

BONN ECON DISCUSSION PAPERS

Discussion Paper 3/2004

Aggregate Behavior and Microdata

by

Werner Hildenbrand, Alois Kneip

February 2004



Bonn Graduate School of Economics
Department of Economics
University of Bonn
Adenauerallee 24 - 42
D-53113 Bonn

The Bonn Graduate School of Economics is
sponsored by the

Deutsche Post  World Net

MAIL EXPRESS LOGISTICS FINANCE

Aggregate Behavior and Microdata

Werner Hildenbrand* and Alois Kneip†

Abstract

It is shown how one can effectively use microdata in modelling the change over time in an aggregate (e.g. mean consumption expenditure) of a large and heterogeneous population. The starting point of our aggregation analysis is a specification of explanatory variables on the micro-level. Typically, some of these explanatory variables are observable and others are unobservable. Based on certain hypotheses on the evolution over time of the joint distributions across the population of these explanatory variables we derive a decomposition of the change in the aggregate which allows a partial analysis: to isolate and to quantify the effect of a change in the *observable* explanatory variables. This analysis does not require an explicit treatment of the unobservable variables.

JEL-Classification: D12, D12, E21

*Forschungsgruppe Hildenbrand, University of Bonn, Lennéstr 37, D-53113 Bonn, Germany. Tel. +49-228-737993, Fax +49-228-736102, e-mail: fgh@wiwi.uni-bonn.de

†Lehrstuhl für Statistik und Mathematik, FB Rechts- und Wirtschaftswissenschaften, University of Mainz, Jakob-Welder-Weg 9, D-55128 Mainz, Germany. Tel. +49-6131-3922979, Fax +49-6131-3923971, e-mail: kneip@wiwi.uni-mainz.de

1 Introduction

It is our goal to model the change over time in an aggregate of a large and heterogeneous population. Examples of such aggregates are the mean consumption expenditure across a population of households or the mean labor demand of a production sector. More precisely, we are looking for explanatory variables for the change $C_t - C_{t-1}$ or the relative change $(C_t - C_{t-1})/C_{t-1}$, where C_t denotes the aggregate in period t .

A microeconomist will argue that decisions are taken by the micro-units and therefore the starting point must be a specification of a complete set of explanatory variables for the relevant response variable *on the micro-level*. The choice of such explanatory variables is based either on experimental economics or on microeconomic theory, i.e. on a model of behavior. In neoclassical microeconomics behavior is modelled by an intertemporal (utility) maximization problem under uncertainty. Then the parameters which define this maximization problem constitute the explanatory variables. An explicit example in the case of consumption expenditure is presented in the appendix. This leads to a micro-relation which can be represented in the form

$$c_t^h = c(x_t^h), \quad h \in H_t, \quad (1.1)$$

where c_t^h and x_t^h denote the response and the vector of explanatory variables, respectively, of the micro-unit h of the underlying population H_t in period t . In this notation the set of explanatory variables determines uniquely the response. The functional relationship c therefore does not depend on h and t , since the complete set of explanatory variables contains everything that is relevant for the decision. Thus x_t^h contains typically unobservable variables such as individual preferences. In empirical work some of the unobservable variables are often summarized by a random term. Note that x_t^h in period t might also include components which refer to periods $t-1, t-2, \dots$.

Given a micro-relation (1.1) the mean response $C_t = \frac{1}{\#H_t} \sum_{h \in H_t} c_t^h$ can be written as

$$C_t = \int c(x) \text{distr}(x | H_t), \quad (1.2)$$

where $\text{distr}(x | H_t)$ denotes the joint distribution of the explanatory variables x_t^h across the population H_t . Consequently, $\text{distr}(x | H_t)$ takes the role of an explanatory "variable" at the aggregate level. Obviously, (1.2) does not provide a feasible basis for applied analysis. The aim of aggregation theory is to simplify (1.2) and to find certain characteristics $\chi_t = (\chi_{t1}, \chi_{t2}, \dots)$ of $\text{distr}(x | H_t)$ as well as a function F such that

$$C_t = \int c(x) \text{distr}(x | H_t) \equiv F(\chi_t) \quad \text{for every period } t \quad (1.3)$$

Whether or not such a simplification is possible for a moderate number of characteristics and a simple function F depends on the functional form of the micro-relation c and/or the way the distributions $\text{distr}(x | H_t)$ evolve over time (see, e.g., Nataf (1948), Gorman (1953), Malinvaud (1993), Stoker (1993) and Blundell and Stoker (2002)). For example, if c were linear then $F \equiv c$ and χ_t is equal to the mean X_t of x_t^h across the population H_t . Typically, however, the micro-relation $c(\cdot)$ is not linear in all micro-specific explanatory variables. Even the simple Linear Expenditure System (Stone (1954)) in demand analysis is not linear in all explanatory variables. Simple aggregation with $\chi_t = X_t$ is also possible if c has a complex nonlinear structure. Indeed, if, for example, the joint distribution of the centered variables $\tilde{x}_t^h = x_t^h - X_t$ is time-invariant, i.e. $\text{distr}(\tilde{x} | H_t)$ is independent of t , say equal to a distribution μ , then again $\chi_t = X_t$ but $F(X) = \int c(\tilde{x} + X)\mu(d\tilde{x})$. In this case the functional form of F may be completely different from that of c .

More generally, let $T_\chi(x)$ denote an invertible transformation of x which is defined in terms of parameters χ , where χ are certain characteristics of the distribution of the explanatory variables x . For example, if $\chi = X$ then $T_X(x) := x - X$ (centered variable) or $T_X = x/X$ (scaled variable). If $\chi = (X, \Sigma)$, where Σ is the non-singular covariance matrix of the distribution of x , then $T_{X,\Sigma}(x) = \Sigma^{-1/2}(x - X)$ (standardized variable). If the distribution of the transformed explanatory variables $\tilde{x}_t^h = T_{\chi_t}(x_t^h)$ across the population H_t is time-invariant, say equal to a distribution μ , then aggregation is possible with

$$F(\chi) = \int c(T_\chi^{-1}(\tilde{x}))\mu(d\tilde{x}) \quad (1.4)$$

Since, in general, a functional form of the micro-relation c is not inherited by the aggregate relation F , it is not useful to start with a specific functional form of c . Our approach is based on the idea of standardizing variables leading to an aggregate relation similar to (1.4). Details are given in sections 2- 4. However, the major steps of our analysis can easily be explained in the simple hypothetical case of time-invariance of the distributions of centered variables. Then (1.4) becomes

$$F(X) = \int c(\tilde{x} + X)\mu(d\tilde{x}) \quad (1.5)$$

In many applications it is justified to assume that aggregate explanatory variables X_{ti} change slowly over time in the sense that either $(\frac{X_{ti} - X_{t-1,i}}{X_{t-1,i}})^2$ or $(X_{ti} - X_{t-1,i})^2$ are negligible. Let the first n variables be of the first type and the remaining m variables of the second type. A first order approximation of the function F at $X = X_{t-1}$ then yields

$$\begin{aligned} \frac{C_t - C_{t-1}}{C_{t-1}} &= \sum_{i=1}^n \beta_{t-1}^i \left(\frac{X_{ti} - X_{t-1,i}}{X_{t-1,i}} \right) + \sum_{j=n+1}^{n+m} \beta_{t-1}^j (X_{ti} - X_{t-1,i}) \\ &+ \text{terms of second order in } \left(\frac{X_{ti} - X_{t-1,i}}{X_{t-1,i}} \right)^2 \text{ and } (X_{ti} - X_{t-1,i})^2 \end{aligned} \quad (1.6)$$

where

$$\beta_{t-1}^i = \frac{X_{t-1,i}}{C_{t-1}} \int \partial_i c(x) \text{distr}(x|H_{t-1}),$$

$$\beta_{t-1}^j = \frac{1}{C_{t-1}} \int \partial_j c(x) \text{distr}(x|H_{t-1})$$

and ∂_k denotes the partial derivative with respect to the k -th variable. Here we use the fact that (1.5) implies $\partial_k F(X_{t-1}) = \int \partial_k c(x) \text{distr}(x|H_{t-1})$.

If all explanatory variables on the micro-level **were** (!) observable and if the distribution $\text{distr}(x|H_{t-1})$ **were** non-degenerate and spread (i.e. the population is large and heterogeneous in all explanatory variables), then individual information on $\{c_{t-1}^h, x_{t-1}^h\}_{h \in H_{t-1}}$, i.e. micro-data in period $t-1$, would give us knowledge about the micro-relation c , since by (1.1) $c_{t-1}^h = c(x_{t-1}^h)$, $h \in H_{t-1}$. Consequently, in this hypothetical case the partial derivatives $\partial_k F(X_{t-1})$, and hence the coefficients β_{t-1} in (1.6) can be related to micro-data in period $t-1$.

Unfortunately, however, not all explanatory variables on the micro-level are observable! At this point it is important to make a distinction between observable and unobservable explanatory variables on the micro-level. We denote by y_t^h the vector of *observable* and micro-specific variables, assuming that the population is heterogeneous in y in the sense that the distribution of y_t^h across the population H_t is non-degenerate. The remaining explanatory variables are either non observable, denoted by the vector v_t^h , or observable, yet not micro-specific, denoted by the vector p_t . With this notation $X_t \equiv (Y_t, V_t, p_t)$ and therefore we obtain $C_t = F(X_t) = F(Y_t, V_t, p_t)$.

The important point now is that those coefficients β_{t-1} in (1.6) which correspond to the observable and micro-specific explanatory variables can still be related to micro-data in period $t-1$. Indeed, consider for example the first partial derivative $\partial_1 F(X_{t-1})$. Assume that y_{t-1}^h and v_{t-1}^h do not correlate across the population H_{t-1} (this assumption is weakened in the paper) then one obtains

$$\partial_1 F(X_{t-1}) = \int \partial_1 \bar{c}_{t-1}(y, p_{t-1}) \text{distr}(y|H_{t-1})$$

where $\bar{c}_{t-1}(y, p_{t-1}) := \int c(y, v, p_{t-1}) \text{distr}(v|H_{t-1}(y))$ is the regression function of c_{t-1}^h given y which can be estimated from individual observations in period $t-1$. Consequently, the partial derivative and hence the coefficient β_{t-1}^1 can be determined from suitable micro-data in period $t-1$. This can be done separately for each period without specifying the structure of unobservables.

This observation plays a key role in our analysis. By generalizing the above simple example we demonstrate that there are explicit ways to incorporate data on the *individual* level into building and analyzing aggregate models.

The practical importance of this point is best seen when first considering the standard time series approach used in applied work to analyze aggregate models. In this context model building is done from a point of view quite different from the above approach that is based on aggregation. The time series $\{C_t\}$ is considered as a realization of a stochastic process and emphasis lies on constructing a valid time series model which links $\{C_t\}$ to an observable multivariate time series $\{Z_t\}$. If relative changes of C_t are of primary interest, such models take the form

$$\log C_t - \log C_{t-1} = \Delta \log C_t = \sum_j \theta_j Z_{tj} + \epsilon_t \quad (1.7)$$

Specification of model components Z_{tj} usually relies on microeconomic reasoning. Typically, means of micro-variables x_t^h are taken as explanatory variables on the aggregate level, and some of Z_{tj} then correspond to components of $\log X_t - \log X_{t-1}$ or $X_t - X_{t-1}$. Often also higher lags $X_{t-1} - X_{t-2}, \dots$ or lags of the mean response $\Delta \log C_{t-1}, \dots$ will be incorporated into the Z_t . Frequently, the error term ϵ_t will be modelled as white noise, but sometimes also a more complex MA-structure is assumed.

Quite obviously at this point there is a formal similarity between (1.6) and (1.7), since relative differences as $(C_t - C_{t-1})/C_{t-1}$ are usually well approximated by differences in logarithms as $\log C_t - \log C_{t-1}$.

On the other hand, in many situations aggregate models (1.7) will also include terms which are not related to any explanatory variable at the micro-level. The reason is that establishing a valid model (1.7) necessarily involves a stochastic analysis of properties of the underlying time series. Additional variable, for example error correction terms, may have to be introduced in order to achieve a proper modelling of the stochastic behavior (for a comprehensive survey of modern time series theory see, for example, Greene (2003)). In the context of consumption analysis important work in this direction is, for example, Davidson et al (1978) and Deaton (1992)).

Time series analysis is a powerful tool but it also has some limitations. Model building is usually not easy and has to rely on a number of specific assumptions. It is well known that in many cases quite different looking models can lead to very similar fits. In principle misspecification of a single component $\theta_j Z_{tj}$ or of the error term may result in inconsistent parameter estimates and invalid economic conclusions. Further problems arise when fitting a highly parametrized model (1.7) to comparably short economic times series by using least squares, maximum likelihood, etc. Due to the possibility of overfitting and finite sample effects, considerable care may be necessary when interpreting model fits or estimated parameters.

In this paper we do not intend to replace time series modelling but to introduce an

additional tool which allows to isolate the effects of some important observable variables, and which may help to achieve a still greater accuracy of macroeconomic modelling by incorporating the rich information which is contained in micro-data.

Our approach concentrates on the micro specific observable variables y_t^h . Since $x_t^h = (y_t^h, v_t^h, p_t)$ the general relation (1.1) becomes

$$C_t = \int c(y, v, p_t) \text{distr}(y, v | H_t)$$

Thus any change in C_t is caused by a change in $\text{distr}(y, v | H_t)$ and/or p_t . Our approach now relies on an explicit modelling of the evolution of the distribution which generalizes our simple example given above. As before, we assume that time changes are not arbitrary, but occur in a "structurally stable" way. This concept is explained in detail in Sections 2 and 3. It is also motivated there that it will often be possible to parametrize changes of the distribution of y_t^h in terms of changes of the corresponding mean values m_t and covariance matrices Σ_t over the population. We will show that structural stability allows to find a *local* solution of the aggregation problem without specifying a functional form of the micro-relation. By applying a first order approximation generalizing (1.6) it is then derived in Section 4 that the following decomposition holds:

$$\Delta \log C_t \approx \frac{C_t - C_{t-1}}{C_{t-1}} = \beta_{t-1}^T (m_t - m_{t-1}) + \text{trace}(\Gamma_{t-1}(\Sigma_t^{1/2} \Sigma_{t-1}^{-1/2} - \mathbb{I})) + \sum_j \theta_j Z_{tj}^* + \text{error} \quad (1.8)$$

The effect of changes in the distribution of the observable and micro specific variables y_t^h is captured by the first two terms on the right hand side of (1.8), where β_{t-1} and Γ_{t-1} are possibly time varying vectors and matrices of coefficients, respectively. The third term $\sum_j \theta_j Z_{tj}^*$ quantifies the influence of other explanatory variables corresponding to v_t^h and p_t . A more specific form of this remainder term is given in the proposition of Section 4.

The crucial point now is that our theory relates the coefficients β_{t-1} and Γ_{t-1} to individual data. They can be determined from derivatives of suitable regression functions which can be estimated from micro observations. A precise definition of the coefficients is given in the Proposition. In principle various kinds of micro data can be used (cross-section, panel or experimental data) provided that the data contain the appropriate variables and that the underlying samples are representative for the population in every time period.

Of course, such micro data also allow to compute means m_t and covariance matrices Σ_t . Therefore, the complete terms $\beta_{t-1}^T (m_t - m_{t-1}) + \text{trace}(\Gamma_{t-1}(\Sigma_t^{1/2} \Sigma_{t-1}^{-1/2} - \mathbb{I}))$ can be estimated from the micro data without invoking any time series fitting. This approach has several attractive features.

- A partial analysis is possible, and the effects of changes in the observable micro-specific variables can be isolated without specifying the structure of the remaining terms $\sum_j \theta_j Z_{tj}^*$ or of the error. This is not possible in a pure time series model, where a consistent estimation of parameters always requires the specification of a complete model.
- Using individual data, calculation of $\beta_{t-1}^T(m_t - m_{t-1}) + \text{trace}(\Gamma_{t-1}(\Sigma_t^{1/2}\Sigma_{t-1}^{-1/2} - \mathbb{I}))$ *does not use any information* about the structure of the time series $\{\Delta \log C_t\}$. Since *no model fitting* takes place, this may provide more precise information about the explanatory power of the observable micro-specific explanatory variables.
- In our model the coefficients β_t and Γ_t are behavioral parameters characterizing the population in period t . Therefore, there is no a priori reason that they will be time invariant. Estimation from micro-data separately for each period will automatically adapt to possible time changes in these coefficients.

This approach is illustrated in Section 5 for the case of consumption expenditure. Using cross-section data from the UK-Family Expenditure Survey, we perform a partial analysis as described above by relying on current income and assets as the observable micro-specific variables. It turns out that these variables explain an essential part of the observed changes in mean consumption.

Of course, a complete model requires to specify the remainder terms $\sum_j \theta_j Z_{tj}^*$ as well as the stochastic error structure. However, different from (1.7), it is only necessary to model the stochastic structure of the *residual series*

$$\Delta \log C_t - \beta_{t-1}^T(m_t - m_{t-1}) - \text{trace}(\Gamma_{t-1}(\Sigma_t^{1/2}\Sigma_{t-1}^{-1/2} - \mathbb{I}))$$

and a lower number of components will have to be fitted from the time series. Note that for prediction purposes the resulting series $\beta_{t-1}^T(m_t - m_{t-1}) + \text{trace}(\Gamma_{t-1}(\Sigma_t^{1/2}\Sigma_{t-1}^{-1/2} - \mathbb{I}))$ may also be analyzed from a time series point of view in order to forecast future values. We will not consider these points in detail, since our paper concentrates on the role of the observable micro-specific variables.

The paper is organized as follows. Our setup is described in Section 2, while in Section 3 we develop the concept of structural stability of distributions. The main theoretical result is presented in Section 4. Section 5 contains an empirical study which applies our theory to modelling consumption expenditure. In the appendix we give an explicit example of a micro relation as used in the paper.

2 Definitions and Notation

The starting point of aggregation analysis is a specification of a complete set of explanatory variables on the micro-level for a certain explicandum (response variable), for example, consumption expenditure of a household or labor demand of a production unit. The choice of the explanatory variables is based either on *experimental economics* or on *microeconomic theory*, that is to say, on a model of behavior. In neoclassical microeconomics behavior is modelled by an intertemporal (utility) maximization problem under uncertainty. Then the parameters which define this maximization problem are the explanatory variables (see Appendix for an explicit example).

Typically, some of the explanatory variables are observable and others are unobservable. For a micro-unit h in period t we denote by y_t^h the vector of *observable* and *micro-specific* variables (e.g., labor income or wealth). The remaining variables are either *unobservable*, denoted by the vector v_t^h (e.g., expected future labor income), or observable, yet not micro-specific, denoted by the vector p_t (e.g., current prices or interest rates).

Note that the vector of explanatory variables y_t^h, v_t^h, p_t in period t might contain components which refer to periods $t - 1, t - 2, \dots$, e.g., past labor income.

We assume that the vector of explanatory variables contains everything that is relevant for the decision. Then, the explicandum (response variable), denoted by c_t^h , is uniquely determined by the explanatory variables (y_t^h, v_t^h, p_t) , that is to say

$$c_t^h = c(y_t^h, v_t^h, p_t). \quad (2.1)$$

We do not need any knowledge about the functional form of this relationship c . We shall assume, however, that c is continuously differentiable in all variables.

The *population* of micro-units in period t is denoted by H_t . Then, the mean response $\frac{1}{\#H_t} \sum_{h \in H_t} c_t^h$ of the population H_t is given by

$$C_t = \int c(y, v, p_t) \text{distr}(y, v | H_t) \quad (2.2)$$

where $\text{distr}(y, v | H_t)$ denotes the joint distribution of the micro-specific explanatory variables (y_t^h, v_t^h) across the population H_t . Analogously, $\text{distr}(y | H_t)$ denotes the observable distribution of y_t^h across H_t .

In addition to the explanatory variables (y, v, p) in the micro-relation (1) we consider certain *observable* micro-specific *attributes* (socio-economic variables, e.g., household size or age of household head). Let $a = (a_1, a_2, \dots)$ denote a finite profile of such attributes. We

allow for a finite set \mathcal{A} of profiles. Let $\text{distr}(y, a | H_t)$ denote the observable joint distribution of (y_t^h, a_t^h) across the population H_t .

By $H_t(y, a)$ we denote the subpopulation of all micro-units in H_t with $y_t^h = y$ and $a_t^h = a$. Then, if $H_t(y, a) \neq \emptyset$, $\text{distr}(v | H_t(y, a))$ denotes the distribution of v_t^h across the subpopulation $H_t(y, a)$. Finally, we define the *regression function* $\bar{c}_t(\cdot, \cdot, p_t)$

$$\bar{c}_t(y, a, p_t) := \int c(y, v, p_t) \text{distr}(v | H_t(y, a)). \quad (2.3)$$

With this definition, the mean response can be written as

$$C_t = \int \bar{c}_t(y, a, p_t) \text{distr}(y, a | H_t). \quad (2.4)$$

For a finite population H_t the regression function is only defined for those variables (y, a) with $H_t(y, a) \neq \emptyset$, i.e., with (y, a) in the support of the distribution $\text{distr}(y, a | H_t)$, which is a finite set. The mathematical analysis is greatly simplified if one assumes that the regression function $c_t(y, a, p_t)$ is a smooth function in y . This requires¹ that the population H_t is “infinitely large” and heterogeneous in the observable explanatory variable y in the sense that the distribution $\text{distr}(y | H_t(a))$ is concentrated on an open domain in \mathbb{R}^n . To be simple and specific one might assume that the support of $\text{distr}(y | H_t(a))$ is equal to \mathbb{R}^n .

Remark: Why stratification by attribute profiles? The reason for introducing observable attributes in addition to the explanatory variables in (2.1) is to justify the assumption (hope) that, by stratifying on y and a , the subpopulation $H_t(y, a)$ becomes “homogeneous” in the unobservable explanatory variable v , either in the strong sense that $v_t^h = v_t(y, a)$ for all $h \in H_t(y, a)$ or, more generally, that the distributions $\text{distr}(v | H_t(y, a))$ are “structurally stable” (see Assumption 1). Furthermore, one might expect that the presence of a strong correlation between v_t^h and (y_t^h, a_t^h) across H_t reduces the time-dependence of $\text{mean}_{H_t(y, a)} v_t^h$, at least for some components of v . Note that time-invariance of $\text{mean}_{H_t(y, a)} v_{t, i}^h$ does not imply time-invariance of $\text{mean}_{H_t} v_{t, i}^h$. It implies however that the change on the aggregate level is caused by the change in the distribution $\text{distr}(y, a | H_t)$. We emphasize that time-invariance of $\text{mean}_{H_t(y, a)} v_{t, i}^h$ has an important consequence: the unobservable explanatory variable v_i does

¹If one insists on a formal mathematical definition, one considers a “continuum of economic agents”, i.e., a measure space (Ω, \mathcal{F}, P) of micro-units (e.g., $[0, 1]$ with Lebesgue measure). The population in period t is then defined by the measurable mappings Y_t, V_t , and A_t of Ω into $\mathbb{R}^n \times \mathbb{R}^m \times \mathcal{A}$, where $Y_t(\omega) = (y_{t, 1}^\omega, \dots, y_{t, n}^\omega)$, $V_t(\omega) = (v_{t, 1}^\omega, \dots, v_{t, m}^\omega)$, and $A_t(\omega) = a_t^\omega$.

The above distributions $\text{distr}(y | H_t)$ and $\text{distr}(y, v | H_t)$ are then defined as the image distribution of P with respect to the mapping Y_t and the mapping (Y_t, V_t) , respectively. The above distribution $\text{distr}(v | H_t(y, a))$ is defined as the conditional distribution of V_t given the mappings Y_t and A_t .

not have to be modelled explicitly since - as we shall show - its influence on the change over time in C_t is fully captured by the "cross-section effect" of the Proposition. Therefore, under time-invariance, one can avoid the delicate problem of postulating an "observable proxy" for an unobservable variable.

3 Structural Stability

With the notation of the last section one obtains for the mean response

$$C_t = \int \left[\int c(y, v, p_t) \text{distr}(v | H_t(y, a)) \right] \text{distr}(y, a | H_t).$$

Thus, given the micro-relation (1), the change over time in C_t is caused by the change over time in the distributions $\text{distr}(y, a | H_t)$ and $\text{distr}(v | H_t(y, a))$ as well as the vector p_t of non-micro-specific explanatory variables.

3.1 The change over time in $\text{distr}(y, a | H_t)$

We emphasize that the distribution $\text{distr}(y, a | H_t)$ is observable and therefore any assumption on the way how these distributions change over time can be falsified.

We shall first consider the change over time in the distribution $\text{distr}(y | H_t)$ of the observable micro-specific explanatory variables y .

Let m_t denote the vector of means of y_t^h across the population H_t , $m_{t,i} := \text{mean}_{h \in H_t} y_{t,i}^h$, and Σ_t the covariance matrix of y_t^h across H_t , $\Sigma_t := \left(\text{cov}_{h \in H_t} (y_{t,i}^h, y_{t,j}^h) \right)_{i,j}$. We assume that the population H_t is sufficiently heterogeneous in y_t^h in the sense that the covariance matrix is non-singular.

The *standardized distribution* of y_t^h across H_t is defined as the distribution of $\tilde{y}_t^h := \Sigma_t^{-\frac{1}{2}} (y_t^h - m_t)$ across H_t .

Thus, $\text{mean}_{H_t} \tilde{y}_t^h = 0$ and $\text{cov}_{H_t}(\tilde{y}_t^h, \tilde{y}_t^h) = \mathbb{I}$, the unit matrix.

Hypothesis 1: Structural Stability of $\text{distr}(y | H_t)$

The standardized distribution of y_t^h across H_t changes sufficiently slowly over time in the sense that the standardized distributions can be considered as time-invariant for two periods s and t that are close to each other.

Obviously, Hypothesis 1 is trivially satisfied if $\text{distr}(y | H_t)$ are multivariate normal distributions. We remark that Hypothesis 1 does not model the dynamics of $\text{distr}(y | H_t)$. Time-invariance of the standardized distributions implies that $\text{distr}(y | H_t)$ in period t is determined by m_t, Σ_t , and $\text{distr}(y | H_s)$ in period s , since, as one easily verifies,

$$\int f(y) \text{distr}(y | H_t) = \int f\left(\Sigma_t^{-\frac{1}{2}} \Sigma_s^{-\frac{1}{2}}(y - m_s) + m_t\right) \text{distr}(y | H_s) \quad (3.1)$$

for any integrable function $f(y)$.

Remark: In our application to consumption expenditure in Section 5 we consider two observable micro-specific explanatory variables; $\eta_{t,1}^h$ income from labor and $\eta_{t,2}^h$ income from assets (property). It is well-known that the observed income distributions of actual economies evolve over time in a surprisingly “structurally stable” way. Income and wealth distributions have been studied extensively in the literature, starting with Pareto (1897). For recent references see Atkinson and Bourguignon (2000). The empirical studies support well Hypothesis 1 for $y_{t,1} := \log \eta_{t,1}^h$ and $y_{t,2} := \log \eta_{t,2}^h$. In this case, the parameter σ_t^2 (variance of $\log \eta_t^h$) can be interpreted as a measure of income dispersion (inequality). For a symmetric log-income distribution the parameter m_t is equal to the logarithm of median income.

In the literature (e.g., Malinvaud (1993)) one considers sometimes a stronger concept of “structural stability”; the time-invariance of the *relative* income distribution, which is defined as $\text{distr}(\eta_t^h / \bar{\eta}_t | H_t)$, where $\bar{\eta}_t$ denotes mean income across H_t . In this case the dispersion σ_t is constant. For time-invariant σ_t one easily shows that time-invariance of the standardized log-income distribution is equivalent to time-invariance of the relative income distribution. We remark, that the concept of “mean-scaled” income distribution as formulated by Lewbel (1990) and (1992) is closely related to the time-invariance of the standardized log income distribution.

For statistical estimates of standardized log income distributions based on FES-data we refer to Hildenbrand and Kneip (1999) and Hildenbrand, Kneip and Utikal (1998).

Next we consider the observable attribute-profile distribution across the subpopulation $H_t(y)$, that is to say, $\text{distr}(a | H_t(y))$. The shape of these distributions, as well as their dependence on y and t , crucially depend on the nature of the attributes, for example, household size or age of household head. Typically, $\text{distr}(a | H_t(y))$ depends on the vector y of the observable micro-specific variables and is not time-invariant (for an example, see Hildenbrand and Kneip (1999)). Obviously, it is problematic to model the change over time in the joint distribution $\text{distr}(y, a | H_t)$, consistent with Hypothesis 1, without being specific about the nature of the observable micro-specific variable y and the observable attribute profile a . Since in this theoretical part of our analysis we want to avoid considering particular

examples, we consider the case where the attribute profile distribution changes much slower than the distribution of the micro-specific explanatory variables y .

This motivates the following

Hypothesis 2:

For two periods s and t that are close to each other, the attribute-profile distribution $\text{distr}(a | H_t(y^t))$ across the subpopulation $H_t(y^t)$ can be considered as equal to the attribute profile distribution $\text{distr}(a | H_s(y^s))$ across the subpopulation $H_s(y^s)$ if y^t and y^s are in the “same position” in the standardized y -distribution of period t and s , respectively, i.e.,

$$\Sigma_t^{-\frac{1}{2}}(y^t - m_t) = \Sigma_s^{-\frac{1}{2}}(y^s - m_s).$$

One easily shows that Hypotheses 1 and 2 imply

$$\int f(y, a) \text{distr}(y, a | H_t) = \int f\left(\Sigma_t^{\frac{1}{2}}\Sigma_s^{-\frac{1}{2}}(y - m_s) + m_t, a\right) \text{distr}(y, a | H_s) \quad (3.2)$$

for any integrable function $f(y, a)$. It is this consequence of Hypotheses 1 and 2 which is used in the proof of our Proposition.

3.2 The change over time in $\text{distr}(v | H_t(y, a))$

In contrast to Subsection 3.1, the distribution $\text{distr}(v | H_t(y, a))$, whose change over time has to be modelled, is now unobservable. Thus, any assumption on the change over time in these distributions is speculative (purely theoretical).

The change over time in the regression function

$$\bar{c}_t(y, a, p_t) = \int c(y, v, p_t) \text{distr}(v | H_t(y, a))$$

is caused by the change in $\text{distr}(v | H_t(y, a))$ and p_t . A trivial way to simplify the time dependence of \bar{c}_t would be to assume that the subpopulation $H_t(y, a)$ is *homogeneous* in the unobservable explanatory variable v_t^h , i.e., $v_t^h = v_t(y, a)$ for every $h \in H_t(y, a)$. Then one obtains

$$\bar{c}_t(y, a, p_t) = c(y, v_t(y, a), p_t).$$

If one views the subpopulation $H_t(y, a)$ as *heterogeneous* in the unobservable micro-specific explanatory variable v_t^h , for example, in the sense that the covariance matrix of v_t^h across $H_t(y, a)$ is non-singular, then one might assume - analogously to Hypothesis 1 of Structural Stability - that the standardized distributions of v_t^h across $H_t(y, a)$ are locally

time-invariant. However, to simplify the analysis (mainly the notation) we shall assume a stronger form of Structural Stability. Instead of the standardized distribution we consider the *centered* distribution which is defined as the distribution of $v_t^h - v_t(y, a)$ across the subpopulation $H_t(y, a)$ where $v_t(y, a)$ denotes the mean of v_t^h across $H_t(y, a)$.

Assumption 1: Structural Stability of $\text{distr}(v | H_t(y, a))$

The centered distribution of v_t^h across $H_t(y, a)$ changes sufficiently slowly over time in the sense that these distributions can be considered as time-invariant for two periods s and t that are close to each other.

Finally, we need in the proof of our Proposition an assumption which specifies how the mean of the unobservable variable $v_{t,i}^h$ across the subpopulation $H_t(y, a)$ depends on y and t . This, obviously, depends on the nature of the unobservable micro-specific explanatory variable v_i .

The most favorable case for our analysis would prevail if one could view the mean $v_{t,i}(y, a)$ as time-invariant (or sufficiently slowly changing). Recall that time-invariance of $v_{t,i}(y, a)$ does not imply time-invariance of $\text{mean}_{H_t} v_{t,i}^h$ (i.e., on the aggregate level). An example might be a structural parameter of the utility function by assuming that for a micro-unit h this parameter is determined by y and a .

On the other hand, one might consider the case where $v_{t,i}^h$ and y_t^h do not correlate across the subpopulation $H_t(a)$. This case trivially prevails if one assumes that the subpopulation $H_t(a)$ is homogeneous in v_t^h (an assumption which is usually made in demand analysis). Then $v_{t,i}(y, a)$ does not depend on y and is equal to $\text{mean}_{H_t(a)} v_{t,i}^h =: v_{t,i}(a)$, which we allow to change over time (otherwise we are back in the above case). The cause of this change is exogenous in our model. The growth rate (not the level!) of future labor income as anticipated in period t might be an example.

The above discussion motivates the following

Assumption 2: Additive Factorization

The mean of v_t^h across $H_t(y, a)$ can be factorized by

$$\text{mean}_{H_t(y,a)} v_t^h =: v_t(y, a) = \varphi(y, a) + \psi(t, a)$$

where the function φ is continuously differentiable in y .

Remark: Depending on the nature of the unobservable explanatory variable it might be

more natural to consider a multiplicative (or even more complex) factorization. To be simple and specific we have chosen the additive form. It will become clear in the proof of our Proposition how the arguments have to be modified in the case of an alternative factorization.

4 The change in mean response C_t

Let us first recall some definitions that are needed in formulating our main result. As in the previous sections, let m_s and Σ_s denote the mean and the covariance matrix, respectively, of the vector y_s^h of the observable and micro-specific explanatory variables across the population H_s . Define $v_s(y, a)$ as the mean of the unobservable explanatory variables v_s^h across the subpopulation $H_s(y, a)$, $v_s(a) := \text{mean}_{H_s(a)} v_s^h = \int v_s(y, a) \text{distr}(y | H_s(a))$ and $v_{t,s}(a) := \int v_t(y, a) \text{distr}(y | H_s(a))$.

Proposition: *Let the micro-relation (2.1) and the regression function (2.3) be continuously differentiable in the explanatory variables y, v , and p . Then Hypotheses 1 and 2 and Assumptions 1 and 2 imply that for two periods s and t that are close to each other the relative change in the mean response C_t can be decomposed in the following form*

$$\begin{aligned} \frac{C_t - C_s}{C_s} &= \beta_s^T (m_t - m_s) + \text{trace} \left[\Gamma_s (\Sigma_t^{\frac{1}{2}} \Sigma_s^{-\frac{1}{2}} - \mathbb{I}) \right] \\ &+ \int \left[(\delta_s^a)^T (v_{t,s}(a) - v_s(a)) \right] \text{distr}(a | H_s) \\ &+ \theta_s^T (p_t - p_s) \\ &+ \text{terms of second order in } \|m_t - m_s\|^2, \\ &\quad \|\Sigma_t^{\frac{1}{2}} \Sigma_s^{-\frac{1}{2}} - \mathbb{I}\|^2, \|v_{t,s}(a) - v_s(a)\|^2, \text{ and } \|p_t - p_s\|^2 \end{aligned}$$

where the vector β_s and matrix Γ_s of coefficients are defined by

$$\beta_s := \frac{1}{C_s} \int \partial_y \bar{c}_s(y, a, p_s) \text{distr}(y, a | H_s)$$

and

$$\Gamma_s := \frac{1}{C_s} \int (y - m_s) [\partial_y \bar{c}_s(y, a, p_s)]^T \text{distr}(y, a | H_s).$$

Remark: The effect on the mean response C_t of the change in the distribution of the observable and micro-specific variables y_t^h is captured by the term

$$\beta_s^T (m_t - m_s) + \text{trace} \left[\Gamma_s (\Sigma_t^{\frac{1}{2}} \Sigma_s^{-\frac{1}{2}} - \mathbb{I}) \right].$$

The vector β_s and the matrix Γ_s of coefficients are directly related to data on the micro-level in period s . Consequently, they do not depend on the postulated micro-relation. By definition these coefficients are mean derivatives of observable regression functions. Therefore they can be estimated separately from cross-section data in every period. Empirical results will be given in Section 5.

We emphasize that β_s and Γ_s are dependent on the chosen set of attribute profiles \mathcal{A} . Indeed, if one conditions on attribute profiles $a \in \mathcal{A}$ one obtains

$$\beta_s^{\mathcal{A}} = \frac{1}{C_t} \int \left[\int \partial_y \bar{c}_s(y, a, p_s) \text{distr}(a | H_s(y)) \right] \text{distr}(y | H_s).$$

If one does not condition at all on attribute profiles, $\mathcal{A} = \emptyset$, then one obtains

$$\beta_s^{\emptyset} = \frac{1}{C_t} \int \partial_y \tilde{c}_s(y, p_s) \text{distr}(y | H_s)$$

where $\tilde{c}_s(y, p_s) = \int c(y, v, p_s) \text{distr}(v | H_s(y)) = \int \bar{c}_s(y, a, p_s) \text{distr}(a | H_s(y))$.

Hence

$$\beta_s^{\emptyset} = \frac{1}{C_t} \int \left[\partial_y \int \bar{c}_s(y, a, p_s) \text{distr}(a | H_s(y)) \right] \text{distr}(y | H_s).$$

Consequently, if $\text{distr}(a | H_s(y))$ depends on y , which typically is the case, then $\beta_s^{\mathcal{A}} \neq \beta_s^{\emptyset}$.

The remaining terms in the Proposition which capture the effect of the change in the unobservable and micro-specific variables v_i naturally also depend on the chosen set \mathcal{A} of attribute profiles. As explained above, the aim of conditioning on $a \in \mathcal{A}$ is to make either these terms negligible or, at least, independent of the change in the distribution of observable explanatory variable y . For example, if one has good reasons to postulate (believe) that for a certain unobservable explanatory variable, say $v_{t,i}^h$, the mean across the subpopulation $H_t(y, a)$ is time-invariant, then the i -th component of the vectors $v_{t,s}(a)$ and $v_s(a)$ are equal. Consequently, the corresponding term $\delta_{s,i}^a(v_{t,s,i}(a) - v_{s,i}(a))$ in the Proposition is zero. Often the unobservable parameters of the utility function are treated this way. For other examples, see the Appendix. Alternatively, one might assume that the mean $v_{s,i}(y, a)$ does not depend on y . (The growth rate of anticipated future labor income might be an example). In this case the corresponding term in the Proposition is not zero, yet it is not effected by the change in the distribution of the observable and micro-specific variables y_s^h . Consequently, if for a given set \mathcal{A} of attribute profiles $v_{s,i}(y, a)$ either is time-invariant or does not depend on y , then the effect on the mean response C_t of a change in the distribution of observable and micro-specific variables can be fully isolated and quantified. Whether such a partial (incomplete) analysis of the relative change in C_t explains an essential part of the observable change in C_t is, of course, an empirical question, which is studied in the case of consumption expenditure in Section 5, where we also argue that terms of second order can be neglected.

Proof. By definition

$$\begin{aligned} C_t &:= \int c(y, v, p_t) \text{distr}(y, v | H_t) \\ &= \int \bar{c}_t(y, a, p_t) \text{distr}(y, a | H_t) \end{aligned} \quad (4.1)$$

with

$$\begin{aligned} \bar{c}_t(y, a, p_t) &:= \int c(y, v, p_t) \text{distr}(v | H_t(y, a)) \\ &= \int c(y, \tilde{v} + v_t(y, a), p_t) \text{distr}(\tilde{v} | H_t(y, a)) \\ &\quad (\text{recall } \tilde{v}_t^h := v_t^h - v_t(y, a) \text{ denotes the centered variable}) \\ &= \int c(y, \tilde{v} + v_t(y, a), p_t) \text{distr}(\tilde{v} | H_s(y, a)) \end{aligned}$$

by Assumption 1 if the periods s and t are close to each other. To shorten the notation, let

$$g_s^a(y, v, p) := \int c(y, \tilde{v} + v, p) \text{distr}(\tilde{v} | H_s(y, a)).$$

Then one obtains

$$\bar{c}_t(y, a, p_t) = g_s^a(y, v_t(y, a), p_t). \quad (4.2)$$

Note that $g_s^a(y, v_s(y, a), p_s) = \bar{c}_s(y, a, p_s)$.

Now Assumption 2 comes into play, i.e., $v_t(y, a) = \varphi(y, a) + \psi(t, a)$.

Let

$$\begin{aligned} v_t(a) &:= \text{mean}_{H_t(a)} v_t^h = \int v_t(y, a) \text{distr}(y | H_t(a)) \quad \text{and} \\ v_{t,s}(a) &:= \int v_t(y, a) \text{distr}(y | H_s(a)). \end{aligned}$$

With this definition one obtains

$$\psi(t, a) - \psi(s, a) = v_{t,s}(a) - v_s(a),$$

and hence,

$$v_t(y, a) = v_{t,s}(a) + \varphi(y, a) - v_s(a) + \psi(s, a).$$

Then (4.2) leads to

$$\bar{c}_t(y, a, p_t) = f_s^a(y, v_{t,s}(a), p_t) \quad (4.3)$$

where $f_s^a(y, v, p) := g_s^a(y, v + \varphi(y, a) - v_s(a) + \psi(s, a), p)$.

Note that $f_s^a(y, v_s(a), p_s) = \bar{c}_s(y, a, p_s)$. Substituting (4.3) in (4.1) one obtains

$$\begin{aligned} C_t &= \int f_s^a(y, v_{t,s}(a), p_t) \text{distr}(y, a | H_t) \\ &= \int f_s^a(\Sigma_t^{\frac{1}{2}} \Sigma_s^{-\frac{1}{2}}(y - m_s) + m_t, v_{t,s}(a), p_t) \text{distr}(y, a | H_s) \end{aligned}$$

by Hypotheses 1 and 2.

Consequently, $C_t - C_s =$

$$\int \left[f_s^a(\Sigma_t^{\frac{1}{2}} \Sigma_s^{-\frac{1}{2}}(y - m_s) + m_t, v_{t,s}(a), p_t) - f_s^a(y, v_s(a), p_s) \right] \text{distr}(y, a | H_s).$$

A first order Taylor expansion of $f_s^a(y, v, p)$ at $(y, v_s(a), p_s)$ then yields

$$\begin{aligned} C_t - C_s &= \int [\partial_y f_s^a(y, v_s(a), p_s)]^T ((\Sigma_t^{\frac{1}{2}} \Sigma_s^{-\frac{1}{2}} - \mathbb{I})(y - m_s) + (m_t - m_s)) \text{distr}(y, a | H_s) \\ &+ \int [\partial_v f_s^a(y, v_s(a), p_s)]^T (v_{t,s}(a) - v_s(a)) \text{distr}(y, a | H_s) \\ &+ \int [\partial_p f_s^a(y, v_s(a), p_s)]^T (p_t - p_s) \text{distr}(y, a | H_s) \\ &+ \text{terms of second order in } \|m_t - m_s\|^2, \\ &\quad \|\Sigma_t^{\frac{1}{2}} \Sigma_s^{-\frac{1}{2}} - \mathbb{I}\|^2, \|v_{t,s}(a) - v_s(a)\|^2, \text{ and } \|p_t - p_s\|^2. \end{aligned}$$

Since $f_s^a(y, v_s(a), p_s) = \bar{c}_s(y, a, p_s)$ the vector β_s of coefficients in the Proposition is defined by

$$\beta_s := \frac{1}{C_s} \int \partial_y \bar{c}_s(y, a, p_s) \text{distr}(y, a | H_s).$$

Since

$$\begin{aligned} &\int [\partial_y \bar{c}_s(y, a, p_s)]^T ((\Sigma_t^{\frac{1}{2}} \Sigma_s^{-\frac{1}{2}} - \mathbb{I})(y - m_s)) \text{distr}(y, a | H_s) \\ &= \text{trace} \left[\int (y - m_s) [\partial_y \bar{c}_s(y, a, p_s)]^T \text{distr}(y, a | H_t) (\Sigma_t^{\frac{1}{2}} \Sigma_s^{-\frac{1}{2}} - \mathbb{I}) \right] \end{aligned}$$

the matrix Γ_s of coefficients in the Proposition is defined by

$$\Gamma_s = \frac{1}{C_s} \int (y - m_s) [\partial_y \bar{c}_s(y, a, p_s)]^T \text{distr}(y, a | H_s).$$

5 Empirical Results

For analyzing our aggregate model of Section 4 we use data from the U.K. Family Expenditure Survey (FES). Each year a total of approximately 7000 households record their

expenditures on a large variety of consumption items. Also included in the survey are different forms of income and household attributes. For a precise definition of the variables, sampling units, sampling designs, interviewing and field work, confidentiality, reliability etc. we refer to the respective yearly FES manuals as well as the Family Survey Handbook of Kemsley et al (1980). We include into the analysis data made available to us for all years between 1968 and 1993 except for the year 1978, where our income variable could not be constructed due to problems in the databasis. Households from Northern Ireland were eliminated for all years.

In the present study we use information on household income and consumption as well as on demographic and socioeconomic variables such as age and occupational status of the household head, household size, etc., included in the yearly surveys. In economic literature most studies focus on consumption of nondurable goods. Following this tradition we will consider nondurable consumption which is defined as total consumption expenditure on all goods and services minus housing costs and durable goods. Following HBAI standards, household incomes are obtained by extracting relevant items from the elementary database¹. We distinguish between current disposable non-property as well as asset income of each household. Our definition of asset income corresponds to the aggregate "investment income" used in the FES. It includes all sources of income which are due to private investments or property. An approximation of household assets is obtained by using the quotient of property income and the corresponding average yearly interest rate. It must be emphasized that, for example, the value of an owner occupied house is not included in this definition of assets. Consumption, assets and income in real prices are determined by dividing by the price index of the respective month in which the household was included in the survey.²

We will concentrate on the effect of changes in the joint distribution of current income and wealth on aggregate consumption. A major complication is the fact that there is a considerable percentage of households in the sample with property income equal to zero. In average over all years this "null group" consists of approximately 40 percent of all households². Our analysis is performed separately for this group and the remaining "non-null" group of household with positive wealth.³

²The task of elaborating the database and specifying consistent variables has mainly been accomplished by Jürgen Arns. His careful work is gratefully acknowledged.

³Also included in the "null group" are households with an extremely small property income of less than 0.02 pounds per week in prices of 1968. There also exists a very small number of households with negative values of either property income or disposable non-property income. These households have been eliminated from the samples. In order to diminish the potential influence of outliers, all households with consumption larger than eight times median consumption were also excluded. In total this procedure leads to an elimination of between 0.1% and 0.3% of all households in the different samples.

For the null group the standardized log income distribution turns out to be very stable over time. This corresponds to previous results for the entire population as shown in Hildenbrand, Kneip and Utikal (1998). For the non-null group one has to study the joint distribution of income and assets. Interestingly the correlation between these variables is extremely small for all years. The average value of the coefficient of correlation is -0.09 . Let $y_{t,1}$ and $y_{t,2}$ denote log-income and log-assets in year t , and let $m_{t,1}^{(1)}, \sigma_{t,1}^{(1)}$ and $m_{t,2}^{(1)}, \sigma_{t,2}^{(1)}$ denote the corresponding means and standard deviations for the non-null group. The joint distribution of $(\frac{y_{t,1}-m_{t,1}^{(1)}}{\sigma_{t,1}^{(1)}}, \frac{y_{t,2}-m_{t,2}^{(1)}}{\sigma_{t,2}^{(1)}})$ changes very slowly over time. This is illustrated in Figures 5.1 and 5.2 which show contour plots of the resulting bivariate densities for the years 1987 and 1989. The structure of the density lines also indicate that the two variables are "almost" independent.

We apply our theoretical approach separately for the null and the non-null group to analyze the dependence of the respective consumptions $C_t^{(0)}$ and $C_t^{(1)}$ on income and assets. We only consider one year predictions with $s = t - 1$. Since yearly changes in the data are of a order of magnitude of less than seven percent, the differences between $\frac{C_t^{(j)} - C_{t-1}^{(j)}}{C_{t-1}^{(j)}}$ and $\Delta \log C_t^{(j)} = \log C_t^{(j)} - \log C_{t-1}^{(j)}$ are negligible for $j = 0, 1$

For the null group there are no assets and our Proposition thus simplifies to

$$\Delta \log C_t^{(0)} = \beta_{t-1,1}^{(0)}(m_{t,1}^{(0)} - m_{t-1,1}^{(0)}) + \gamma_{t-1,1}^{(0)} \frac{\sigma_{t,1}^{(0)} - \sigma_{t-1,1}^{(0)}}{\sigma_{t-1}^{(0)}} + \text{remainder term} \quad (5.1)$$

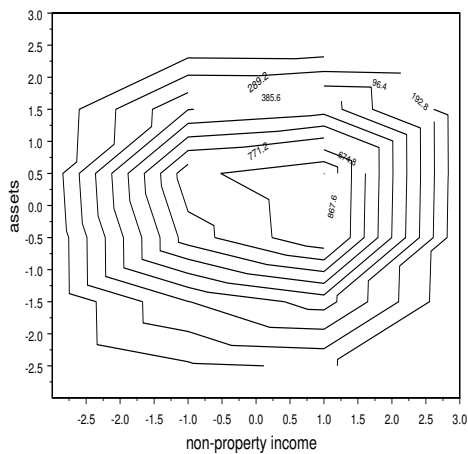


Fig. 5.1: density contours for 1987

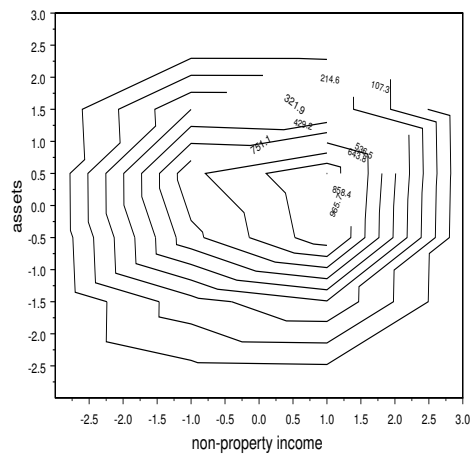


Fig. 5.2: density contours for 1989

Here $m_{t,1}^{(0)}, \sigma_{t,1}^{(0)}$ are mean and standard deviation of log income for the null group. As mentioned above, there is only a very small correlation between income and assets for the

non-null group. When assuming that already the joint distribution of $(\frac{y_{t,1}-m_{t,1}^{(1)}}{\sigma_{t,1}^{(1)}}, \frac{y_{t,2}-m_{t,2}^{(1)}}{\sigma_{t,2}^{(1)}})$ is approximately time-invariant the terms in our expansion depending on differences of the covariance matrices simplify in the sense that only changes of the respective variances have to be taken into account. We then obtain

$$\begin{aligned} \Delta \log C_t^{(1)} = & \beta_{t-1,1}^{(1)}(m_{t,1}^{(1)} - m_{t-1,1}^{(1)}) + \beta_{t-1,2}^{(1)}(m_{t,2}^{(1)} - m_{t-1,2}^{(2)}) \\ & + \gamma_{t-1,1}^{(1)} \frac{\sigma_{t,1}^{(1)} - \sigma_{t-1,1}^{(1)}}{\sigma_{t-1,1}^{(1)}} + \gamma_{t-1,2}^{(1)} \frac{\sigma_{t,2}^{(1)} - \sigma_{t-1,2}^{(1)}}{\sigma_{t-1,2}^{(1)}} + \text{remainder term} \end{aligned} \quad (5.2)$$

In (5.1) and (5.2) the influence of the additional explanatory variables v and p is summarized by writing "remainder term". Note that our general theory does not provide any information on the stochastic structure of this term. In particular, it is **not** reasonable to assume that these remainder terms can be treated as i.i.d. error terms.

As already mentioned above our aim in this section is a partial analysis. We want to capture the effect of changes in the joint distribution of current income and wealth on aggregate consumption. However, this goal requires to specify a valid way to determine the parameters β_t and γ_t .

Following usual macroeconomic analysis parameter estimation has to be based on time series models for $\{\Delta \log C_t^{(j)}, m_t^{(j)}, \dots\}$. However, from this point of view "models" (5.1) and (5.2) are incomplete and do not allow any consistent parameter estimation. In order to establish a valid time series model it will be necessary to specify the additional variables hidden in the "remainder term" and to study their stochastic behavior. Further assumptions will have to be made concerning the possible variation of the parameters β_t and γ_t , which in our general approach are allowed to change from period to period. Of course, at any stage of such a process of model building one encounters the inherent danger of misspecifications. Incorrect models may lead to false conclusions.

Our approach offers a way to estimate the parameters without a time series modelling of $\beta_t^{(1)}$. As has been explained in the theoretical part the values of β_t , and γ_t are to be obtained from suitable derivatives of regression functions. Separately for each year t they can be **estimated from the cross-section data** on individual income and assets provided by the FES. Figures 5.3 and 5.4 show the resulting estimates $\hat{\beta}_{t,1}^{(0)}$, $\hat{\beta}_{t,1}^{(1)}$, $\hat{\beta}_{t,2}^{(1)}$, $\hat{\gamma}_{t,1}^{(0)}$, $\hat{\gamma}_{t,1}^{(1)}$ and $\hat{\gamma}_{t,2}^{(1)}$ for nondurable consumption of the null as well as the non-null group. Details of the estimation procedure are described in Subsection 5.1.

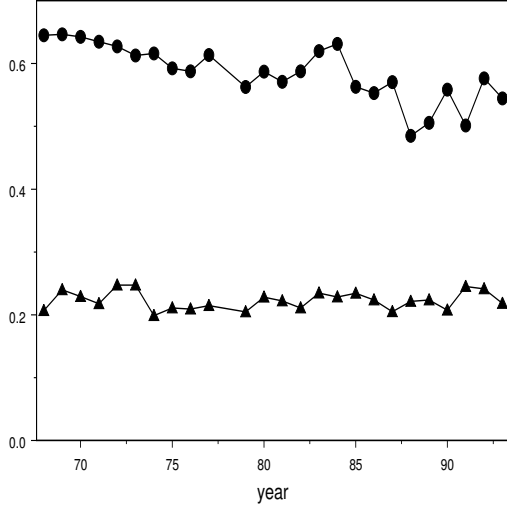


Fig. 5.3: Estimated values of $\beta_{t,1}^{(0)}$ ("•"), and $\gamma_{t,1}^{(0)}$ ("▲")

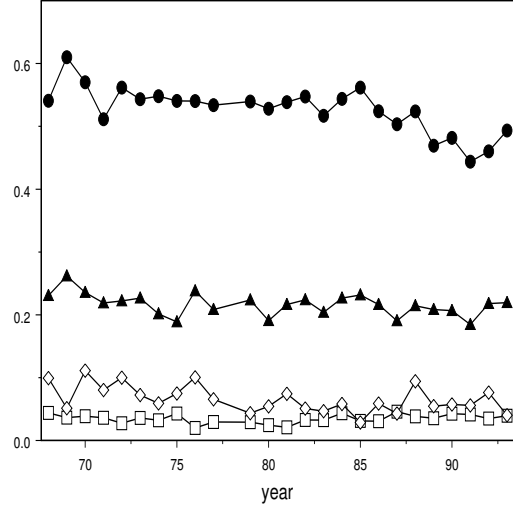


Fig. 5.4: Estimated values of $\beta_{t,1}^{(1)}$ ("•"), $\gamma_{t,1}^{(1)}$ ("▲"), $\beta_{t,2}^{(1)}$ ("□") and $\gamma_{t,2}^{(1)}$ ("◇")

When considering disposable income the average values of $\hat{\beta}_{t,1}^{(0)}$, $\hat{\beta}_{t,1}^{(1)}$, $\hat{\gamma}_{t,1}^{(0)}$, and $\hat{\gamma}_{t,1}^{(1)}$ are 0.59, 0.53, 0.22, and 0.22. Changes of assets seem to possess a comparably much smaller influence on consumption. The average values of $\hat{\beta}_{t,2}^{(1)}$ and $\hat{\gamma}_{t,2}^{(1)}$ are 0.035 and 0.068. Since, however, our data only allows a rather rough approximation of household assets some care is necessary when interpreting these results.

One recognizes that the estimates $\hat{\beta}_{t,1}^{(0)}$ as well as $\hat{\beta}_{t,1}^{(1)}$ possess a slightly falling trend. In view of our theory this is quite easily interpretable. First we note that the time series $\{m_{t,1}^{(j)}\}$ determined from our data have a pronounced *increasing* trend. At the same time it is easily seen from its definition in Proposition 4 that $\hat{\beta}_{t,1}^{(j)}$ can be interpreted as a mean income elasticity of consumption across the respective population in period t . A *falling* trend of $\hat{\beta}_{t,1}^{(j)}$ therefore seems to indicate that the mean income elasticity becomes smaller when the general level of income, as quantified by the mean $m_{t,1}^{(j)}$ of log income, increases. This is certainly not implausible.

We now consider the question which proportion of consumption is explained by changes in the current income distribution. We will consider the approximations $\widehat{\Delta \log C_t^{(0)}}$ and $\widehat{\Delta \log C_t^{(1)}}$ of $\Delta \log C_t^{(0)}$ and $\Delta \log C_t^{(1)}$ obtained by the different components of models (5.1) and (5.2). For $j = 0, 1$ we use two measures to quantify the remaining differences, the average absolute error (*AE*) and the relative residual sum of squares (*RRSS*):

$$AE = 100 \cdot \frac{1}{T} \sum_t |\Delta \log C_t^{(j)} - \widehat{\Delta \log C_t^{(j)}}|, \quad RRSS = \frac{\sum_t |\Delta \log C_t^{(j)} - \widehat{\Delta \log C_t^{(j)}}|^2}{\sum_t |\Delta \log C_t^{(j)}|^2}$$

$RRSS$ measures the sum of squared residuals relative to the original squared differences $|\Delta \log C_t^{(j)}|^2$. In a standard parametric regression model we have $RRSS = 1 - R^2$. Obviously, the better a model the smaller the values of AE and $RRSS$. In order to obtain a detailed picture the two and the four terms on the right hand side of (5.1) and (5.2), respectively are added one by one. Hence, the second row of Table 5.1 refers to the approximation $\widehat{\Delta \log C_t}^{(j)}$ obtained when using only the first term $\hat{\beta}_{t-1,1}^{(j)}(m_{t,1}^{(j)} - m_{t-1,1}^{(j)})$ only. The last row corresponds to the complete model.

In addition, the final predictions $\widehat{\Delta \log C_t}^{(j)}$ from the complete models (5.1) and (5.2) were used to approximate changes $\Delta \log C_t$ of aggregate consumption for the **total** population of all households. We therefore use the approximation

$$\Delta \log C_t \approx \pi_{t-1}^{(0)} \frac{C_{t-1}^{(0)}}{C_{t-1}} \widehat{\Delta \log C_t}^{(0)} + \pi_{t-1}^{(1)} \frac{C_{t-1}^{(1)}}{C_{t-1}} \widehat{\Delta \log C_t}^{(1)}, \quad (5.3)$$

where $\pi_{t-1}^{(0)}$ and $\pi_{t-1}^{(1)}$ are the respective proportions of households in the null and non-null group in period $t - 1$.

	null group		non-null		total	
	AE	$RRSS$	AE	$RRSS$	AE	$RRSS$
$\Delta \log C_t$	2.24		2.52		2.22	
$\hat{\beta}_{t-1,1}(m_{t,1} - m_{t-1,1})$	1.90	0.711	1.34	0.325		
$\hat{\beta}_{t-1,1}(m_{t,1} - m_{t-1,1}) + \hat{\beta}_{t-1,2}(m_{t,2} - m_{t-1,2})$			1.34	0.300		
$\hat{\beta}_{t-1,1}(m_{t,1} - m_{t-1,1}) + \hat{\beta}_{t-1,2}(m_{t,2} - m_{t-1,2})$ $+ \hat{\gamma}_{t-1,1} \frac{\sigma_{t,1}^{(1)} - \sigma_{t-1,1}^{(1)}}{\sigma_{t-1,1}^{(1)}}$	1.58	0.584	1.52	0.427		
$\hat{\beta}_{t-1,1}(m_{t,1} - m_{t-1,1}) + \hat{\beta}_{t-1,2}(m_{t,2} - m_{t-1,2})$ $+ \hat{\gamma}_{t-1,1} \frac{\sigma_{t,1} - \sigma_{t-1,1}}{\sigma_{t-1,1}} + \hat{\gamma}_{t-1,2} \frac{\sigma_{t,2} - \sigma_{t-1,2}}{\sigma_{t-1,2}}$			1.50	0.378	1.38	0.392

Table 5.1

The following figure shows the yearly errors obtained when predicting nondurable consumption for the whole population by (5.3).

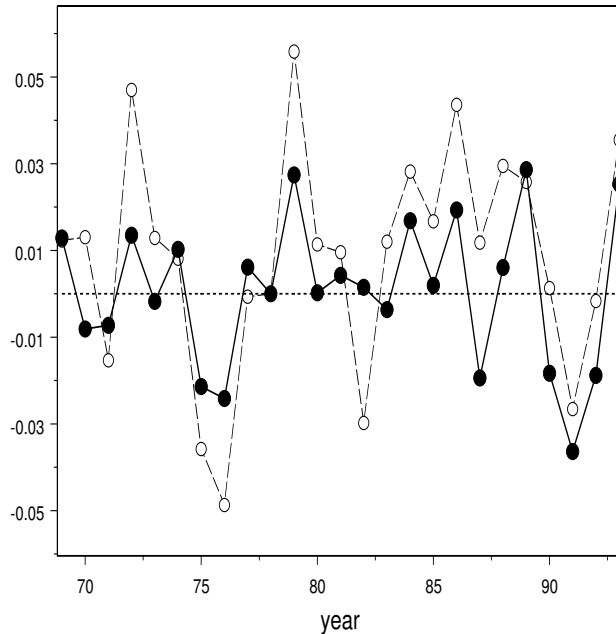


Fig. 5.4: $\Delta \log C_t$ ("o") and final approximation error ("•")

We want to note that similar results are obtained when considering total consumption of all goods and services instead of only nondurable consumption. In this case AE and *RRSS* for the final model are 1.60 and 0.333.

Table 5.1 shows that based on (5.1) and (5.2) changes in income and assets explain a considerable part of the variation of $\Delta \log C_t^{(j)}$, $j = 0, 1$. This is a remarkable result which may help to settle the long-lasting discussion in consumption theory whether or not aggregate income possess an influence on aggregate consumption. The point is that when interpreting the table it must be taken into account that there exists a crucial difference to usual model fits obtained from standard time series methods. Recall that our approach does not rely on fitting (5.1) or (5.2) to the observed time series $\{\Delta \log C_t^{(j)}\}$. Indeed, parameter estimates $\hat{\beta}_t$ and $\hat{\gamma}_t$ are computed from cross-section data and calculation of $\widehat{\Delta \log C_t^{(j)}}$ does not incorporate any information about the structure of $\{\Delta \log C_t^{(j)}\}$. From a statistical, data-analytic point of view there is thus no mechanism which enforces a small approximation error. The possible values of *RRSS* are theoretically unbounded.

When considering the table in more detail one recognizes an ambivalent role of the term $\hat{\gamma}_{t-1,1} \frac{\sigma_{t,1}^{(1)} - \sigma_{t-1,1}^{(1)}}{\sigma_{t-1,1}^{(1)}}$, which quantifies the influence of the changing income dispersion on consumption expenditure. When adding this variance term, the approximation error decreases for the null group and increases for the non-null group. There is a possible economic in-

terpretation for this effect. A potentially important explanatory variable not considered in (5.1) and (5.2) is the uncertainty of anticipated future income. In the Appendix a theoretical argument is given which relates the general level of uncertainty in period t to

$$V_t = \text{var}(\log y_t^h - \log y_{t-1}^h)$$

which is the variance of the differences $\log y_t^h - \log y_{t-1}^h$ across the population. An increasing/decreasing value of V_t then indicates an increasing/decreasing general level of uncertainty. One may note that commonly used measures of uncertainty, as for example the unemployment rate, are related to this variable. More unemployment will usually result in higher values of V_t .

Let us now consider the role of $\sigma_t^2 - \sigma_{t-1}^2$. It is immediately seen from Figure 5.1 that all estimated γ_t are positive and therefore the *direct* effect of an increasing variance of the income distribution stimulates consumption. However, one easily verifies that

$$\sigma_t^2 - \sigma_{t-1}^2 = V_t + 2 \text{cov}_H(\log y_t^h - \log y_{t-1}^h, \log y_{t-1}^h)$$

Although the covariance term may act as a nuisance, large values of $\sigma_t^2 - \sigma_{t-1}^2$ may thus tend to go along with large values of V_t and high income uncertainty. The negative effect of higher uncertainty on consumption may well explain the empirical results of Table 5.1. Due to the existence of savings it seems to be reasonable to assume that households in the non-null group generally have a more "forward-looking" behavior and thus are more sensitive to uncertainty than households in the null group.

Of course, these arguments also give additional weight to the fact that (5.1) and (5.2) are incomplete. A theoretically sound consumption model will have to include the effects of changing interest and inflation rates as well as aggregate proxies for expectation and uncertainty of future income. Such proxies for expectations of future income may, for example, include lagged values of mean log income. All parameters quantifying the influence of such additional variables have to be estimated from the *residual time series* $\{\Delta \log C_t^{(j)} - \widehat{\Delta \log C_t^{(j)}}\}$. Establishing a valid time series model incorporating all relevant variables obviously requires a considerable amount of additional work which is not in the scope of the present paper.

5.1 Cross-section estimation of coefficients

Assume that for each period s there are data $(c_s^h, y_{s1}^h, y_{s2}^h, a_s^h)$, $h = 1, \dots, n_s$ about current consumption, log income, log assets, and household attributes from an independent sample of n_s households. Since the value of p_s in period s does not depend on h , definition of

$\beta_s = (\beta_{s1}, \beta_{s2})$ implies that with $y = (y_1, y_2)$

$$\beta_s = \frac{1}{C_s} \int \partial_y c_s(y_1, y_2, a, p_s) \text{distr}(y_1, y_2, a | H_s) = \frac{1}{C_s} \int \partial_y \bar{c}_s(y_1, y_2, a) \text{distr}(y_1, y_2, a | H_s)$$

where $\bar{c}_s(\cdot) \equiv c_s(\cdot, p_s)$ is the regression function of c_s^h on (y_s^h, a_s^h) . Estimates \hat{c}_s and $\partial_y \hat{c}_s$ of \bar{c}_s and its derivative with respect to y can thus be obtained by suitable parametric or nonparametric regression methods. Indeed, from a statistical point of view the problem of estimating β_s falls into the domain of average derivative estimation (see, for example, considered Härdle and Stoker (1989) or Stoker (1991)).

We use a generalized version of a "direct" average derivative estimator. In order to guard against misspecifications in the relation between c and y estimation relies on a semi-parametric model of the form

$$c_s^h = \bar{c}_s(y_{s1}^h, y_{s2}^h, a_{sj}^h) + \epsilon_s^h = f_1(y_{s1}^h) + f_2(y_{s2}^h) + \sum_j \vartheta_j a_{sj}^h + \epsilon_s^h$$

The household attributes a_{sj}^h used are age, age² and indicator variables referring to household size, employment status, occupation, month in which the household was recorded, and region. For approximating the unknown functions f_j , $j = 1, 2$, we rely on a quadratic spline function with a prespecified number k of knots $i_{j0}, i_{j1}, \dots, i_{jk}$. The knot locations are chosen in such a way that in each interval $[i_{j,l-1}, i_{jl}]$ there are approximately the same number of observations y_{s1}^h or y_{s2}^h , respectively. The spline parameters as well as the ϑ_j are then estimated by least squares, and with $\partial_{y_1} \hat{c}_s(y_{s1}^h, y_{s2}^h, a_{sj}^h) = \hat{f}'_1(y_{s1}^h)$, $\partial_{y_2} \hat{c}_s(y_{s1}^h, y_{s2}^h, a_{sj}^h) = \hat{f}'_2(y_{s2}^h)$ an estimate of β_s is then determined by

$$\hat{\beta}_{s1} = \frac{\sum_{h=1}^{n_s} \hat{f}'_1(y_{s1}^h)}{\sum_{h=1}^{n_s} c_s^h}, \quad \hat{\beta}_{s2} = \frac{\sum_{h=1}^{n_s} \hat{f}'_2(y_{s2}^h)}{\sum_{h=1}^{n_s} c_s^h}$$

By similar arguments reasonable estimates of the coefficient matrix Γ_s are obtained by

$$\hat{\Gamma}_{s,ij} = \frac{\sum_{h=1}^{n_s} (y_{si}^h - \hat{m}_{si}) \hat{f}'_j(y_{sj}^h)}{\sum_{h=1}^{n_s} c_s^h}$$

where \hat{m}_s denotes the sample average of y_s^h . The results presented in Section 5.1 turn out to be stable when choosing a number of knots between $k = 6$ and $k = 25$.

Appendix: Expected intertemporal utility maximization: The consumption function of a forward looking household

Given two non-negative stochastic processes $(\eta_\tau)_{\tau=0,\dots,T}$ and $(\rho_\tau)_{\tau=0,\dots,T}$ on a probability space (Ω, \mathcal{F}, P) with (non-random) starting points $\eta_0(\omega) = y_0$ and $\rho_0(\omega) = r_0$ and a (utility) function in $T + 2$ real variables.

Let $(c_\tau)_{\tau=0,\dots,T}$ denote a non-negative stochastic process such that c_τ is \mathcal{F}_τ -measurable, where $\mathcal{F}_\tau = \sigma(y_\tau, y_{\tau-1}, \dots, \rho_\tau, \rho_{\tau-1}, \dots) \subset \mathcal{F}$.

For given starting points y_0, r_0 and $W_0 \geq -L$ consider the following maximization problem:

$$\max \int u(c_0, c_1(\omega), \dots, c_T(\omega), W_{T+1}(\omega)) P(d\omega)$$

subject to the sequence $(\tau = 0, \dots, T)$ of budget constraints

$$W_{\tau+1} = (1 + \rho_\tau)(W_\tau + \eta_\tau - c_\tau) \geq -L, \quad P - \text{a.e.}$$

A solution (c_τ^*) and hence, in particular, its first component c_0^* is determined by $y_0, W_0, r_0, (\rho_\tau, \eta_\tau)_{\tau=1,\dots,T}, u$, and L , i.e.,

$$c_0^* \equiv c_0^*[y_0, W_0, r_0, (\rho_\tau, \eta_\tau)_{\tau=1,\dots,T}, u, L]. \quad (\text{A.1})$$

If one is interested in an explicit solution, then, of course, one has to make specific assumptions on the stochastic process (η_τ, ρ_τ) and, in particular, on the utility function.

Now, we consider a household h in period t with current real income y_t^h , real financial wealth $W_t^h \geq -L_t^h$ and an intertemporal utility function u_t^h in $T_t^h + 2$ variables.

In making the decision c_t^h on current real consumption expenditure the household looks into the future $\tau = t + 1, \dots, t + T_t^h$. Let $\eta_\tau^h(t)$ and $\rho_\tau^h(t)$ denote the uncertain real income and uncertain real interest rate, respectively, in the future period τ *as anticipated in period t*. The stochastic process $(\eta_\tau^h(t))_\tau$ describes the subjective beliefs about future income of household h . If consumption behavior on the household level is modelled by the above maximization problem then real current consumption expenditure c_t^h is determined by $y_t^h, W_t^h, r_t, (\eta_\tau^h(t), \rho_\tau^h(t))_{\tau=t+1,\dots,t+T_t^h}, u_t^h$, and L_t^h :

$$c_t^h \equiv c_t^h[y_t^h, W_t^h, r_t, (\eta_\tau^h(t), \rho_\tau^h(t))_{\tau=t+1,\dots,t+T_t^h}, u_t^h, L_t^h]. \quad (\text{A.2})$$

It will be appropriate to make a change of variable. Consider the stochastic growth rate of anticipated income which is defined by

$$z_\tau^h(t) := \log \eta_\tau^h(t) - \log \eta_{\tau-1}^h, \quad \tau > t$$

where $\eta_\tau^h(t) := y_t^h$. Since

$$z_\tau^h(t) = \log \left(1 + \frac{\Delta \eta_\tau^h(t)}{\eta_{\tau-1}^h(t)} \right) \approx \frac{\Delta \eta_\tau^h(t)}{\eta_{\tau-1}^h(t)},$$

$100 \cdot z_\tau^h(t)$ can be interpreted as the uncertain percentage change in anticipated income. The stochastic process $(\eta_\tau^h(t))_\tau$ is determined by y_t^h and the stochastic process $(z_\tau^h(t))_\tau$. Hence current consumption expenditure c_t^h is determined by $y_t^h, W_t^h, r_t, (z_\tau^h(t), \rho_\tau^h(t))_{\tau=t+1, \dots, t+T_t^h}, u_t^h$, and L_t^h :

$$c_t^h[y_t^h, W_t^h, r_t, (z_\tau^h(t), \rho_\tau^h(t))_{\tau=t+1, \dots, t+T_t^h}, u_t^h, L_t^h]. \quad (\text{A.3})$$

Classification of the explanatory variables: the explanatory variables y_t^h (income), W_t^h (wealth) and r_t (interest rate) are observable. All other explanatory variables in (A.3) are viewed as unobservable.

As we showed in the Remark to the Proposition in Section 4, a particular favorable case for our decomposition of the change in mean consumption expenditure C_t prevails if the mean of an unobservable explanatory variable v_t^h across the subpopulation $H_t(y, W, a)$ is time-invariant. As we showed, such an explanatory variable has no effect on the *change* in mean consumption expenditure C_t . Note that time-invariance of $\text{mean}_{H_t(y, W, a)} v_t^h$ does not imply time-invariance of $\text{mean}_{H_t} v_t^h$, since $\text{distr}(y, W, a | H_t)$ is changing over time. If one can expect that an explanatory variable v_t^h is determined by income y , wealth W and the attribute profile a , i.e., $v_t^h = v(y, W, a)$ for all households in $H_t(y, W, a)$ in every period t , then, trivially, one obtains the desired time-invariance. One might view the unobservable variables L_t^h, T_t^h and u_t^h in (A.3) to be of this type.

It is standard practice in aggregate consumption analysis to model the expectations about future real interest rates by $\rho_\tau^h(t) \equiv r_t$, that is to say, one postulates that all households make their decisions under the assumption that future real interest rates are equal to the current real interest rate. Thus, by assumption the unobservable expectational variable $\rho_\tau^h(t)$ becomes observable.

It remains to discuss the modelling of the stochastic future growth rates $z_\tau^h(t)$ as anticipated in period t .

In the "Rational Expectations" literature one starts from the assumption that *past and future* income of a household is a realization of an autonomous stochastic process (\tilde{y}_s^h) . The probability distribution of $(\eta_\tau^h(t))_{\tau=t+1, \dots}$ (future real income as anticipated in period t) is then defined as the conditional probability distribution of $(\tilde{y}_{t+1}^h, \tilde{y}_{t+2}^h, \dots)$ given \mathcal{F}_t^h , i.e., given observations of $\tilde{y}_t^h, \tilde{y}_{t-1}^h, \dots$. Thus, in particular $\mathbb{E}\eta_\tau^h(t) = \mathbb{E}(\tilde{y}_\tau^h | \tilde{y}_t^h, \tilde{y}_{t-1}^h, \dots)$. This view might be appropriate for the fiction of a "representative" household whose income in period t is the mean income across the population H_t .

Alternatively, one might consider directly the future stochastic growth rate $z_\tau^h(t)$ as anticipated in period t .

Assume, for example, that every household in H_t believes, that is to say, makes its decision in period t under the assumption, that its future log-income $\log \eta_\tau^h(t)$ is determined by a random walk with drift μ_t^h . Then, $z_\tau^h(t) = \epsilon_\tau^h + \mu_t^h$ where $\epsilon_\tau^h \sim IIN(0, \nu_{t,h}^2)$. The stochastic process $(z_\tau^h(t))$ is fully determined by the two parameters μ_t^h and $\nu_{t,h}^2$; $\mathbb{E}z_\tau^h(t) = \mu_t^h$, $V(z_\tau^h(t)) = \nu_{t,h}^2$ and $\text{cov}(z_\tau^h(t), z_\tau^{h'}(t)) \equiv 0$. Consequently, in (A.3) the unobservable parameters $(\mu_t^h, \nu_{t,h}^2)$ take the role of $(z_\tau^h(t))_\tau$. Obviously, in this case - in contrast to the case of L_t^h , T_t^h or u_t^h - it is hard to justify why $\text{mean}_{H_t(y,W,a)} \mu_t^h$ and $\text{mean}_{H_t(y,W,a)} \nu_{t,h}^2$ should be time-invariant. Rather one might assume now that these means are independent of y and W , yet changing over time. What then might cause the change in the mean of μ_t^h and $\nu_{t,h}^2$ across the subpopulation $H_t(a)$? Naturally one can never exclude a general change in opinion about the future which can not be related to any of the explanatory variables considered up to now. One can, however, argue that at least some part of a change in $\text{mean}_{H_t(a)} \mu_t^h$ and $\text{mean}_{H_t(a)} \nu_{t,h}^2$ might be attributed to a change in $\text{mean}_{H_t(a)} \Delta \log y_t^h$ and $\text{variance}_{H_t(a)} \Delta \log y_t^h$, respectively. For example, if $\mu_t^h = \mu_t^a$ and $\nu_{t,h}^2 = \nu_{t,a}^2$ for all $h \in H_t(a)$ and if all households believe that their experienced growth rate $\Delta \log y_t^h$ in period t is an independently drawn sample from $N(\mu_t^a, \nu_{t,a}^2)$, then the common expectational variables μ_t^a and $\nu_{t,a}^2$ must satisfy $\text{distr}(\Delta \log y_t^h | H_t) \sim N(\mu_t^a, \nu_{t,a}^2)$.

References

- Atkinson, A. B.; Bourguignon, F. (eds).** *Handbook of Income Distribution*. Amsterdam, Elsevier, 2000.
- Blundell, R. W; Stoker, T. M.**, 2002. "Models of Aggregate Economic Relationships That Account For Heterogeneity", mimeo, Institute for Fiscal Studies, London, to appear in *Handbook of Econometrics*, Vol. 6.
- Davidson, J. E. H.; Hendry, David F.; Srba, Frank and Yeo, S.** "Econometric Modelling of the Aggregate Time-Series Relationship between Consumers' Expenditure and Income in the United Kingdom", *Economic Journal*, December 1978, (88), pp. 661-692.
- Deaton, A. S.** *Understanding Consumption*. Oxford, Clarendon Press, 1992.
- DSS.** 1995. Household Below Average Income 1979-1992/93: A Statistical Analysis, London, Her Majesty's Stationery Office.
- Gorman, W. M.** "Community Preference Fields", *Econometrica*, January 1953, (21), pp. 63-80.
- Greene, W. H.** *Econometric Analysis (Fifth Edition)*. New Jersey, Prentice Hall, 2003.
- Härdle, W. and Stoker, T. M.** "Investigating Smooth Multiple Regression by the Method of Average Derivatives", *Journal of the American Statistical Association*, 1989, (84), pp. 986-995.
- Her Majesty's Stationery Office.** *Family Expenditure Survey*. Her Majesty's Stationery Office. Annual Report. Material from the Family Expenditure Survey is Crown Copyright; has been made available by the Office for National Statistics through the Data Archive; and has been used by permission. Neither the ONS nor the Data Archive bear any responsibility for the analysis or interpretation of the data reported here.
- Hildenbrand, W. and Kneip, A.** "Demand Aggregation under Structural Stability", *Journal of Mathematical Economics*, February 1999, (31, 1), pp. 81-109.
- Hildenbrand, W.; Kneip, A. and Utikal, K.** "Une Analyse non Parametrique des Distributions du Revenu et des Caracteristiques des Menages", *Revue de Statistique Applique*, 1998, (47), pp. 39-56.
- Kemsley, W.F.F.; Redpath, R.U. and Holmes, M.** *Family Expenditure Survey Handbook*. London, Her Majesty's Stationery Office, 1980.
- Lewbel, A.** "Income Distribution Movements and Aggregate Money Illusions", *Journal of Econometrics*, 1990, (43), pp. 35-42.

- Lewbel, A.** "Aggregation with Log-linear Models", *Review of Economic Studies*, 1992, (59), pp. 635-642.
- Malinvaud, E.** "A Framework for Aggregation Theories", *Ricerche Economiche*, Vol. 47, (1993), pp. 107-135.
- Nataf, A.** "Sur la possibilité de construction de certains macromodèles", *Econometrica*, 1948, (16), pp. 232-244.
- Pareto, V.** *Cours d'économie politique*. Rouge, Lausanne, 1897.
- Stoker, T. M.** "Empirical Approaches to the Problem of Aggregation over Individuals", *Journal of Economic Literature*, 1993, (Vol. XXXI (December)), pp. 1827-1874.
- Stoker, T. M.** "Equivalence of Direct, Indirect and Slope Estimators of Average Derivatives," in *Nonparametric and semiparametric methods in econometrics and statistics*. Barnett, W.A. and Powel, J.L. and Tauchen, G. Cambridge, Cambridge University Press, 1991.
- Stone, R.** "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand", *Economic Journal*, 1954, (Vol. 64), pp. 511-27.