dummies will not solve the problem and we will need a long time series dimension to obtain valid estimates of the model. Therefore, if the effect of aggregate shocks varies over individuals, we may obtain different estimates for the parameters of the model for different periods of time, even if we include time dummies to pick up these effects. We have used a second unbalanced panel from a previous series of consumer surveys for

Spain (Encuesta Permanente de Consumo) which were carried out between 1978 and 1983, at a very different part of the cycle, relative to the period 1985-89 when the economy was booming. On the basis of these two data sets we can check the stability of our results.

This chapter is organized as follows. In section 2 we present the life-cycle model of consumer behaviour that will be used in the chapter. In section 3 we describe the information contained in the Spanish family expenditure surveys and how the variables of the model have been constructed. In section 4 we analyze some econometric issues on the estimation of the model. The results are presented in section 5. Section 6 concludes.

2.- THE MODEL

The decision problem faced by the consumer is how to allocate consumption over time to maximize the expected intertemporal utility, i.e.

$$\begin{aligned} &\max E_t\left[\sum_{k=0}^{T-t} \delta^k u(c_{t+k},\eta_{t+k})\right] \\ &\text{subject to} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2.1) \\ &A_{t+1+k}=(1+i_{t+k})(A_{t+k}+y_{t+k}-p'_{t+k}c_{t+k}) \qquad k=0,..,T-t \\ &A_{T+1}\geq 0 \end{aligned}$$

Where c_s is a vector of consumption of n groups of commodities in period s, y_s is income, i_s is the nominal interest rate and p_s is a vector of prices. A_s are assets at the beginning of period s, η_s is a vector of family characteristics and δ is the discount rate. E_t is the conditional expectation operator, conditional on information known by the consumer in period t.

The set of Euler equations for this problem is

$$E_t\left[\frac{(1+r_{jt})\delta\, u_j(c_{t+1},\eta_{t+1})}{u_j(c_t,\eta_t)}\right]=1 \qquad j=1,\ldots,n \tag{2.2}$$

Where r_{jt} is the commodity specific real interest rate $(1+r_{jt}) = (1+i_t)p_{jt}/p_{jt+1}$ and u_j is the partial derivative of the utility function with respect to consumption of commodity j. We can write (2.2) in terms of the actual values as

$$\frac{(1+r_{jt})\delta\, u_j(c_{t+1},\eta_{t+1})}{u_j(c_t,\eta_t)}=1+\epsilon_{jt+1} \quad E_t(\epsilon_{jt+1})=0 \qquad j=1,\ldots,n \tag{2.3}$$

Consider the following instantaneous utility function, which is not additive (given the normalization we use) but simple to guarantee an approximate log-linear Euler equation

$$u(c_{1t},\ldots,c_{nt},\eta_t)=c_{1t}^{\alpha_1}..c_{nt}^{\alpha_n}\phi(\eta_t) \tag{2.4}$$

Where ϕ is a function of the vector of family characteristics, which will be parameterized as an exponential[12]. We can write the set of Euler equations in (2.3) for the utility function (2.4). Taking logarithms and using a second order taylor approximation for $\log(1+\epsilon_{jt+1})$, we obtain

[12] Notice that even though this utility function is weakly separable, the normalization we use implies non-separability in the Euler equations among consumption goods.

$$(\alpha_j - 1)\Delta\log c_{jt+1} + \sum_{k \neq j} \alpha_k \Delta\log c_{kt+1} + \log\delta + \frac{1}{2}\sigma_j^2 + \log(1+r_{jt}) + \Delta\log\phi(\eta_t) = \epsilon_{jt+1} \tag{2.5}$$

where $E_t(\epsilon_{jt+1}) = 0$, $\sigma^2 = E(\epsilon^2_{jt+1})$ and Δ is the first differences operator[13]. All the variables except the interest rate are household specific, but we have omitted the household subscript to simplify notation.

We are implicitly assuming additive separability between consumption of durable and non-durable goods in the utility function. The reasons why we do not include expenditure in durables in our model are first, the additional econometric problems involved in the treatment of durables (infrequency of purchases); and secondly, the fact that the inclusion of durables would complicate the specification of preferences (see Hayashi (1985)). On the other hand, we are mainly concerned in testing the life-cycle permanent income hypothesis, rather than in modelling consumption patterns for different goods.

Another assumption in this model is separability between consumption and leisure. To our knowledge, there is not much formal evidence about this issue. If we look at the results in Browning, Deaton and Irish (1985), the evidence about non-separability between consumption and leisure is not very reliable; the reason is that, even though the cross-price effects are significant, their signs are contradictory. We have tried to

[13] This approximation has been criticized by Altug and Miller (1990). They argue that $\log(1+\epsilon_{jt+1})$ is correlated with past information invalidating the instruments widely used to estimate this kind of models.

overcome potential shortcomings by including dummy variables for labour market status of the household head and the wife as additional regressors in our equation. These dummies will pick up to some extent the potential differences in consumption behaviour among households with different labour force participation status.

3.- THE DATA

The first data set that we have used is the Spanish family expenditure survey (Encuesta Continua de Presupuestos Familiares (ECPF)). This survey is carried out by personal interview on a quarterly basis, from the 1st quarter 1985. The survey contains very detailed information on family expenditures, information on household characteristics and family income. In this application we have used 20 quarters of the survey, from 1st quarter 1985 to 4th quarter 1989.

Every quarter, about 3000 families are interviewed. The data set is a rotating panel, since in principle on each quarter 1/8 of the households are renewed. A family stays in the sample at most eight periods but there is quite an important percentage of attrition in earlier quarters, mainly during the first two years of the survey. In this research we have considered families that report full information at least for five consecutive periods. The reason why we have dropped households with less than five responses is that we need lagged information to instrument the endogenous variables of the model.

Consumption patterns can be very different for households with different characteristics (family members, age, etc.). To overcome this problem, we can either work with a small sample of "homogeneous" consumers, or else we can assume that we know how preferences depend on family characteristics. Using the first approach will mean a reduction on the sample size and hence a worse performance of the estimators[14]. The problem with the second procedure is to combine flexibility with parsimony. Our approach will be a compromise between these two approaches.

Taking into account the considerations above, we keep in our sample married couples with or without children, such that the husband is coded head of the household. We drop households whose head is either very young (younger than 25), or else quite old (older than 65). We also condition on another demographic and labour force variables as explained below.

For the purpose of this research we have used only expenditure in non-durables and services which we have aggregated in three groups of commodities: the first one includes food, alcoholic and non-alcoholic drinks and tobacco; the second clothing and footwear; and the third energy and transport[15].

---

[14] Provided the sample selection is based on exogenous variables, otherwise we will have additional sample selection problems.

[15] Energy and transport is the group of non-durables whose definition is more homogeneous in the two data sets we use in this chapter. Unfortunately it does not include exactly the same expenses. In the ECPF, energy and transport includes car repairing and parking expenditures, while in the second survey these expenses are not included in this group.

The data set includes information on labour market status but not on hours of work for any member of the household. We have included as regressors dummy variables on labour market status for the household head and the wife. We have considered three dummies (full time employed, part time employed and unemployed) for each spouse. There is also information on the sex and age for each member of the household. We have assumed that preferences can also depend on demographics and we have included age and age squared of the household head, the number of babies (between 0-2 years old), children (3-17 years old), elderly people (older than 65), and family size.

The second data set is an unbalanced panel from a previous series of consumer surveys for Spain (Encuesta Permanente de Consumo (EPC)). This survey was carried out from the first quarter 1978 to the fourth quarter 1983. We observe some households for 24 quarters, however, on each period, part of the sample is renewed. Due to the reasons explained above, we keep families reporting full information for at least five consecutive quarters. The subsample we have considered was obtained using the same criteria that we used for the ECPF. The EPC does not contain information on income, and therefore we can not use this data set to test excess sensitivity of consumption growth to anticipated income growth. This survey does not provide any information on the labour market status of the wife. Therefore, when we use this data set we can only consider labour market dummies for the husband.

In the appendix we present descriptive statistics for demographic characteristics, expenditures and income, for the two data sets. If we compare the means or the medians of real expenditure on energy and transport (table A1), we can see that the figures are higher for 1985-89 (ECPF data), than for 1978-83 (EPC data). As we comment above, in the ECPF this group comprises some expenditures that are not contained in this group in the EPC. However, this fact does not explain completely the observed differences between these two periods. The figures for expenditure on clothing are also higher in 1985-89, even though this group includes the same expenses. We are aware of these differences which may cast some doubts on the comparison between the results obtained from the two surveys.

The price index for each group of commodities is derived from the disaggregated consumer retail price index for Spain published by the National Institute of Statistics (Instituto Nacional de Estadística), using the same weights that are used to construct the general index. The nominal interest rate is an interest rate on deposits provided by Cuenca (1991).

4.- ECONOMETRIC ISSUES

The set of Euler Equations in (2.5) is estimated using the Generalized Method of Moments (GMM). If the only component of the error terms in these equations were an expectational error, we could use as instruments for the model all the variables dated t-1 and earlier. However, if consumption is measured with error, additional terms are added to the disturbances and even assuming that these measurement errors are serially uncorrelated, we can only use as instruments endogenous variables dated t-2 and earlier. Another source of stochastic variability in the model are random preferences, i.e. individual heterogeneity is not perfectly observed. We can model this fact by adding an error term to the vector η_t in (2.5). This would add an extra component to the disturbances, and as it happens in the presence of measurement errors in consumption, random preferences can also invalidate the use of endogenous variables dated at t-1 as valid instruments for the model. We have tried several instrument sets, and a detailed explanation is provided in the next section.

As we discussed above, the presence of aggregate shocks will invalidate the econometric results based on cross-section averages. Therefore, we include time dummies in our regression equations, which will pick up the effect of the aggregate shocks, provided that their influence is similar across households. However, as we mention earlier, if the effect of aggregate shocks is different for different families, the estimated coefficients will be biased. We estimate the model using the ECPF and the EPC and we compare the results. In the presence of aggregate shocks we could expect to reject the stability of the coefficients, given that these shocks could bias the estimates in different ways for different periods.

The two data sets we have available are incomplete panels but they do not overlap. In order to have a longer time series dimension, we could join the information contained in the two samples by constructing cohorts of families according to the year of birth (see Browning, Deaton and Irish

(1985), Blundell, Browning and Meghir (1994) amongst others). As we explained in Chapter 2, the population can be divided in groups with fixed membership over time (cohorts), and the sample means for each cohort on each time period can be treated as a panel subject to measurement errors. The classical estimators for panel data can be modified in a convenient way to obtain consistent estimators using the cohort means. We leave this approach for future research.

5.- RESULTS

We have estimated two equations, one for food consumption and another one for energy and transport. In both equations we condition on the growth rate of consumption of clothing and footwear. The food equation to be estimated is

$$\Delta \log c_{it}^{f} = \beta_1 \Delta \log c_{it}^{e} + \beta_2 \Delta \log c_{it}^{c} + \theta \log(1+r_t^{f}) + \gamma' \Delta \eta_{it} + seas + \epsilon_{it} \qquad (5.1)$$

where c_{it}^{f}, c_{it}^{e} and c_{it}^{c} are consumption by household i in period t of food, energy and transport, and clothing respectively; r_t^{f} is the commodity-specific real interest rate; η_{it} is a vector of family characteristics, which includes the number of babies, children, and household members older than 64, family size, husband age and age squared and dummies for the labour market status of the household head and the wife; seas are seasonal dummies and ϵ_{it} is the disturbance term. The equation for energy and transport is analogous. As we mention earlier we estimate the Euler equations by GMM[16].

We have to choose a set of instruments that are uncorrelated with the disturbance term. In our application the instrument set comprises lagged values of the endogenous variables and contemporaneous values of the exogenous variables as it is explained in detail below. The instruments will provide a set of moment restrictions, $E(Z_{it}^{j}\epsilon_{it})=0$, $j=1,..,J$, where Z_{it}^{j} is the j-th instrument for household i in period t and ϵ_{it} is the disturbance term. These restrictions can be seen as a system of equations relating the parameters of the model. In the overidentified case (when we have more restrictions than parameters to estimate), as it is in our application, the system will not have a solution once we replace the moment restrictions by their sample counterparts. The GMM estimator minimizes a quadratic form in these sample moments using any positive definite matrix as a weighting matrix. The GMM estimates reported in the tables below are two-step estimates, i.e. they are obtained in a second iteration using as weighting matrix the inverse of a consistent estimate of the variance-covariance matrix of the moment restrictions, and the standard errors that we present are robust to general forms of heteroscedasticity and serial correlation.

As mentioned earlier, in the absence of measurement errors in consumption, we could use as instruments for the model any endogenous variable dated t-1 or earlier. Since we consider demographic variables as

[16] We use the DPD program written in Gauss by Arellano and Bond (1988).

exogenous[17], we only have to instrument consumption variables, the interest rate and the labour force dummies using past information. The results we obtained estimating these equations by GMM, and including in the instrument set: real income, the nominal interest rate, real consumption of food, clothing, and energy and transport, and the labour force dummies in t-1 and t-2, as well as contemporaneous values of the exogenous variables, clearly suggested the inadequacy of some instruments. We obtained very large values for the Sargan test of overidentifying restrictions. In order to see whether this rejection was due to the fact that endogenous variables dated t-1 were not valid instruments, or else that a particular instrument was not valid, we estimated the model excluding from the instrument set real income, real consumption of food (in the equation for food), or real consumption of energy and transport (in the equation for energy and transport), and we reached very similar results. Furthermore, the negative first order serial correlation of the residuals also indicated that endogenous variables dated t-1 were not valid instruments for the model. These two issues are indicative of measurement errors in consumption.

The presence of measurement errors in consumption will add extra terms to the disturbance. These extra terms will have an MA(1) structure, provided that the measurement errors are serially uncorrelated. If this is the case, consumption variables dated at t-1 will not be valid instruments because measured consumption at t-1 will be correlated with the error term

[17] In principle, this assumption is not very convincing in the case of children. However, the exogeneity of children when we are modelling consumption does not seem to be such an important issue as it is in the context of female labour supply (see Browning (1992)).

through its measurement error. However, if the measurement errors are serially uncorrelated, measured consumption at t-2 or earlier will be a valid instrument for the model. Something similar happens in the presence of random preferences. If the unobservable component of η_{it} in equation (5.1) is serially uncorrelated an extra MA(1) term will be added to the disturbance, invalidating endogenous variables dated at t-1 as instruments for the model. Alternative plausible assumptions are that this component is constant over time or that it has a random walk structure. In the first case the unobservable heterogeneity will vanish since it enters the equation as a change over time. In the random walk case, a white noise term will be added to the disturbance and variables dated t-1 will still be valid instruments. We consequently decided to estimate the model using the same instruments as above but excluding consumption and income in t-1.

We have estimated the model including different lags of the endogenous variables in the instrument set. The results obtained for the different instrument set specifications were quite similar. The results based on the ECPF are presented in tables 1 and 2. The instrument set used in both equations includes: income, consumption of food, consumption of clothing, and consumption of energy and transport in t-2 and t-4; the interest rate and the labour force dummies in t-2 and t-3; and contemporaneous values of the exogenous variables. In columns (3) and (4) we have included time dummies to pick up the effect of aggregate shocks which are not explained due to fluctuations on the interest rate. The set of time dummies is significant, however the results are only slightly altered when we introduce these dummies, both in the equation for food (compare columns (1) and (2) with (3) and (4) in table 1) and in the equation for energy and transport (compare columns (1) and (2) with (3) and (4) in table 2). In columns (2) and (4) we include contemporaneous income growth as an additional regressor, the estimated coefficient is not significant providing evidence of no excess sensitivity of consumption growth to income.

The Sargan test of overidentifying restrictions does not reject the instrument set. This result indicates that endogenous variables dated t-2 and earlier are valid instruments as we could expect if measurement errors are white noise. Furthermore, the values of the m1 and m2 statistics for first and second order serial correlation of the residuals provide evidence of first order but not of second order correlation, reinforcing the evidence of white noise measurement errors and hence the validity of the instrument set[18].

None of the labour market dummies are significant at 5% in any of the specifications. This can indicate that changes in labour force status do not influence consumption growth but there is not enough evidence to guarantee that. The demographic variables do not seem to play an important role in explaining consumption growth. In the food equation none of these variables is significant at 5%, and in the equation for energy and transport only the change in the number of children is significant at 5%. The reason why the demographic variables are not significant is probably

---

[18] These statistics are asymptotically distributed as standard normals, see Arellano and Bond (1991).

because Family composition does not change over time for most of the households in our sample (the change in the demographic variables that we have considered is not zero only for about 4% of the observations).

The growth rate of consumption of energy and transport is significant in the food equation, and so is the growth rate of food consumption in the equation for energy and transport. This fact provides evidence of non-separability in the Euler equations. This result was also obtained for the US by Attanasio and Weber (1992).

Given that we do not find evidence of excess sensitivity of consumption growth to income, we could think, as we mention earlier, that the evidence found in the studies which consider an additive separable utility function could be due to this sort of misspecification of the normalization of the utility function. However, when we do not include conditioning commodities in the Euler equations, we do not find evidence of excess sensitivity either.

As we commented earlier, the presence of aggregate shocks that influence different families in different ways will invalidate the results relying in cross-section asymptotics. If this were the case, we would expect to obtain different values for the estimated parameters of the model when we use the second data set (the EPC). In order to test the stability of the parameters, we have estimated the Euler equations for food, and for energy and transport using the data from the EPC.

The EPC does not provide information on the labour market status of the wife. Furthermore, the husband is considered working if he was working for at least 13 hours during the reference week. This definition coincides with the definition of working full time in the ECPF. To be able to test the stability of the parameters, we have to use the same set of regressors for the two samples. Therefore, we have estimated the two Euler equations conditioning in just to labour market dummies (husband full time employed and husband unemployed).

In tables 3 and 4 we present these results for the ECPF (columns (1) and (3)) and the EPC (columns (2) and (4)). In columns (3) and (4) we have included time dummies. The instrument set that we have used is the following: income, consumption of food, consumption of clothing, and consumption of energy and transport in t-2 and t-4; the interest rate and the labour force dummies in t-2 and t-3; and contemporaneous values of the exogenous variables. For the EPC we have excluded from the instrument set the lags of the dependent variable.

The estimated parameters for the EPC look different that those for the ECPF. In the food equation (table 3), the only variables that are significant, when we use the EPC, are the number of children and the interest rate. In the equation for energy and transport (table 4), none of the variables are significant when we use the EPC. The signs of some of the parameters are also different. However, the Wald test of the equality of the parameters does not lead to a rejection of the null hypothesis (except when we compare columns (1) and (2) in table 4). The reason why we fail to reject the null hypothesis is probably the low precision of our estimates.

Table 1

Food, Alcohol and Tobacco

	(1)	(2)	(3)	(4)
Δln(clothing)	-0.0166	-0.0200	-0.0285	-0.0366
	(0.0520)	(0.0558)	(0.0555)	(0.0607)
Δln(entr)	0.2779	0.2751	0.3128	0.3081
	(0.1093)	(0.1102)	(0.1199)	(0.1215)
Δbabies	0.0050	0.0008	0.0063	-0.0015
	(0.0506)	(0.0561)	(0.0526)	(0.0597)
Δchildren	0.0683	0.0645	0.0757	0.0689
	(0.0406)	(0.0447)	(0.0430)	(0.0477)
Δelder	0.0908	0.0911	0.0895	0.0891
	(0.1076)	(0.1095)	(0.1158)	(0.1197)
Δfsize	0.0567	0.0664	0.0583	0.0772
	(0.0355)	(0.0583)	(0.0373)	(0.0630)
hage	-0.0009	-0.0009	-0.0011	-0.0011
	(0.0006)	(0.0006)	(0.0006)	(0.0007)
Δhfullemp	-0.9720	-1.0023	-1.2756	-1.3472
	(0.6113)	(0.6263)	(0.6853)	(0.7138)
Δhpartemp	-0.5150	-0.5431	-0.8249	-0.8956
	(0.6656)	(0.6890)	(0.7682)	(0.8103)
Δhunemp	-0.7432	-0.7935	-0.9856	-1.0931
	(0.4972)	(0.5466)	(0.5604)	(0.6252)
Δwfullemp	0.2812	0.3101	0.3150	0.3725
	(0.4142)	(0.4212)	(0.4392)	(0.4476)
Δwpartemp	0.1259	0.1107	0.1263	0.1005
	(0.3809)	(0.3895)	(0.4045)	(0.4169)
Δwunemp	0.5896	0.5673	0.6795	0.6453
	(1.0918)	(1.0643)	(1.1539)	(1.1143)
ln(1+r)	0.0917	0.1123	-	-
	(0.5312)	(0.5436)		
lnrinc	-	-0.0613	-	-0.1149
		(0.2600)		(0.2795)
time dummies	no	no	yes	yes
seasonal dum.	yes	yes	no	no
Sargan Test	14.483	14.247	11.574	11.028
df	13	12	12	11
m1	-7.412	-7.371	-6.798	-6.592
m2	0.197	0.205	0.516	0.539

The dependent variable is the growth rate of food consumption.
Numbers in parentheses are standard errors.
The description of the variables is provided below table 4.

Table 2

Energy and transport Consumption

	(1)	(2)	(3)	(4)
Δln(clothing)	0.0144	0.0122	0.0373	0.0443
	(0.1020)	(0.1108)	(0.1087)	(0.1203)
Δln(food)	1.0494	1.0428	1.1139	1.1276
	(0.4498)	(0.4550)	(0.4733)	(0.4842)
Δbabies	-0.0745	-0.0766	-0.0655	-0.0589
	(0.0980)	(0.1079)	(0.1016)	(0.1157)
Δchildren	-0.1778	-0.1796	-0.1838	-0.1793
	(0.0736)	(0.0798)	(0.0768)	(0.0843)
Δelder	-0.3040	-0.3011	-0.2782	-0.2777
	(0.2201)	(0.2178)	(0.2335)	(0.2356)
Δfsize	0.0124	0.0173	-0.0021	-0.0168
	(0.0775)	(0.1283)	(0.0819)	(0.1385)
hage	0.0011	0.0011	0.0014	0.0014
	(0.0012)	(0.0012)	(0.0013)	(0.0013)
Δhfullemp	1.0290	0.9966	1.5850	1.6476
	(1.2568)	(1.2892)	(1.4199)	(1.4788)
Δhpartemp	0.3204	0.2980	0.9108	0.9692
	(1.3714)	(1.3965)	(1.5735)	(1.6322)
Δhunemp	0.8704	0.8337	1.3150	1.4037
	(0.9966)	(1.1240)	(1.1280)	(1.2874)
Δwfullemp	-0.1762	-0.1330	-0.3031	-0.3444
	(0.9257)	(0.9408)	(0.9582)	(0.9697)
Δwpartemp	0.2639	0.2554	0.2558	0.2851
	(0.8339)	(0.8615)	(0.8665)	(0.9032)
Δwunemp	-1.4303	-1.3688	-1.4876	-1.4275
	(2.7039)	(2.8518)	(2.7669)	(2.8762)
ln(1+r)	0.4760	0.4710	-	-
	(1.8277)	(1.9099)		
lnrinc	-	-0.0328	-	0.0796
		(0.5469)		(0.5770)
time dummies	no	no	yes	yes
seasonal dum.	yes	yes	no	no
Sargan Test	13.485	13.621	10.764	10.650
df	13	12	12	11
m1	-8.392	-8.452	-7.571	-7.400
m2	0.382	0.374	0.571	0.584

The dependent variable is the growth rate of consumption of energy and transport.
Numbers in parentheses are standard errors.
The description of the variables is provided below table 4.

Table 3

	Food, Alcohol and Tobacco			
	(1)	(2)	(3)	(4)
Δln(clothing)	-0.0394	0.0056	-0.0537	0.0109
	(0.0605)	(0.0409)	(0.0658)	(0.0415)
Δln(entr)	0.3683	-0.0302	0.4271	-0.0418
	(0.1410)	(0.1539)	(0.1600)	(0.1538)
Δbabies	0.0120	0.0361	0.0145	0.0362
	(0.0538)	(0.0353)	(0.0572)	(0.0356)
Δchildren	0.0795	-0.0065	0.0898	-0.0073
	(0.0442)	(0.0220)	(0.0482)	(0.0219)
Δelder	0.1031	0.0062	0.1017	0.0193
	(0.1120)	(0.0464)	(0.1251)	(0.0466)
Δfsize	0.0510	0.0869	0.0517	0.0831
	(0.0387)	(0.0375)	(0.0421)	(0.0377)
hage	-0.0008	-0.0003	-0.0010	-0.0002
	(0.0006)	(0.0003)	(0.0007)	(0.0004)
Δhfullemp	-0.8418	0.0675	-1.2391	0.2029
	(0.6729)	(0.7490)	(0.7863)	(0.7731)
Δhunemp	-0.7199	-0.1027	-1.0397	0.0724
	(0.5305)	(0.8366)	(0.6251)	(0.8590)
ln(1+r)	0.2008	1.6324	-	-
	(0.5450)	(0.6991)		
time dummies	no	no	yes	yes
seasonal dum.	yes	yes	no	no
Sargan Test	9.426	8.570	6.766	0.560
df	7	5	6	4
m1	-5.885	-15.530	-5.308	-16.523
m2	0.405	-0.258	0.700	-0.262

Wald test for stability of the parameters (chi-square distribution)

H_0: column (1) = column (2), statistic = 17.48 df = 14 p-value = 0.2313.
H_0: column (3) = column (4), statistic = 11.99 df = 9 p-value = 0.2140.

The dependent variable is the growth rate of food consumption.
Numbers in parentheses are standard errors.
The description of the variables is provided below table 4.

Table 4

Energy and transport

	(1)	(2)	(3)	(4)
Δln(clothing)	0.0528	0.0568	0.0793	0.0515
	(0.1109)	(0.0711)	(0.1171)	(0.0712)
Δln(food)	1.0939	0.0235	1.1611	0.0134
	(0.4586)	(0.2010)	(0.4798)	(0.1998)
Δbabies	-0.0775	0.0289	-0.0656	0.0295
	(0.0950)	(0.0617)	(0.0986)	(0.0615)
Δchildren	-0.1743	0.0104	-0.1800	0.0111
	(0.0702)	(0.0362)	(0.0737)	(0.0360)
Δelder	-0.2915	0.0012	-0.2555	-0.0032
	(0.1981)	(0.0806)	(0.2138)	(0.0804)
Δfsize	0.0014	0.0872	-0.0167	0.0893
	(0.0778)	(0.0581)	(0.0830)	(0.0584)
hage	0.0009	0.0000	0.0013	0.0000
	(0.0012)	(0.0006)	(0.0013)	(0.0006)
Δhfullemp	0.8236	0.9609	1.5642	0.8763
	(1.2497)	(1.6091)	(1.4431)	(1.6115)
Δhunemp	0.8951	1.0850	1.4521	0.9945
	(0.9805)	(1.8007)	(1.1289)	(1.7881)
ln(1+r)	0.6007	-0.1555	-	-
	(1.7219)	(0.3108)		
time dummies	no	no	yes	yes
seasonal dum.	yes	yes	no	no
Sargan Test	8.516	1.414	5.941	1.3041
df	7	5	6	4
m1	-8.978	-13.864	-7.909	-14.175
m2	0.328	-0.294	0.549	-0.265

Wald test for stability of the parameters (chi-square distribution)

H_0: column (1) = column (2), statistic = 32.48 df = 14 p-value = 0.0034.
H_0: column (3) = column (4), statistic = 14.06 df = 9 p-value = 0.1202.

The dependent variable is the growth rate of consumption of energy and transport.
Numbers in parentheses are standard errors.
The description of the variables is provided below.

Notes to tables

- *clothing*, *entr* and *food* are real consumption of clothing, energy and transport,and food respectively.
- *babies*, *children* and *elder* are the number of babies (between 0 and 2 years old), the number of children (between 3 and 17), and the number of household members older than 64 respectively.
- *fsize* is family size.
- *hage* is the age of the household head.
- *hfullemp* (full-time employed), *hpartemp* (part-time employed) and *hunempl* (unemployed), are dummy variables for the labour market status of the household head. Analogously *wpartemp*, *wfullemp* and *wunemp* for the wife.
- *r* is the commodity specific real interest rate.
- *income* is real income.
- *Sargan Test* is the Sargan test of overidentifying restrictions. It is distributed as a chi-square with *df* degrees of freedom.
- *m1* and *m2* are test statistics for first and second order serial correlation, their distribution is standard normal (see Arellano and Bond (1991) for a description of these tests).

6.- CONCLUSIONS

The results we have obtained for Spain add new evidence reinforcing the life-cycle permanent income hypothesis. We allow for non-separabilities among consumption goods in the Euler equations and we do not find evidence of excess sensitivity.

We estimate the Euler equations using two data sets corresponding to different periods of time. On the basis of a Wald test we do not reject the stability over time of our results, once we include time dummies in the model. This fact suggests that the effect of aggregate shocks, which are not explained by fluctuations on the interest rate, can be captured by the time dummies. However, some of the coefficients are not very well determined, a fact that makes a rejection difficult.

Measurement errors in consumption and non-separabilities seem to be an important issue, and they may be responsible for the failure of the model commonly found in the literature.

Data Appendix

Table A1

Descriptive Statistics for Quarterly Expenditures and Income

Real Expenditures on Food, Alcohol and Tobacco (1983 pesetas)

year	Median	Mean	St. Dev.	Minimum	Maximum
1978	98070	106565	49129	14408	503645
1979	97743	109235	53489	11632	532677
1980	102094	112094	53198	18050	557679
1981	100998	112143	57812	9080	743292
1982	100886	111760	56470	13364	714264
1983	99758	109438	53701	18774	596327
1985	96412	105515	49196	8662	710813
1986	93099	102685	49769	13509	468421
1987	93582	102749	51253	11025	763150
1988	92790	101222	47393	8079	484540
1989	93426	101464	48489	11112	500322

Real Expenditures on Energy and Transport (1983 pesetas)

year	Median	Mean	St. Dev.	Minimum	Maximum
1978	21313	30009	29183	213	327059
1979	23290	31277	27972	561	219373
1980	21953	29279	25916	478	247346
1981	21552	28991	25201	435	237739
1982	23873	30074	24650	371	236680
1983	22915	30096	25064	406	321644
1985	38079	49061	45958	801	776900
1986	36817	48460	45284	939	532463
1987	36455	47589	44833	230	660692
1988	36920	48883	45554	461	513408
1989	37351	50494	49049	1524	856798

Real Expenditures on Clothing (1983 pesetas)

year	Median	Mean	St. Dev.	Minimum	Maximum
1978	22721	28702	23276	16	192953
1979	21024	26782	32850	59	1198863
1980	19335	24629	21768	26	282816
1981	17377	22691	20493	67	271292
1982	17676	22531	19474	40	204691
1983	16877	21638	19798	100	301717
1985	28884	40953	43027	59	645991
1986	29142	43062	51106	58	1579519
1987	29323	42113	45059	42	532827
1988	28091	41723	44990	85	491009
1989	29260	43165	46318	213	639557

Real Income (1983 pesetas)

year	Median	Mean	St. Dev.	Minimum	Maximum
1985	267653	304181	180087	20487	2332197
1986	265720	307253	174370	23247	2297062
1987	282074	319876	180439	30420	4043017
1988	295250	331436	172158	31121	1961325
1989	309161	355995	194707	34725	2237499

Table A2

Descriptive Statistics for Demographic characteristics

year	babies	child	elder	famsize	hage	hemp	hunemp
1978	0.152 (0.404)	1.504 (1.280)	0.125 (0.374)	4.346 (1.409)	45.96 (8.819)	0.938 (0.242)	0.021 (0.142)
1979	0.128 (0.365)	1.504 (1.297)	0.132 (0.391)	4.370 (1.472)	46.58 (8.866)	0.936 (0.244)	0.019 (0.136)
1980	0.093 (0.307)	1.493 (1.279)	0.150 (0.408)	4.363 (1.395)	46.86 (8.646)	0.926 (0.262)	0.023 (0.150)
1981	0.100 (0.325)	1.479 (1.283)	0.139 (0.393)	4.378 (1.419)	47.03 (8.529)	0.903 (0.296)	0.038 (0.192)
1982	0.092 (0.320)	1.423 (1.288)	0.129 (0.379)	4.317 (1.422)	47.52 (8.624)	0.903 (0.296)	0.039 (0.193)
1983	0.078 (0.289)	1.350 (1.243)	0.115 (0.358)	4.229 (1.394)	48.19 (8.787)	0.877 (0.328)	0.056 (0.230)
1985	0.145 (0.376)	1.458 (1.293)	0.130 (0.379)	4.441 (1.477)	45.52 (9.991)	0.806 (0.395)	0.080 (0.271)
1986	0.144 (0.371)	1.395 (1.228)	0.113 (0.359)	4.366 (1.438)	45.44 (10.23)	0.818 (0.386)	0.074 (0.261)
1987	0.157 (0.393)	1.351 (1.202)	0.105 (0.344)	4.331 (1.390)	45.53 (10.09)	0.828 (0.377)	0.078 (0.268)
1988	0.147 (0.393)	1.282 (1.150)	0.113 (0.357)	4.298 (1.358)	45.73 (10.19)	0.832 (0.374)	0.067 (0.250)
1989	0.146 (0.379)	1.261 (1.119)	0.120 (0.364)	4.283 (1.345)	45.99 (10.24)	0.839 (0.367)	0.051 (0.219)

Sample means. Sample Standard deviations in parenthesis.

CHAPTER 4. INTERTEMPORAL DEMANDS: SOME PANEL DATA ESTIMATES

1. INTRODUCTION

In this chapter we develop a system of demand equations for the simultaneous determination of the inter-temporal and intra-temporal allocation of expenditure - that is, a system of Frisch demands. We then go on to estimate the parameters of this system using a panel data set that has full consumption information. The advantage of having such estimates is that we can predict the responses to price changes that allow for changes in within period allocation (traditional uncompensated or Marshallian responses) and for changes in allocation between periods. Thus a tax change by the government that changes prices will typically lead agents to adjust how much they spend in any period and how that amount is allocated amongst different goods. Estimates of Marshallian systems are necessarily silent on the first adjustment.

Although the ideas behind the estimation of Frisch demand systems have been understood for some time (see, for example, MaCurdy (1983) and Browning, Deaton and Irish (1985)) there have been very few attempts to implement this on consumption data. Indeed, the only other example we know that uses micro data is Blundell, Browning and Meghir (1994). Partly the relative paucity of estimates of Frisch demand systems seems to be because of data problems - we need panel data on individual demands[19]. The other reason is that there is a wide-spread feeling that when we consider inter-temporal allocation we can treat total expenditure within any period as a composite commodity ('consumption'). Whilst this is true under restrictions that are not too onerous (see Gorman (1959)) it seems worthwhile exploring the more general case.

Apart from its congruence with theory, there are also other potential advantages from estimating a Frisch system. First, in neo-classical models of demand the equation governing inter-temporal allocation (the 'consumption function') shares parameters with the system governing the spending of total expenditure (the (Marshallian) 'demand system'). There is thus a potential gain in efficiency in estimating the two together. Second, the specification and estimation of a Frisch system brings a coherence to any discussion of the response to price and income changes. We can take account of the full response rather than the usual 'holding total expenditure constant' response. Finally, from the estimates of a Frisch system we can recover all of the parameters of interest for both inter-temporal and intra-temporal allocation. Thus we can estimate both the inter-temporal substitution elasticity (ISE) and the usual uncompensated own-price and cross-price effects.

In sections 2 and 3 we discuss the different ways that have been suggested to estimate models of the intra-temporal and inter-temporal

---

[19] Blundell *et al* overcome this by using a quasi-panel data set constructed from a long time series of Family Expenditure Surveys.

allocation of expenditures. We establish a set of criteria which should ideally be satisfied by a Frisch demand system in terms of consistency with the theory, flexibility and econometric tractability. Guided by these criteria we choose a functional form for the Frisch system that allows us to estimate a set of exact Euler equations. That is, we first difference the marginal utility of money λ and not some function of λ (as in Blundell *et al* (1994) and Browning *et al* (1985), for example). Since it is λ that follows a random walk according to the traditional life cycle theory this exactness property gives a closer fit between theory and practice. Furthermore, as we document below, our specification is more flexible than others that have been suggested in the literature in that it does not restrict either the ISE or within period cross price effects.

In the empirical section of the chapter we use a panel data set that has detailed information on individual demands; family income and family characteristics. As far as we are aware this the first time such a data set has been used in this context. Based on our estimates we calculate the ISE for different family compositions and different levels of expenditure. We find that both have an important effect on inter-temporal allocation.

2. ESTIMATING FRISCH DEMAND SYSTEMS

We start with a system of Frisch demands (see appendix A):

$$q_{iht} = f^i(p_{ht}, \lambda_{ht}) \qquad i = 1,2,..,n \tag{2.1}$$

where q_{iht} is the quantity of good i consumed in period t by household h;

p_{ht} is a vector of discounted prices for household h in period t and λ_{ht} is the (unobservable) marginal utility of money for household h in time t. The prices differ across households since they are discounted by nominal rates and different households may face different interest rates (either because there are different nominal rates for different households or because of taxes). In the next section we shall discuss at length how to choose a parameterization for the $f^i(.)$'s; in this section we concentrate on estimating the parameters of $f^i(.)$ in (2.1). We are aware of four methods of estimation.

The usual way to estimate the parameters of (2.1) is to use the Euler equation:

$$\lambda_{h(t-1)} = E_{h(t-1)}(\lambda_{ht}) \tag{2.2}$$

where $E_{h(t-1)}(.)$ is the expectations operator conditional on information available to household h in time (t-1). One particularly attractive way to use this condition is to parameterize (2.1) so that some known function of the Frisch demands and prices ($\tau^i(q,p)$, say) is additive in the marginal utility of money and then first difference:

$$\Delta\tau^i(q_{iht},p_{ht}) = \beta_i \Delta\lambda_{ht} \tag{2.3}$$

Some forms that satisfy this condition and are consistent with a utility maximising assumption are given in Browning, Deaton and Irish (1985). The parameters of (2.3) can be estimated using the condition that $\Delta\lambda_{ht}$ should be orthogonal to all information dated (t-1) or earlier for household h. We shall call this the conventional Euler equation approach. As far as we know, the only papers that estimate an exact Euler equation based on (2.3) are the macro papers that use quadratic preferences (see Hall (1978)). The papers that use micro data estimate approximate Euler equations based on first differences of some known function of λ.

Very often the conventional Euler equation approach is used with supplementary preference structure restrictions. For example, if good 1 is leisure it is often assumed that this is additively separable form other goods so that only the discounted price of good 1 (that is, the discounted wage) enters (2.3) for good 1.

A second way to estimate the parameters of (2.1) is to recognise that this Frisch system combines both the inter-temporal allocation problem and the intra-temporal allocation problem (loosely, the consumption function and the demand system respectively). To see that all of the parameters of the two 'stages' of allocation can be recovered from (2.1) note first that we can multiply both sides by the price of good i and add over goods to derive total expenditure in period t:

$$x_{ht} = \sum_i p_{iht} q_{iht} = \sum_i p_{iht} f^i(p_{ht}, \lambda_{ht}) = g(p_{ht}, \lambda_{ht}) \qquad (2.4)$$

This expression allows us to model the inter-temporal allocation of total expenditure (or consumption if we divide both sides by a price index) using conventional Euler equation techniques. Thus (2.4) is a Frisch consumption function. Indeed, this is the function that is implicitly estimated in the majority of Euler equation studies that assume just one good. In these studies the prices are restricted to enter only through a (linear homogeneous) price index h(.):

$$x_{ht} = \tilde{g}(h(p_{ht}), \lambda_{ht})$$

To recover the parameters of intra-temporal demand note that we can invert on the unobservable λ in (2.4)[20] and derive an expression for this variable in terms of observables:

$$\lambda_{ht} = \chi(p_{ht}, x_{ht}) \qquad (2.5)$$

Substituting this into (2.1) gives a conventional (Marshallian or uncompensated) demand system:

$$q_{iht} = f^i(p_{ht}, \chi(p_{ht}, x_{ht})) = g^i(p_{ht}, x_{ht}) \qquad (2.6)$$

The important thing to note here is that knowledge of the Frisch demands $f^i(.)$ allows us construct the Marshallian demands $g^i(.)$. The converse is not, of course, true. Also, note that the $g^i(.)$'s in (2.6) and the $g(.)$ in (2.4) share parameters so there are obvious gains in efficiency from estimating them together.

We can also turn this procedure on its head and derive the parameters of (2.1) in two stages. First we estimate a conventional demand system. When doing this we need to take into account the possible endogeneity of total expenditure in (2.6). Given these estimates we can define the marginal utility of money in any period up to the parameters of the normalisation of the utility function (the latter can never, of course, be derived from the demand system alone). Given this we can then estimate the parameters of the normalisation from a single Euler equation approach

---

[20] A sufficient condition for this is that the cardinalisation of the utility function for the intertemporally additive representation is strictly concave.

on (2.4); see MaCurdy (1981) or Blundell, Browning and Meghir (1994). We shall call this approach the two stage budgeting approach since it takes its inspiration from the two stage procedure analyzed by Gorman (1959).

A third way to estimate the parameters of (2.1) was suggested by Altonji (1986). It is similar to the last procedure except that rather than inverting on total expenditure we invert on one demand function (good 1, say):

$$\lambda_{ht} = \xi(p_{ht}, q_{1ht}) \qquad (2.7)$$

and then substitute in for this in the other equations:

$$q_{iht} = f^i(p_{ht}, \xi(p_{ht}, q_{1ht})) = h^i(p_{ht}, q_{1ht}) \qquad i=2,3,..,n \qquad (2.8)$$

This allows one to estimate the demands for the last (n-1) goods conditional on the demand for good 1. We term this the conditional approach. Since the demand for good 1 is likely to be endogenous for the other demands we need to find instruments to consistently estimate the parameters of these conditional demands[21]. Even then this does not give all of the parameters of (2.1); just as in the previous case we cannot identify the normalisation of the utility function (unless we impose some constraint like additivity between good 1 and all other goods). To do that we need to estimate the Frisch demand for good 1. This can be done using conventional Euler equation techniques.

The final way to estimate the parameters of (2.1) is rather different

[21] These instruments may differ from those for total expenditure in the two stage budgeting approach.

from the three outlined above; it was first used in Attfield and Browning (1985)[22]. It starts from the fact that λ in (2.1) is unobservable and then uses latent variable techniques to allow for this. The identification comes from imposing some of the conditions derived form the maximisation problem 'behind' (2.1). Amongst the integrability conditions that the demands in (2.1) have to satisfy if they are to be consistent with utility maximisation are:

Homogeneity: $f^i(.)$ is zero homogeneous in p and λ^{-1} ,

Symmetry: $\frac{\partial f^i}{\partial p_j} = \frac{\partial f^j}{\partial p_i}$ for all i and j.

It turns out that with three goods the system is just identified and with four or more goods the system is over-identified (see Attfield and Browning (1985) for details). We term this the latent variable approach.

The principal advantage of the latent variable approach is that in estimation we do not need to assume anything about the orthogonality of $\Delta\lambda$ to past information. Thus these latter conditions, which are maintained in the first three approaches above, need not be imposed and are thus testable.

3. CHOOSING A FRISCH DEMAND SYSTEM

We wish to estimate the parameters of a system of Frisch demands where each demand q_i is defined implicitly by:

---

[22] In fact, that is the only place it has ever been used to the best of our knowledge.

$$F^i(q_i,p_1,...,p_n,r;\theta) = F^i(q_i,p,r;\theta) = 0 \tag{3.1}$$

where r is the inverse of the marginal utility of expenditure[23], the p_i's are prices and θ is a vector of unknown parameters.

In this system each good has its own Frisch elasticity, defined in the usual way:

$$\phi_i = \frac{\partial \log q_i}{\partial \log p_i}$$

If there is only one good (n = 1) then this is usually known as the intertemporal substitution elasticity (ISE) ϕ. In Appendix B, we show that for a system of Frisch demands the ISE can be derived from the individual demands by first defining total expenditure x as the sum of individual expenditures ($= \sum p_k q_k$) and then using:

$$\phi = \frac{\partial \log x}{\partial \log r} \tag{3.2}$$

This elasticity occupies a central position in the analysis of intertemporal allocation.

How should we choose a functional form for the $F^i(.)$'s? The usual procedure is to specify a utility function and then to derive (3.1) as the solutions to the system of first order conditions for a constrained optimisation problem. We adopted an alternative approach. We first set down a list of seven criteria for Frisch demands. We conjecture that there is no Frisch system that satisfies all seven of these criteria

[23] For reasons that will became clear below it is easier to work with the inverse rather than the level of the marginal utility of money.

simultaneously[24]. If this conjecture is true then we necessarily have to trade off amongst our criteria and search for a good Frisch system rather than a best one.

In practice we spent many, many hours trying different functional forms. We ended up choosing the following system:

$$q_i = d^i(p;\beta) + \mu_i\left[\frac{\tilde{\mu}(p)}{r} - 1\right]^{\theta} \quad \text{where } \tilde{\mu}(p) = \textstyle\sum_k \mu_k p_k \qquad (3.3)$$

where $(\mu_1,..,\mu_k)$ and β are vectors of parameters. Note that for this to be defined for all θ we require that $\tilde{\mu}(p) > r$, and therefore all the μ_k's have to be positive. We now list the seven criteria that guided our choice of (3.3). We shall illustrate the criteria with this system as we go along. The first of our criteria is motivated by a concern for ensuring that the resulting demands can be consistent with an underlying utility framework. Criteria 2 to 4 are more concerned with flexibility whilst the final three criteria are concerned with econometric tractability.

1. Consistency with theory. As noted in the last section, to be consistent with utility maximisation the demands should be zero homogeneous in (p,r) and symmetric in p[25]:

$$\frac{\partial q_i}{\partial p_j} = \frac{\partial q_j}{\partial p_i} \quad \text{for all } i\text{ , } j$$

---

[24] We have been unable to show this formally.

[25] We discuss the negativity conditions on the Frisch demand under the next criterion.

For the system given in (3.3), these conditions are equivalent to each $d^i(p)$ being zero homogeneous and satisfying $d^i_j = d^j_i$ for $i \neq j$, where d^i_j is the partial of d^i with respect to p_j. This in turn implies that for each i, $d^i(p)$ is the partial with respect to p_i of some linear homogeneous function d(p). We postpone the choice of d(p) until the empirical section, but we impose this condition in the rest of this section.

2. <u>Flexible inter-temporal preferences</u>. It will be desirable to have each demand increasing in r; this corresponds to each good being normal with respect to lifetime wealth. This seems a reasonable requirement for the broad commodity groupings we shall be using. For the system given in (3.3) we have:

$$\frac{\partial q_i}{\partial r} = -\theta\mu_i\left[\frac{\tilde{\mu}}{r}-1\right]^{\theta-1}\frac{\tilde{\mu}}{r^2}$$

Given that all the μ_i's have to be positive, if θ is negative each demand will be increasing in r.

We also require that the Frisch own price response be negative

$$\frac{\partial q_i}{\partial p_i} = d^i_i + (\mu_i)^2\theta\left[\frac{\tilde{\mu}(p)}{r}-1\right]^{\theta-1}\frac{1}{r} \tag{3.4}$$

If $\theta<0$ then a sufficient condition for this is that $d^i_i \leq 0$.

As we noted above an important parameter connected with intertemporal allocation is the ISE. This determines how willing households are to substitute across time. It is entirely reasonable that this should be dependent on lifetime wealth. Thus the functional forms in (3.1) should be flexible enough to allow the ISE to be increasing or decreasing in r since the latter is increasing in lifetime wealth. Our own prior is that wealthy households are more likely to be willing to substitute across time so that we certainly would not wish to use functional forms that restrict ϕ to be decreasing in r.

To derive the intertemporal substitution elasticity for our system we first multiply each side of (3.3) by p_i and sum over i.

$$x = \sum_i p_i q_i = d(p) + \tilde{\mu}\left[\frac{\tilde{\mu}}{r} - 1\right]^{\theta} \tag{3.5}$$

where $\Sigma p_i d^i = d(p)$ by linear homogeneity. This gives total expenditure in each period in terms of the marginal cost of utility and prices; it can be thought of as a Frisch consumption function. Using the definition of the ISE given in (3.2) we have:

$$\phi = -\frac{\theta\tilde{\mu}}{(\tilde{\mu} - r)\left[\frac{d}{\tilde{\mu}}\left(\frac{\tilde{\mu}}{r} - 1\right)^{-\theta} + 1\right]} \tag{3.6}$$

If all the μ_i's are positive and θ is negative (normal goods) then ϕ is positive and increasing in r. Thus our system restricts the intertemporal substitution elasticity to be increasing in lifetime wealth; for the reasons given in the last paragraph we do not regard this as being too restrictive.

3. Flexible intra-temporal preferences As noted in the last section, associated with any Frisch demand system there is a (unique) conventional Marshallian demand system. We require that the Marshallian demand

system associated with our Frisch system not be too restrictive. Thus we would not want use a functional form that imposes, say, homotheticity or additive separability across goods or that implies a diagonal Slutsky matrix (that is, zero Hicksian substitution).

For the system given in (3.3) we can use (3.5) to substitute for r to give the Marshallian system:

$$q_i = d^i(p) + \frac{\mu_i}{\bar{\mu}}(x-\bar{d}) \qquad (3.7)$$

Thus Engel curves are linear in total expenditure - our system implies quasi-homothetic preferences. We regard this as being the most restrictive aspect of our specification. On the other hand, in the data we use we have limited price variability so that we have to use very broad aggregates of goods. The assumption of quasi-homotheticity for such broad aggregates may be more acceptable than for finer categories of goods. Note that if the $d^i(.)$'s are chosen to be flexible then price responses are not restricted. Thus this form is less restrictive than that given in Browning *et al* (1985); the latter was criticised for its lack of price flexibility by (amongst others) Blundell *et al* (1986) and Nickell (1985).

From (3.5) and (3.7) we could use a two-stage budgeting approach to estimation. Alternatively, we could use a conditional approach by inverting (3.3) on good 1 and substituting in the other demands:

$$q_i = (d^i(p) - \frac{\mu_i}{\mu_1}d^1(p)) + \frac{\mu_i}{\mu_1}q_1 \qquad (3.8)$$

This gives a particularly simple form for conditional demands that can then be estimated with (3.3) for good 1 to derive the parameters of the Frisch system.

4. Intertemporal preferences not identified from cross-section. The Marshallian system associated with any Frisch system is independent of the normalisation of the within period utility function (that is, the mapping from Frisch to Marshallian is many-one). In general the ISE should not be independent of this normalisation. Hence we shall require that there are some parameters in (3.1) that are not identified from the Marshallian demand system.

From the Marshallian system given in (3.7) we see that only the ratio of the individual μ_i's are identified and that the parameter θ does not appear in the Marshallian system. Thus the system given in (3.3) satisfies this criterion since the identification of the ISE requires identification of the μ_i's and θ (see (3.6)). Thus this form is also less restrictive that the second specification proposed in Browning *et al* (1985) (expenditures linear in r form), given that in their model the ISE is identified from the Marshallian system.

5. Additivity in the marginal utility of money. The variable r is not observed. To take care of this it will be convenient to have the $F^i(.)$'s in (3.1) additive in some function of r. In particular the Euler equation for intertemporal allocation under uncertainty has that the inverse of r follows a random walk; hence it would be very convenient for estimation to have

each $F^i(.)$ additive in r^{-1}.

Re-arranging (3.3) we have:

$$\left[\left[\frac{q_i-d^i}{\mu_i}\right]^{\frac{1}{\theta}}+1\right]\frac{1}{\bar{\mu}} = \frac{1}{r} \tag{3.9}$$

Thus we can find some non-linear function of (q_i,p) that is equal to the inverse of r. This means that we can use an exact Euler equation approach in estimation. Very often other investigators have first differenced other functions of r than the inverse (in particular, the log form has been much used). This requires auxiliary assumptions on the distributions of future prices and other variables that are used as instruments. We regard this exactness property of our specification as being one of its principal strengths.

6. Linearity in parameters. As usual it would facilitate estimation to have the demands linear in parameters. Indeed, this is necessary for some ways of estimating Frisch demands (for the latent variable approach). It will be clear that the specification given in (3.3) or (3.9) is non-linear.

7. Allowing for measurement error. We require a form for (3.1) that allows us to take into account the fact that the reported expenditure on good i (= p_iq_i) is very likely measured with error. On the other hand we are

willing to assume that the prices are not measured with error[26]. Clearly, multiplying each side of (3.3) by p_i gives a form that allows for measurement error in the usual way. Unfortunately the 'exact' form given in (3.9) does not lend itself to accounting for measurement error in a simple way. In practice, we are forced to make the artificial assumption that

$$\left[\left[\frac{q_i-d^i}{\mu_i}\right]^{\frac{1}{\theta}}+1\right]\frac{1}{\tilde{\mu}} \qquad (3.10)$$

is measured with an additive error that is uncorrelated with all other variables.

4. THE DATA

The data set used is the Spanish Family Expenditure Survey (Encuesta Continua de Presupuestos Familiares) conducted by the National Statistics Office (Instituto Nacional de Estadística). This survey is carried out by personal interview on a quarterly basis, from the first quarter of 1985. Each family is visited four times in a week. During this week all members of the household have to note down their expenditures on a diary. On the intermediate visits, a very detailed information on family characteristics, income and expenditures on goods with a reference period longer than a week is recorded. On the last visit, the agent checks all the

---

[26] This is not strictly defensible since agents face different prices in different areas and over the quarter of observation but taking account of measurement errors seems infeasible at present.

information and collects the diary.

The data set is a rotating panel: in each quarter about 3000 families are interviewed, and 1/8 of the household are replaced by a new random sample. We observe families for at most eight consecutive quarters. However, there is an important percentage of attrition in earlier quarters, mainly during the first two years of the survey. The replacement procedure was very irregular during those two years and households were replaced before they complete their eighth interview. It was only in the fourth quarter of 1986 that the sample started to be stable and a group of households who complete their eighth interview were observed for the first time (in the third quarter of 1988). For this research we have considered families reporting full information for eight quarters and hence we only use thirteen waves of the survey (from fourth quarter 1986 to fourth quarter 1989).

As we explained in Chapter 3, the data set we use has important advantages over other data sets that have expenditure information. The obvious advantage with respect to the U.K. Family Expenditure Survey and the Canadian FAMEX is that we observe households more than once. The U.S. Consumer Expenditure Survey (CEX) does follow households over time, but only for four quarters; thus we cannot observe annual changes. Given the likely importance of annual re-planning in the determination of saving and expenditure this makes the CEX less useful than the longer panel we have. Moreover the CEX has poor income information. Finally, the data set to hand has much fuller information on expenditures than other panel data sets; in particular, the PSID.

In this chapter we are dealing with the estimation of a flexible specification of a Frisch demand system. There is no doubt that preferences are not homogeneous among consumers; they depend on observed and unobserved family characteristics. In applied work we usually take into account the observable heterogeneity by assuming that certain parameters of the model depend on family characteristics (i.e. we assume we know how preferences depend on households characteristics). However, if the sample used is very heterogeneous we will have a very large number of parameters to estimate. This problem is specially serious in non-linear models which are computationally expensive. Our empirical model is quite complex and therefore the estimation procedure is very slow. For this reason we decided to consider a subsample of homogeneous households. We keep in our sample married couples such that the husband was full time employed in a non-agricultural activity and coded head of the household, and the wife was not working in the labour market during the sample period. In all we end up with 215 households.

For the purpose of this research we have used only expenditure in non-durables and services which we have aggregated in two groups of commodities: the first one includes food, alcoholic and non-alcoholic drinks and tobacco; the second other non-durables and services. The latter group comprises expenditure in transport, energy, leisure and non-durable house related expenditures (not including rent). The reason why we have used only two broad groups of commodities is the limited price variability during the sample period.

The price index for each group of commodities is derived from the disaggregated consumer retail price index for Spain published by the National Statistics Office, using the same weights that are used to construct the general index. The nominal interest rate used to deflate prices is an interest rate on deposits provided by Cuenca (1991).

5. THE EMPIRICAL MODEL AND ECONOMETRIC ISSUES

In section 2, we outlined some different methods that have been used in the literature to estimate Frisch demand systems. Unfortunately, the Frisch system in (3.3) is non-linear in parameters, and therefore we can not use the latent variable approach. However, the functional form that we have chosen allows us to estimate a system of two exact Euler equations, and therefore, as we explained above, we do not need to make any of the unpleasant assumptions underlying the empirical studies based on approximate Euler equations (see Altug and Miller (1990) for details).

The d^i functions have to be homogeneous in prices. We use the same flexible form as in Browning *et al* (1985)

$$d^i = \gamma_{ii} + \gamma_{ij}\left(\frac{p_j}{p_i}\right)^{\frac{1}{2}} \tag{5.1}$$

Notice that d^i is symmetric in prices if $\gamma_{ij} = \gamma_{ji}$. Refering back to (3.4), we see that negativity holds if $\gamma_{ij} \geq 0$.

The Euler equation for inter-temporal allocation under uncertainty implies that

$$E_{t-1}\left[\frac{1}{r_t}\right] = \frac{1}{r_{t-1}}$$

where E_{t-1} is the rational expectation operator conditional on information at t-1. Using the Frisch system in (3.3) and the specification for the d^i's in (5.1), we obtain the following system of Euler equations

$$\left[\left(\frac{q_{1t}-\gamma_{11}-\gamma_{12}\left(\frac{p_{2t}}{p_{1t}}\right)^{\frac{1}{2}}}{\mu_1}\right)^{\frac{1}{\theta}}+1\right]\frac{1+i_t}{\mu_1 p_{1t}+\mu_2 p_{2t}} - \left[\left(\frac{q_{1t-1}-\gamma_{11}-\gamma_{12}\left(\frac{p_{2t-1}}{p_{1t-1}}\right)^{\frac{1}{2}}}{\mu_1}\right)^{\frac{1}{\theta}}+1\right]\frac{1}{\mu_1 p_{1t-1}+\mu_2 p_{2t-1}} = \epsilon_{1t} \tag{5.2}$$

$$\left[\left(\frac{q_{2t}-\gamma_{22}-\gamma_{21}\left(\frac{p_{1t}}{p_{2t}}\right)^{\frac{1}{2}}}{\mu_2}\right)^{\frac{1}{\theta}}+1\right]\frac{1+i_t}{\mu_1 p_{1t}+\mu_2 p_{2t}} - \left[\left(\frac{q_{2t-1}-\gamma_{22}-\gamma_{21}\left(\frac{p_{1t-1}}{p_{2t-1}}\right)^{\frac{1}{2}}}{\mu_2}\right)^{\frac{1}{\theta}}+1\right]\frac{1}{\mu_1 p_{1t-1}+\mu_2 p_{2t-1}} = \epsilon_{2t} \tag{5.3}$$

where ϵ_{1t} and ϵ_{2t} are rational expectation errors orthogonal to information available at time t-1, and i_t is the nominal interest rate. Good 1 is food, alcohol and tobacco (FAT) and good 2 is other non-durables (OND). We have eliminated the household specific index to simplify notation. We allow that γ_{11} and γ_{22} are linear functions of the number of children (nch), a dummy for families with at least one child (cdch), the number of adults (nad), a dummy for households with more than two adults (cdad), and seasonal dummies. Unfortunately we could not construct a family specific interest rate given that we do not have information on taxes paid by the household. Our data set does not have information on the region of residence of the family and therefore we have had to use national wide price indexes.

We estimate the system in (5.2) and (5.3) by the Generalized Method of Moments (GMM)[27], using a set of orthogonality conditions based on the lack of correlation between the disturbances and past information available to the household. We tried to estimate all the parameters of the model, however, when we minimized the objective function associated to the set of orthogonality conditions, either convergence was not achieved, or else we obtained very large values for the Sargan test of overidentifiying restrictions. The problem of our specification is the non-linearity which probably implies, in this case, several local minima for the objective function. In order to avoid this problem, we decided to use a grid in $1/\theta$, i.e. minimize the objective function holding θ fixed, and repeat the optimization for several values of θ. We use the grid $1/\theta$=-2.0,-1.9,-1.8,..,2.0 and for the different sets of instrument that we used, the minimum was clearly achieved between -1.1

[27] We have used the "Gauss GMM Package" by Hansen, Ogaki and Heaton (1992).

and -0.8. Then we use a thicker grid between these limits[28].

As we mention above, If the error terms in equation (5.2) and (5.3) were pure conditional expectations errors, they should be orthogonal to all information dated (t-1) or earlier. However, when we include in the instrument set consumption of FAT and OND in period t-1, the Sargan test of overidentifying restrictions rejects the instrument set. This problem disappears when we consider consumption at t-2 or t-3. We know that for the non-linear specification that we are using, we cannot rely on measurement errors on expenditures to explain this fact. However, it seems to be the case that consumption lagged one period is correlated with the error terms in equations (5.2) and (5.3), but consumption lagged two periods is orthogonal to the disturbances.

The results are presented in table 1[29]. In columns (1) and (2), we present the estimates of the model for different sets of instruments, and the results are quite robust to the instrument set specification. The signs of the estimated coefficients are consistent with the theory. The coefficients of the price ratios are positive which implies that the negativity conditions ($\partial q_i/\partial p_i<0$) are satisfied. We have also tested the symmetry condition $\gamma_{ij}=\gamma_{ji}$ and we could not reject the null hypothesis. However, due to the lack of price variability in our sample period, these coefficients are not very well determined. The parameter θ is negative and therefore both

[28] We do not present the results for the different values of $1/\theta$, they are available on request.

[29] The results correspond to the value of $1/\theta$ for which the minimum was achieved.

aggregate commodities (FAT and OND) are normal with respect to life-time wealth[30]. Family characteristics seem to play an important role explaining intertemporal allocation of consumption. This is consistent with the evidence found for some other countries (see for example Blundell *et al* (1994)).

Many applied papers dealing with the intertemporal allocation of expenditures use a functional form for household preferences which restrict the ISE to be constant (see Zeldes (1989) and Runkle (1991) among others). Some other authors restrict the ISE to be deceasing in life-time wealth (see Browning *et al* (1985)) which seems to be also quite unrealistic. Our prior is that wealthy people are more willing to substitute consumption over time (see Blundell *et all* (1994) for some evidence), and we use a functional form for the Frisch demand system which implies that the ISE is increasing in life-time wealth.

We can rewrite (3.6) in terms of total expenditure.

$$\phi = -\theta\left(1-\frac{d}{x}\right)\left[1+\left(\frac{x-d}{\bar{\mu}}\right)^{-\frac{1}{\theta}}\right]$$

and use this expression to calculate the intertemporal substitution elasticity. In table 2 we present the ISE for different household compositions and for the 25th, 50th and 75th percentiles of total expenditure in our sample. The ISE has been calculated using the estimated values of the parameters presented in table 1 column (1). As we comment

[30] The μ_i's were constrained to be positive, otherwise (3.3) is not defined for all values of θ.

above, the ISE is increasing in total expenditure and depends also on family characteristics.

Table 1

FAT and other non-durables

	(1) $1/\theta=-0.85$		(2) $1/\theta=-0.99$	
	FAT	OND	FAT	OND
const	-0.4068 (7.1245)	-1.5796 (0.0890)	-1.0479 (0.7236)	-0.4779 (1.7265)
nch	0.0711 (0.0115)	0.0406 (0.0005)	0.0663 (0.0044)	0.0369 (0.0107)
cdch	-0.0376 (0.0992)	0.0156 (0.0026)	-0.0135 (0.0193)	0.0232 (0.0203)
nad	0.1037 (0.0148)	0.1387 (0.0022)	0.1045 (0.0088)	0.1399 (0.0025)
cdad	0.0732 (0.0686)	0.0948 (0.0050)	0.0781 (0.0119)	0.0919 (0.0339)
price ratio	0.4620 (7.5513)	1.2869 (0.0828)	1.1309 (0.7705)	0.2465 (1.6325)
$\mu 1$	0.4717 (0.6384)		0.3148 (0.1892)	
$\mu 2$	0.0004 (0.0024)		0.0049 (0.0061)	
Sargan Test	3.5109		15.6929	
df	4		10	

Numbers in parenthesis are standard errors. Standard errors robust to heteroscedasticity and time series correlation.

nch is the number of children, *cdch* is a dummy for households with at least one child, *nad* is the number of adults and *cdad* is a dummy for households with more than two adults. *Price ratio* is $(p_2/p_1)^{1/2}$ in the FAT equation and $(p_1/p_2)^{1/2}$ in the OND equation.

Instruments:

column (1): consumption of FAT and OND in t-2 and t-3; nch, cdch, nad, cdad in t-1; a constant and three seasonal dummies.

column (2): income and consumption of FAT and OND in t-2 and t-3; nch, cdch, nad, cdad, and $(p_2/p_1)^{1/2}$ in t-1; a constant and three seasonal dummies.

Table 2

Estimated Intertemporal Substitution Elasticity

		(1)	(2)	(3)
Number of children	Number of adults	ISE	ISE	ISE
0	2	2.31	3.07	3.99
0	3	0.89	1.66	2.61
0	4	0.36	1.01	1.92
1	2	1.94	2.71	3.66
1	3	0.66	1.40	2.35
1	4	0.25	0.80	1.69
2	2	1.53	2.31	3.27
2	3	0.44	1.10	2.03
2	4	0.07	0.58	1.42
3	2	1.16	1.94	2.89
3	3	0.26	0.87	1.76
3	4	0.01	0.37	1.16

The ISE has been obtained using the estimated values of the parameters in column (1) table 1.

In column (1) the ISE has been calculated for the 25th percentile of total expenditure.
In column (2) the ISE has been calculated for the 50th percentile of total expenditure.
In column (3) the ISE has been calculated for the 75th percentile of total expenditure.

6. CONCLUSIONS

This chapter analyzes the different approaches that have been used in the literature to estimate models of inter-temporal and intra-temporal allocation of consumption. We discuss a set of criteria that Frisch demands should ideally verify and we use them as a guide to chose a functional form which is consistent with the theory, rather flexible and tractable from the econometric point of view. Our specification turns to be less restrictive than others in terms of the criteria mentioned above.

In the empirical section we estimate the exact Euler equations associated to the Frisch demand system using panel data and we calculate the ISE for different family composition and different levels of expenditure. Household characteristics seem to play an important role in the allocation of consumption over time.

APPENDIX A. Intertemporal Demands (Frisch Demands)

According to the life-cycle model of consumption, consumers allocate consumption over time to maximize the expected value of the sum of present and future utilities. The maximization problem to be solved is

$$\max E_t\left[\sum_{s=t}^{T-1} V_s(x_s)\right]$$
$$\text{subject to} \qquad (A.1)$$
$$A_{s+1}=(1+i_s)(A_s+y_s-x_s) \qquad s=t,\dots,T-1$$
$$A_T \geq 0$$

Where x_s is total expenditure in period s, y_s is income, i_s is the nominal interest rate, A_s are assets at the beginning of period s and V_s is the indirect utility function. E_t is the conditional expectation operator, conditional on information known by the consumer in period t.

If $A_T=0$, we can rewrite the constraints as

$$A_t+y_t+\sum_{s=t+1}^{T-1} R_s y_s = x_t+\sum_{t+1}^{T-1} R_s x_s$$

where

$$R_s=\prod_{j=t}^{s-1} \frac{1}{1+i_j}$$

We define lifetime wealth as

$$W_t=A_t+y_t+\sum_{s=t+1}^{T-1} R_s y_s$$

then

$$W_{t+1}=(1+i_t)(W_t-x_t)$$

We can write the value function for the optimization problem (A.1) as

$$\begin{aligned} V_t^*(W_t) &= \max_{x_t}[V_t(x_t) + E_t(V_{t+1}^*(W_{t+1}))] \\ \text{s.t.} \quad W_{t+1} &= (1+i_t)(W_t - x_t) \end{aligned} \tag{A.2}$$

Taking derivatives with respect to lifetime wealth, we obtain the Euler equation

$$\frac{\partial V_t^*(W_t)}{\partial W_t} = E_t\left[\frac{\partial V_{t+1}^*(W_{t+1})}{\partial W_{t+1}}(1+i_t)\right] \tag{A.3}$$

$\lambda_t = \partial V_t^*/\partial W_t$ is the marginal utility of wealth. We can define the price of utility $r_t = 1/\lambda_t$, and we can write the Euler equation (A.3) as

$$E_t\left[\frac{1}{r_{t+1}}(1+i_t)\right] = \frac{1}{r_t}$$

From (A.2) we obtain

$$\frac{\partial V_t^*(W_t)}{\partial x_t} = 0 = \frac{\partial V_t(x_t)}{\partial x_t} - E_t\left[\frac{\partial V_{t+1}^*(W_{t+1})}{\partial W_{t+1}}(1+i_t)\right]$$

and therefore

$$\frac{1}{r_t} = \frac{\partial V_t^*(W_t)}{\partial W_t} = \frac{\partial V_t(x_t)}{\partial x_t} = g(x_t, p_{1t}, \ldots, p_{kt})$$

Where $p_{1t}, \ldots, p_{kt}$ are the prices. Then we can write total expenditure as a function of the prices and the price of utility

$$x_t = f(r_t, p_{1t}, \ldots, p_{kt})$$

Taking derivatives with respect to prices, we obtain the Frisch demand system

$$\frac{\partial x_t}{\partial p_{it}} = q_{it} = f_i(r_t, p_{1t}, \ldots, p_{kt}) \qquad i=1,\ldots,k$$

APPENDIX B. The Intertemporal Substitution Elasticity

Let V(p,x) be the indirect utility function. The inverse of the marginal utility of money (r) is defined as $(V_x(p,x))^{-1}$, where V_x denotes the partial derivative of the indirect utility function with respect to total expenditure. If V(p,x) is strictly concave in x then we can invert on x, and write x as a function of p and r

$$x = \chi(p,r)$$

by definition we have

$$rV_x(p,\chi(p,r)) \equiv 1 \tag{B.1}$$

In a multigood setting the ISE is defined by

$$\phi = -\frac{V_x}{xV_{xx}}$$

(see Browning (1989)). Taking partial derivatives with respect to r in (B.1)

$$rV_{xx}\frac{\partial x}{\partial r} + V_x = 0$$

and hence

$$\phi = -\frac{V_x}{xV_{xx}} = \frac{\partial x}{\partial r}\frac{r}{x} = \frac{\partial \ln x}{\partial \ln r}$$

CONCLUSIONS

In the first part of the thesis (Chapter 2), we deal with the estimation of dynamic models using time series of cross-sections. We propose different types of measurement error corrected estimators and we analyze their asymptotic properties. We calculate the asymptotic biases of the non-corrected estimators for the AR(1) model. The size of the biases depends on the parameters of the model, and the measurement error correction appears to be sometimes crucial to obtain valid estimates. We have also carried out Monte Carlo experiments and we have obtained similar results in small samples.

In Chapter 3, we estimate a set of Euler equations derived from the life-cycle model of consumption. Using a specification that allows for non-separabilities in the Euler equations among consumption goods, we do not find evidence of excess sensitivity of consumption growth to income. We have used two data sets for Spain for two different periods of time (1978-83 and 1985-89), and we have tested the stability of our results. Once we include time dummies in the model we cannot reject the stability of the coefficients over the two periods. However, some of the parameters are not very well determined, and therefore, our results are not very conclusive.

In the last chapter, we discuss the properties that a Frisch demand system should ideally verify. On the basis of these criteria, the data available and econometric tractability, we chose a nonlinear functional form for our system. Our specification allows us to estimate a set of exact Euler equations, contrary to the usual practice in the literature. We estimate the model using a rotating panel for Spain, and we calculate the elasticity of intertemporal substitution for different levels of expenditures and different household compositions. We can conclude that the decisions on consumption allocation over time are partly determined by the particular demographic characteristics of the family.

REFERENCES

Altonji, J.G. (1986): "Intertemporal Substitution in Labour Supply: Evidence from Micro-Data", *Journal of Political Economy*, 94, s176-s215.

Altug, S. and R.A. Miller (1990): "Household choices in Equilibrium", *Econometrica*, 58, 543-570.

Arellano, M. and S. Bond (1988): "Dynamic Panel Data Estimation Using DPD-A Guide for Users", IFS Working Paper 88/15, London.

Arellano, M. and S. Bond (1991): "Some Test of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 58, 277-297.

Arellano, M. and O. Bover (1994): "Another Look at the Instrumental-Variable Estimation of Error-Components Models", *Journal of Econometrics*, forthcoming.

Attanasio, O.P. and G. Weber (1992): "Consumption Growth and Excess Sensitivity to Income: Evidence from US Micro Data", IFS Working Paper 92/15, London.

Attanasio, O.P. and G. Weber (1993): "Consumption Growth, the Interest Rate and Aggregation", *Review of Economic Studies*, 60, 631-649.

Attfield, C.L.F. and M. Browning (1985): "A Differential Demand System, Rational Expectations and the Life Cycle Hypothesis", *Econometrica*, 53, 31-48.

Blundell, R., M. Browning and C. Meghir (1994): "Consumer Demand and the Life-Cycle Allocation of Household Expenditures", *Review of Economic Studies*, 61, 57-80.

Blundell, R., V. Fry and C. Meghir (1986): "Preference Restrictions in Empirical Models of Life-Cycle Behaviour under Uncertainty", Department of Economics, UCL, London, mimeo.

Browning, M. (1989): "The Intertemporal Allocation of Expenditure on Non-Durables, Services and Durables", *Canadian Journal of Economics,* 22, 22-36.

Browning, M. (1992): "Children and Household Economic Behavior", *Journal of Economic Literature,* vol. XXX, 1434-1475.

Browning, M., A. Deaton and M. Irish (1985): "A Profitable Approach to Labor Supply and Commodity Demand over the Life-Cycle", *Econometrica*, 53, 503-543.

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

Chamberlain, G. (1984): "Panel Data" in Z. Griliches and M.D. Intrilligator (eds.), *Handbook of Econometrics* vol. 2, Elsevier.

Cuenca, J.A. (1991): "La construccion de variables financieras para el análisis del sector monetario de la Economía Española", Banco de España, Servicio de Estudios, mimeo.

Deaton, A. (1985): "Panel Data from Time Series of Cross-Sections", *Journal of Econometrics*, 30, 109-126.

Deaton, A. (1991): "Household Saving in LDC's: Credit Markets, Insurance, and Welfare", Research Program in Development Studies, Center of International Studies, Princeton University, Processed.

Deaton, A. (1992): *Understanding Consumption*, Clarendon Press.

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

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

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

Hall, R.E. and F.S. Mishkin (1982): "The Sensitivity of Consumption to

Transitory Income: Estimates from Panel data on Households", *Econometrica*, 50, 461-481.

Hansen, L.P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, 50, 1029-1054.

Hayashi, F. (1985): "The permanent Income Hypothesis and Consumption Durability: Analysis based on Japanese Panel data", *The Quarterly Journal of Economics*, 100, 1083-1113.

Holtz-Eakin, D., W. Newey and H. Rosen (1988): "Estimating Vector Autoregressions with Panel Data", *Econometrica*, 56, 1371-1395.

Hsiao, C. (1986): Analysis of Panel Data. Cambridge University Press.

MaCurdy, T.E. (1981): "An Empirical Model of Labor Supply in a Life-Cycle Setting", *Journal of Political Economy*, 89, 1059-1085.

MaCurdy, T.E. (1983): "A simple Scheme for Estimating an Intertemporal Model of Labour Supply and Consumption in the Presence of Taxes and Uncertainty", *International Economic Review*, 24, 265-289.

Moffitt, R. (1993): "Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross Sections", *Journal of Econometrics*, 59, 99-123.

Nickell, S. (1981): "Biases in Dynamic Models with Fixed Effects", *Econometrica*, 49, 1417-1426.

Nickell, S.J. (1985): "Intertemporal Substitution in Labour Supply", Econometric Society World Congress, Invited Symposium, Boston.

Runkle D.E. (1991):"Liquidity constraints and the Permanent Income Hypothesis. Evidence from Panel Data", *Journal of Monetary Economics*, 27, 73-98.

Verbeek, M. and T. Nijman (1992): "Can Cohort Data be Treated as Genuine Panel Data?", *Empirical Economics*, 17, 9-23.

Verbeek, M. and T. Nijman (1993): "Minimun MSE Estimation of a Regression Model with Fixed Effects from a Series of Cross-Sections", *Journal of Econometrics*, 59, 125-136.

Zeldes, S.P. (1989): "Consumption and Liquidity Constraints: an Empirical Investigation", *Journal of Political Economy*, 97, 305-346.