Aggregate Consumers' Expenditure and Income

Werner Hildenbrand

and

Alois Kneip

January 19, 1995

1 Introduction

Attempts by economists and econometricians to understand the saving and consumption patterns, either at the level of the individual household or at the level of the economy as a whole, have lead to an extensive literature, theoretical as well as applied. An excellent exposition of the state of the art is presented in DEATON'S recent book *Understanding Consumption* (1993). On the last page of this book, p.221, DEATON writes "I should like to conclude this summary with one of the recurring themes of the book, the importance of aggregation. I believe that future progress is most likely to come when aggregation is taken seriously, and when macroeconomic questions are addressed in a way that uses the increasingly plentiful and informative microeconomic data. ...They [the microeconomic data] also contain a great deal more information than the already over-used aggregate time-series data."

In this paper we try to take aggregation seriously and we use extensively microeconomic data. In emphasizing aggregation we might have neglected – perhaps too much to the taste of some readers – today's requirement that macroeconomic analysis should have a "microeconomic foundation". If the methodological approach of this paper turns out to be fruitful, then certainly, we would not hesitate to remedy this deficiency.

Our main goal in this paper is to answer the following question: under what circumstances can one predict aggregate consumption in period t from observations that refer to periods prior to period t. Or, more modestly, under what circumstances can one predict aggregate consumption in period t conditional on a hypothetical level of aggregate income in period t from observations prior to period t? The problem thus is *to establish a stable,* that is to say, *time invariante functional relationship K which links aggregate income in period t and other "determinants" of consumption* – which refer to periods prior to period t – *with aggregate consumption in period t.*

By "establishing" such a *consumption function* we understand to derive it from hypotheses which should be falsifiable by empirical data (time series of cross-section data) and actually are supported by empirical evidence. Before one can discuss the form of a functional relationship one has to specify the "variables" that are involved. Hence, the hypotheses on which the theory is build should lead to a specification of the nature of the other "determinants" that enter the consumption function (e.g., past income and/or past consumption).

The basic hypothesis on which the analysis of this paper is built is that of a *structurally stable evolution*. In our context this hypothesis says that the normalized income distributions and the suitably transformed Engel consumers' expenditure curves change slowly over time (for a precise statement see section 3). This hypothesis is not rejected by empirical evidence.

Structural stability is a strong restriction on the evolution of the cross-section micro data of an economy. The hypothesis allows to predict "approximately" mean consumers' expenditure in period t from the microecomomic data of a nearby period s (see the Proposition in section 3).

Any theory of consumption function should *imply* the observed regularities of aggregate consumption and aggregate income data such as the high linear association of aggregate income and consumption for short-run time series (see Empirical Fact A., section 2). Such a theory should also *allow* for the relative stability of the ratio of aggregate consumers' expenditure and aggregate income (average propensity to consume, APC) that is claimed for the U.S. economy. In addition, of course, a satisfactory theory should lead to accurate predictions.

The paper is organized as follows: in section 2 we recall the well-known empirical facts on aggregate consumption and income. Since in this paper we need *time series of cross-section data*, we use the data from the U.K. Family Expenditure Survey (FES) from 1968 to 1984. Some information on these data is given in section 3. The main contribution of this paper is presented in section 4. The hypothesis of a structurally stable evolution is discussed and the Proposition summarizes the consequences of this hypothesis. Finally, in section 5 we illustrate the predictive power of the model of section 4.

2 Empirical facts on aggregate consumption and income

The relationship between aggregate consumers' expenditure and aggregate disposable income is one of the most thoroughly researched topics in quantitative economics. The numerous empirical studies come to quite different conclusions which sometimes seem to be contradictory. Yet the following two empirical facts of descriptive statistics are well established:

A. If "aggregate consumers' expenditure" is plotted against "aggregate disposable income" for a short- or medium-run series of periods, then this scatter diagram shows a *strong linear association* (i.e., the correlation coefficient is near to one). Furthermore, the scatter diagram "matches approximately" a straight line with a slope less than one and strictly positive intercept.

How is this straight line determined? A natural candidate is the S-D line[1] (standard deviation line). In the literature one usually refers to the L-S line[2] (least square line). If the correlation coefficient r is near to one then the difference is small since the slope of the L-S line is r times the slope of the S-D line (see Figure 1 and 5).

B. The average propensity to consume (i.e., the ratio of aggregate consumers' expenditure and aggregate disposable income) is "relatively stable" for very long series of periods.

These claims require, of course, a definition of "aggregate disposable income" and "aggregate consumers' expenditure". Does "aggregate" mean "total" or "mean" (i.e., per capita)? Are income and expenditure measured in constant prices or current prices?

We use the following notation:

[1]The S-D line goes through the point of averages and the slope is equal to the ratio of the standard deviation of consumers' expenditure and the standard deviation of disposable income, see e.g. FREEDMAN et al. (1978).

[2]The L-S line minimizes the mean square error. The L-S line also goes through the point of averages and its slope is r times the slope of the S-D line.

$\bar{x}_t$ denotes mean disposable income in period t

$\bar{c}_t$ denotes mean consumers' expenditure in current prices in period t (if expenditures on durables are excluded we will explicitely say so, otherwise expenditure means total spendings on all items of consumption)

π_t denotes a price index for period t with respect to some base period $\tau, \pi_\tau = 1$.

$\tilde{x}_t = \bar{x}_t/\pi_t$ denotes "real" mean disposable income in period t

$\tilde{c}_t = \bar{c}_t/\pi_t$ denotes "real" mean consumers' expenditure

If "mean" is replaced by "total" we use capital letters. Thus, $\bar{X}_t$ is total disposable income in period t, $\bar{C}_t$ is total consumers' expenditure in current prices...

Does the empirical fact A refer to data points $(\bar{x}_t, \bar{c}_t)$, $(\tilde{x}_t, \tilde{c}_t)$, $(\bar{X}_t, \bar{C}_t)$ or $(\tilde{X}_t, \tilde{C}_t)$? Does it matter? Most macro economic textbooks use *total* real income and *total* real consumers' expenditure, i.e., $(\tilde{X}_t, \tilde{C}_t)$. For example, EVANS (1969), PARKIN and BADE (1982) or DORNBUSCH and FISCHER (1987). SAMUELSON argued that one should use *mean* real income and *mean* real consumers' expenditure, i.e., $(\tilde{x}_t, \tilde{c}_t)$, " the same real income divided up among more people cannot be expected to yield the same real consumption expenditure", SAMUELSON (1941), p. 252. In Figure 1 is shown the scatter diagram of $(\tilde{x}_t, \tilde{c}_t)$ for the U.S. from 1950-69. The correlation coefficient is $r = 0.9985$; thus the difference between the S-D line and the L-S line is so small that the two lines cannot be distinguished in the diagram.

Figure 1: Scatter diagram of $(\tilde{x}_t, \tilde{c}_t)$ for the U.S. from 1950-69 in 1958 dollars. The data are from *The Economic Report of the President 1974.*

In the formulation of the empirical facts the time structure of the sequence $(\tilde{x}_t, \tilde{c}_t)_{t\in T}$ was completely ignored. We emphasize that we did not interpret the L-S line as a "Keynesian consumption function"! In particular, we did not interpret the slope of the L-S line (or the S-D line) as a "marginal propensity to consume". We claim a story *linear association*, however, not a *causation*, between the variables $\tilde{x}_t$ and $\tilde{c}_t$.

Up to now we looked at the series of points $(\tilde{x}_t, \tilde{c}_t)_{t\in T}$ without a theoretical model of a relationship between income and consumption. In econometrics one considers the observed time series $(\tilde{x}_t, \tilde{c}_t)_{t\in T}$ as a realization of a stochastic process $(x_t, c_t)_{t\in T}$ that might have a complicated stochastic time structure (even more generally, (x_t, c_t) can be the projection of a multi-variable stochastic process). If one takes this view, then it is not justified to link the coefficients of the L-S line to the structural parameters of the stochastic process (x_t, c_t). To be more specific, consider, as an example, the special case where the stochastic process $(x_t, c_t)_{t\in T}$ is such that the conditional expectation of c_t given $x_t = \tilde{x}_t$ depends only on $\tilde{x}_t$ and, furthermore, is linear in $\tilde{x}_t$, i.e.,

$$\mathbb{E}(c_t | X_t = \tilde{x}_t) = c_1 + c_2 \tilde{x}_t \ ,$$

where c_1 and c_2 are (time invariant) parameters of the stochastic process. Even in this hypothetical case it is problematic without further very restrictive assumptions on the time structure of the stochastic process (x_t, c_t) to conclude from the empirical fact A that the slope (intercept, respectively) of the L-S line can be used as an estimator for the parameter c_2 (c_1, respectively) of the stochastic process (x_t, c_t), see, e.g., Granger and Newbold (1974). Such unjustified conclusions are however often made in the earlier literature and are then called "stylized" empirical facts, which now refer to the parameters of a model. An unjustified conclusion from an empirical fact does not become a "stylized" fact but remains what it is; an unjustified conclusion!

The empirical fact B is essentially based on Kuznets (1942) who analyzed data for the U.S. from 1869 to 1938. He considered moving ten year averages of the ratio of consumption and "national" income and concluded that this sequence is "relatively stable", varying between 0.84 and 0.89. Stability of the consumption-income ratio has been reconfirmed for the U.S.

by GOLDSMITH (1955) for "personal" income with a somewhat lower value. Later studies on historical data for other countries than the U.S. have not confirmed the long-run "relative stability" of the average propensity to consume. There seems to be a downward trend for several countries; see in particular MADDISON (1992).

The empirical facts A and B are obviously not compatible with the naive notion of a "Keynesian" consumption function $\tilde{c}_t = K(\tilde{x}_t)$, where K is a stable (time invariant) function linking in every period t real mean disposable income to real mean consumers' expenditure. This apparent conflict between the empirical facts and the naive notion of a consumption function was the main motivation for further developements of the notion of a consumption function (DUESENBERRY (1949), MODIGLIANI and BRUMBERG (1954) and FRIEDMAN[3] (1957)). For a history on the subject see THOMAS (1989) or SPANOS (1989).

This view is shared by the profession. For example, DAVIDSON et al. (1978) write:"...we do wish to stress that most theories of the consumption function were formulated to reconcile the low short-run MPC with the relative stability claimed for the APC over medium to long data periods."

The present paper is an attempt in the same tradition.

[3] FRIEDMAN (1957) p.4 "Current consumption expenditure was highly correlated with income, the marginal propensity to consume [the slope of the L-S line] was less than unity, and the marginal propensity was less than the average propensity to consume, so the percentage of income saved increased with income. But then a serious conflict of evidence arose. Estimates of savings in the United States by Kuznets for the period since 1899 revealed no rise in the percentage of income saved during the past century despite a substantial rise in real income..."

3 The Data

The model of consumers' expenditure that we shall propose in the next section is not formulated in terms of aggregates only, but also distributions of individual income and expenditure are involved. If one wants to confront such a model with empirical data then one needs *time series of cross-section data.* For this reason we use the data from the U.K.-Family Expenditure Survey.[4] The use of survey data in analysing aggregate expenditure and aggregate income can be criticized since it is well-known that these data suffer from more or less serious under reporting of the households in the sample. The reliability of these survey data has been questioned mainly for disposable income, in particular, for self-employment income and occupational pensions. For a critical valuation of the income data we refer to ATKINSON and MICKLEWRIGHT (1983). Inspite of the deficiency of the FES data, we did not make any adjustments of the FES data. To compensate somewhat the deficiency in the income data we consider not only the whole population but also subpopulations, for example, all households with head of households not self-employed and/or not retired. Figure 2 illustrates the time path of mean disposable income $\bar{x}_t$ and mean consumers' expenditure $\bar{c}_t$ *in current prices* from 1968 to 1984.

Figure 2: Time path of mean disposable income and mean consumers'expenditure in current prices.

Figure 3 illustrates the time path of mean disposable income $\tilde{x}_t$ and mean consumers' expenditure $\tilde{c}_t$ *in constant prices.* From 1968 to 1976 real expenditure follows well the movements of real income. From 1976 to 1977 real

[4] KEMSLEY et al. (1980).

income falls but real expenditure raises and from '79 to '80 real income raises sharply and real expenditure falls.

Figure 3: Time path of mean disposable income and mean consumers'expenditure in constant prices.

Figure 4-a shows the scatter diagram of $(\bar{x}_t, \bar{c}_t)$. Mean consumers' expenditure includes expenditure on durables. The correlation coefficient is very high, $r = 0.9997$. The S-D line and the L-S line cannot be distinguished in the diagram. The slope is 0.93 and both intercepts are positive.

Figure .3.4-a: Average consumption (incl. durables) and income in current prices, standard deviation and regression line.

Figure .3.4-b: Average consumption (excl. durables) and income in current prices, standard deviation and regression line.

Figure 5-a shows the scatter diagram of $(\tilde{x}_t, \tilde{c}_t)$. Mean consumers' expenditure includes expenditure on durables. The correlation coefficient is $r = 0.9535$. The S-D line has a slope of 0.72 and is slightly different from the L-S line, whose slope is 0.69. Both intercepts are positive.

Figure .3.5-a: Average consumption (incl. durables) and income in constant prices, standard deviation and regression line.

Figure .3.5-b: Average consumption (excl. durables) and income in constant prices, standard deviation and regression line.

Figure 6 gives the APC for 1968 to 84. The mean is 0.95 and the standard deviation is 0.018.

Illustrations of the distribution of disposable income and expenditures are given in the next section.

4 A simple dynamic model of consumers' expenditure

We want to model consumers' expenditure for a large and heterogenous population of households. Households may have quite different consumption

Figure .3.6: APC, Average Propensity to Consume, FES 1968-1984.

behavior and they have typically not the same disposable income.

We denote by ρ_t the *density* of the *distribution of disposable income* in period t and by $\bar{x}_t = \int x\rho_t(x)dx$ the *mean disposable income* in period t.

The "expenditure" of a household in period t in current prices is the total spending on "all" consumption goods and "all" services. Sometimes we shall exclude expenditure on durables. For every income level x, we denote by $c_t(x)$ the *mean expenditure* in period t in current prices of *all households with income level* x. The function c_t is called the *Engel consumers' expenditure curve* in current prices in period t.

The *mean consumers' expenditure* in current prices in period t is then given by

$$\bar{c}_t = \int c_t(x)\rho_t(x)dx \ .$$

It is important to emphasize that the income density ρ_t and the Engel consumers' expenditure curve c_t can be estimated from cross-section data as given, for example, by the U.K. Family Expenditure Survey.

The expenditure in period t of a household depends on its disposable income of that period, but also on the current prices of consumption goods and services and on many other determinants of demand, like, preferences, past income, past consumption and past prices, expected future income (permanent income or life cycle income) and expected future prices as well as on demographic variables. For the purpose of this section we do not need to model consumer's expenditure as a function of all these determinants of demand. Among all these determinants of demand, disposable income plays a special rôle in so far as we stratify the population by disposable income. Yet this does not mean that for the explanation of individual expenditure disposable income is the most relevant variable. We call the model of this section "simple", since consumer's expenditure is not defined in terms of a microeconomic model of demand, like in HILDENBRAND (1994).

We now consider a sequence of periods $t \in T$ and the corresponding sequence $(\rho_t, c_t)_{t \in T}$ of income densities and Engel consumers' expenditure curves. Obviously the functions ρ_t and c_t change over time. The question, of course, is how do they change? Can one describe the evolution of (ρ_t, c_t) over time? For example, are these functions time invariant after certain suitable transformations? The numerous empirical studies on income distributions and income-inequality contain valuable information on the evolution over time of income distributions, see, e.g., ATKINSON (1976) and ATKINSON (1983). To our knowledge, there are only very few studies on the evolution over time of Engel consumers' expenditure curves. The most relevant paper in our context is HÄRDLE and JERISON (1990).

To answer the above question we estimated with non-parametric methods the income density and the Engel curve for every year from 1968 to 1984 for the U.K. using the data from the FES. We used kernel density and kernel regression estimators to determine income densities and Engel curves. For details of the particular kernel methods that we applied see, e.g., ENGEL and KNEIP (1994). The empirical findings of this data analysis, that are relevant in the present context, can be summarized as follows:

I. *The normalized income densities*

$$\rho_t^*(\xi) = \bar{x}_t \rho_t(\bar{x}_t \xi) \qquad , t \in T$$

change very slowly over time, that is to say

$$\rho_s^* \approx \rho_t^* \tag{1}$$

for two nearby periods s and t.

The literaturee on income distributions and income inequality contains in various forms empirical evidence for the slow change of the normalized income densities ρ_t^*.

In Figure 7 are shown the estimates of ρ_t^* for $t = 1968 \ldots 1984$, and in Figure 8 are shown estimates of ρ_t^* for five consecutive years, 1968-1872, 1969-1973 etc. The income domain $[0.2, 3.0]$ contains ca. 98% of the households in the sample.

Figure .4.7: Normalized income densities, FES 1968-1984.

Clearly the normalized densities are not time invariant yet they change quite slowly over time. The pronounced bimodality of the income densities disappears if one leaves out retired households; see Figure 9. As remarked above, the data for disposable income of self-employed households is quite unreliable. We therefore estimated also the income densities for the sub-population consisting of non-retired and not self-employed households, see Figure 10. Due to large outliers in the income data (for example, the sample of 1982 contains a household whose disposable income is 56 times the mean disposable income!) the estimates of the variance (or higher moments) of the income distribution are not robust. If one is interested in the evolution of "income-inequality" one might want to look at the evolution of some of the various measures of inquality. For example, the estimates of the interquartile range of ρ_t^* (which are robust) have a very small tendency to enlarge over time. For the purpose of this paper, however, we are satisfied with the observation (by visual inspection) that the normalized income densities ρ_t^* change slowly over time.

Figure .4.8: Normalized income densities, FES 1968-1984.

Figure .4.9: Normalized income densities, FES 1968-1984 head of household not retired.

Figure .4.10: Income densities, FES 1968-84, head of household not retired and not self-employed.

Figure .4.11: π-scaled Engel consumers' expenditure curve of period t in constant prices (including durables).

Figure .4.12: π-scaled Engel consumers' expenditure curve of period t in constant prices (excluding durables).

Figure .4.13: π-scaled Engel consumers' expenditure curve of period t in constant prices (without durables); exluding self-employed and retired households.

Figure .4.14: π-scaled Engel consumers' expenditure curves in constant prices from 1968 to 84.

Figure .4.15: $\bar{\pi}$-scaled Engel consumers' expenditure curve of period t in constant prices (including durables).

II. *The π-scaled Engel consumers' expenditure curves in constant prices* c_t^π, i.e., *the functions*

$$\xi \mapsto \frac{1}{\pi_t}\bar{c}_t(\pi_t\xi) =: c_t^\pi(\xi) \ ,$$

change slowly over time, that is to say

$$c_s^\pi \approx c_t^\pi \tag{2}$$

for two nearby periods s and t.

In the Figures 11-13 are shown the estimates of the π-scaled Engel curves c_t^π for five consecutive years. With the exception of the years 1968 and 1970 the π-scaled Engel curves change remarkably slowly. Surely they actually change over time as can be seen from Figure 14 where the estimates of all π-scaled Engel curves from 1968 to 1984 are shown.

Remark: In the search for a suitable transformation of the Engel curves c_t that leads to approximate time invariance, the "π-scaled" Engel curves did not come out of an trial and error process. Actually, we considered first the transformation

$$\xi \mapsto c_t(\bar{x}_t\xi) \ .$$

The estimates of the sequence of these curves suggested a multiplicative factor, as was already observed by Härdle and Jerison (1990). Hence we considered the transformation

$$\xi \mapsto \frac{1}{\pi_t}c_t(\bar{x}_t\xi) \ .$$

The estimates of these functions changed slowly over time. In order to understand why this particular ad hoc choosen transformation led to approximate invariance we considered a micro economic model of consumer demand. A simple hypothesis on the evolution of the distribution of consumers' characteristics then strongly suggested the time invariance of the π-scaled Engel curves.

The main idea is simple and the heuristic argument can be sketched now. The expenditure in current prices of an individual consumer is defined by $p_t \cdot f(p_t, x, \alpha_1, \alpha_2, \ldots)$ where p_t denotes the price vector in period t and $f(p_t, x, \alpha_1, \alpha_2, \ldots)$ the demand vector which depends not only on current prices and income but also on many other determinants of demand $\alpha = (\alpha_1, \alpha_2, \ldots) \in A$ as explained above.

The population of consumers in period t is described by a joint distribution μ_t of the consumers' characteristics x and α. Let $\mu_t|x$ denote the conditional distribution of α given the income level x. Then we obtain

$$c_t(x) = p_t \cdot \int_{\mathcal{A}} f(p_t, x, \alpha) d\mu_t|x \ .$$

An evolution of the population of consumers over time is now described by an evolution (μ_t) of the joint distributions of consumers' characteristics. Assume that the marginal distributions of disposable income evolve over time as described in I. It remains to specify the evolution of the conditional distributions $\mu_t|x$. A simple and natural hypothesis would be

$$\mu_t|\bar{x}_t\xi = \mu_s|\bar{x}_s\xi$$

for all income levels ξ and nearby periods t and s. This models a "pure income change"; households "keep in average their characteristics". Then one obtains

$$\int_{\mathcal{A}} f(p_t, x, \alpha) d\mu_t|x = \int_{\mathcal{A}} f(p_t, x, \alpha) d\mu_s|\frac{\bar{x}_s}{\bar{x}_t}x \ .$$

If one assumes that consumers' demand functions are homogeneous in (p, x), i.e., $f(p_t, x, \alpha) = f(\lambda p_t, \lambda x, \alpha)$ and if one considers the ideal case where prices change proportional, i.e., $p_t/\pi_t = p_s/\pi_s$, then it follows that

$$\int_{\mathcal{A}} f(p_t, x, \alpha) d\mu_t|x = \int_{\mathcal{A}} f(p_s, \frac{\pi_s}{\pi_t}x, \alpha) d\mu_s|\frac{\bar{x}_s}{\bar{x}_t}x \ .$$

We shall show later that for nearby periods t and s one has $\pi_s/\pi_t \approx \bar{x}_s/\bar{x}_t$; they are, of course, not equal, yet not very different. Consequently, on the right hand side of the above equation we can substitute either π_s/π_t for $\bar{x}_s/\bar{x}_t$ (i.e., replace $\mu_s|\frac{\bar{x}_s}{\bar{x}_t}$ by $\mu_s|\frac{\pi_s}{\pi_s}\pi_t$) or $\bar{x}_s/\bar{x}_t$ for π_s/π_t (i.e., replace $f(p_s, \frac{\pi_s}{\pi_t}x, \alpha)$ by $f(p_s, \frac{\bar{x}_s}{\bar{x}_t}, \alpha)$). In which case is the approximation better? Income has a direct impact on demand while the dependence of $\mu_t|x$ might be relatively weak. Consequently, one might expect that

$$\int_{\mathcal{A}} f(p_t, x, \alpha) d\mu_t|x \approx \int_{\mathcal{A}} f(p_s, \frac{\pi_s}{\pi_t} x, \alpha) d\mu_s|\frac{\pi_s}{\pi_t} x \ ,$$

and since we considered the case where $p_t/\pi_t = p_s/\pi_s$, we obtain

$$\frac{1}{\pi_t} p_t \cdot \int f(p_t, x, \alpha) d\mu_t|x \approx \frac{1}{\pi_s} p_s \cdot \int f(p_s, \frac{\pi_s}{\pi_t} x, \alpha) d\mu_s|\frac{\pi_s}{\pi_t} x$$

i.e.,

$$\frac{1}{\pi_t} c_t(x) \approx \frac{1}{\pi_s} c_s \left(\frac{\pi_s}{\pi_t} x \right) \ ,$$

for all income levels x, hence $c_t^\pi \approx c_s^\pi$.

This heuristic argument suggests that the π-scaled Engel curves are approximately time invariant, and that the extent of the invariance of the π-scaled Engel curve is stronger than the extent of the invariance of the $\bar{x}$-scaled Engel curves in constant prices

$$\xi \mapsto \frac{1}{\pi_t} \bar{c}_t(\bar{x}_t \xi) \ .$$

The empirical findings confirm this claim. Compare the estimates of Figure 11 and 15.

III. *The price-income evolution is "normal"*, i.e. , Δ_t/π_t (where $\bar{x}_t = \Delta_t \cdot \bar{x}_0$ and $\pi_0 = 1$) *change moderately overtime,* that is to say, $\bar{x}_t/\bar{x}_s \cdot \pi_s/\pi_t$ is in the order of magnitude of 1 for two nearby periods s and t, e.g.

$$\frac{\bar{x}_t}{\bar{x}_s} \cdot \frac{\pi_s}{\pi_t} \in [0.8, 1.2] \quad \text{for } |t - s| \leq 5 \tag{3}$$

Table 1 gives the values of Δ_t/π_t.

Δ_t/π_t	1	1.013	1.026	1.020	1.073	1.128	1.131	1.094	1.060	1.049
Year	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977
Δ_t/π_t	1.122	1.135	1.169	1.157	1.148	1.164	1.161			
Year	1978	1979	1980	1981	1982	1983	1984			

Table 1:

It follows from Table 1 that

$$\max_{\substack{s<t \\ |t-s|\leq 5}} \frac{\bar{x}_t}{\bar{x}_s}\frac{\pi_s}{\pi_t} = 1.113$$
$$\min_{\substack{s<t \\ t\leq 5}} \frac{\bar{x}_t}{\bar{x}_s}\frac{\pi_s}{\pi_t} = 0.927$$

The above empirical findings I-III are strong restrictions on the evolution of

$$(\rho_t, c_t, \pi_t)_{t\in T}$$

An evolution of (ρ_t, c_t, π_t) satisfying I to III is called *structurally stable*.

We emphasize that structural stability does not mean that the normalized income densities ρ_t^* and the π-scaled Engel consumers' expenditure curves c_t^π do not change at all. It is evident from Figure 7 and 14 that over a long series of periods these functions are not time invariant. Structural stability means only that the (up to now still unmodelled) changes of these functions are sufficiently slow in the following sense: if one considers any period s and *all nearby* perdiods t, say $t \in T(s) = \{s, s \pm 1, s \pm 2\}$, then it is "approximately" true that the functions ρ_t^* and c_t^π are time invariant.

Structural stability means short – or medium – run invariance; it does not mean long-run invariance. It should be clear that we do not claim that the evolution of (ρ_t, c_t, π_t) is always structurally stable. Such a claim could easily be rejected by empirical evidence. There are various reasons why a structural stable evolution might be interrupted or disturbed. For example, a drastic income tax reform in period s might lead (in the very short-run) to an income density ρ_{s+1} such that ρ_{s+1}^* is quite different from ρ_s^*. Also a drastic change in households behavior in period s due, for example, to a change in preferences or expectations on future prices or income, might lead to an Engel curve c_{s+1} such that c_{s+1}^π is quite different from c_s^π. Thus, structural stability refers to a "normal" evolution without such disturbances. The notion of structural stability, as used here, has been described in the literature, most clearly and explicitly in MALINVAUD (1983)[5] and (1991).[6]

IV. *The Engel consumers' expenditure curve is well approximated on the relevant domain of the income distribution by the following functional form*

$$c_t(x) \approx \alpha_t x + \beta_t x \cdot \log x \tag{4}$$

with $\alpha_t > 0$ and $\beta_t < 0$.

This functional form of an Engel curve, first used by WORKING (1943), is frequently used in empirical studies. In the case of expenditure on all consumption goods and all services this simple functional form seems to be quite acceptable. If one considers, however, Engel expenditure curves for certain subgroups of commodities, like expenditure on "food" or "clothes", then one needs a more general functional form. It has been shown by KNEIP (1993) that in this case the functional form

$$\sum_{n=0}^{3} \alpha_n x(log x)^n$$

is well justified (using the U.K. Family Expenditure Survey and the French Enquête Budget de Famille). This functional form can also be derived theoretically for a "heterogenous" population, Kneip (1993). In the Appendix we discuss alternative assumptions on the functional form of the Engel curves.

[5] MALINVAUD (1983), pp. 71-72 "Depuis longtemps divers auteurs ont attiré l'attention sur le fait que, mis à part un coefficient d'échelle évidemment variable, les distributions statistiques observées dans le monde économique et social présentaient une grande permanence. Elles se modifient peu avec le temps. Bien entendu cette assertion repose sur des preuves qui sont limitées par la difficulté que l'on a à trouver des statistiques suffisamment présises et comparables. Elle ne doit pas non plus être considérée comme absolue: de petites déformations dans les lois de distribution ont parfois leur importance sur les phénomènes étudiés. Néanmoins la stabilité des distributions statistiques est suffisamment générale pour fournir le plus souvent une bonne justification du raisonnement sur grandeurs agrégées".

[6] MALINVAUD (1991), p.166 "Pour justifier leurs raisonnements, les macroéconomistes invoquent d'ailleurs la 'stabilité des structures'. L'expression, rarement définie avec précision, fait référence, en particulier, à la lenteur avec laquelle les distributions statistiques se déforment autour des tendances générales de leurs grandeurs respectives".

Remark: If one accepts assumption (4) then the hypothesis $c_t^\pi = c_s^\pi$ implies that $\beta_t = \beta_s$ and $\alpha_t = \alpha_s + \beta_s \log \pi_s/\pi_t$. Consequently, if one accepts assumption (4) and the hypothesis of a structurally stable evolution, then one expects a high negative corrolation between α_t and π_t since $\alpha_t \approx \alpha_s \beta_s (1 - \dfrac{\pi_t}{\pi_s})$.

The following mathematical proposition shows that there exists a time invariant function K in the variables $\tilde{x}, \pi, \rho$ and c such that

$$\tilde{c}_t = K(\tilde{x}_t, \pi_s, \rho_s, c_s)$$

if the evolution is strictly structurally stable between the periods s and t.

Proposition. *Let the evolution of the income densities ρ_t and the Engel consumers' expenditure curves c_t be such that for the two periods s and t, the two normalized income densities ρ_s^* and ρ_t^* and the two π-scaled Engel curves c_s^π and c_t^π coincide, that is to say,*

(1) $$\rho_s^* = \rho_t^*$$

and

(2) $$c_s^\pi = c_t^\pi$$

Then one obtains

$$\bar{c}_t = \frac{\pi_t}{\pi_s} \int c_s \left(\frac{\pi_s \bar{x}_t}{\pi_t \bar{x}_s} \xi \right) \rho_s(\xi) d\xi \qquad (\bar{\mathrm{a}})$$

that is to say, mean consumers' expenditure in current prices in period t is a (time invariant) function in mean income $\bar{x}_t$, the price index π_t and the data in period s (i.e., π_s, c_s and ρ_s). In constant prices this implies

$$\tilde{c}_t = \frac{1}{\pi_s} \int c_s \left(\frac{\tilde{x}_t}{\tilde{x}_s} \xi \right) \rho_s(\xi) d\xi \qquad (\tilde{\mathrm{a}})$$

that is to say, mean consumers' expenditure in constant prices in period t is a function in mean real income $\tilde{x}_t$ and the data in period s.

Thus, strict structural stability between the periods s and t allows to predict $\bar{c}_t$ or $\tilde{c}_t$ conditional on $\bar{x}_t$ and π_t provided the data in period s are known.

If (on the relevant domain of the income distribution) the Engel consumers' expenditure curve of period s has the form

$$c_s(x) = \alpha_s x + \beta_s x \log x \tag{4}$$

then one obtains

$$\bar{c}_t = \frac{\bar{c}_s}{\bar{x}_s} \cdot \bar{x}_t + \beta_s \bar{x}_t \log\left(\frac{\bar{x}_t}{\bar{x}_s}\frac{\pi_s}{\pi_t}\right) \tag{$\bar{\text{b}}$}$$

or, in constant prices,

$$\tilde{c}_t = \left(\frac{\bar{c}_s}{\bar{x}_s} - \beta_s \log \tilde{x}_s\right)\tilde{x}_t + \beta_s \tilde{x}_t \log \tilde{x}_t \tag{$\tilde{\text{b}}$}$$

Thus, the functional form of the relation which links real consumption expenditure $\tilde{c}_t$ and real income $\tilde{x}_t$ given the data in period s i.e. $(\pi_s, \bar{c}_s, \beta_s, \bar{x}_s)$ is of the same type (yet with a different coefficient of the linear term) as the Engel consumers' expenditure curve.
Furthermore, one obtains

$$\tilde{c}_t = a_s \tilde{x}_t + b_s + \beta_s \tilde{x}_s\, g\left(\frac{\tilde{x}_t}{\tilde{x}_s}\right) \tag{c}$$

with $a_s = \dfrac{\bar{c}_s}{\bar{x}_s} + \beta_s < 1$, $b_s = -\beta_s \tilde{x}_s > 0$ *and* $g(z) = z - 1 - z\log z$. Around 1 the values of the function g are very small *(see Table 2).*

x	0.60	0.70	0.80	0.90	0.95	1	1.05	1.10	1.20	1.30	1.40
$g(x)$	0.093	0.050	0.021	0.005	0.001	0	0.001	0.004	0.018	0.041	0.071

Table 2:

Consequently, if the price-income evolution is normal, than one obtains

$$\tilde{c}_t \approx a_s \tilde{x}_t + b_s \ .$$

Proof. By definition of mean consumers' expenditure $\bar{c}_t$ we have

$$\bar{c}_t = \int c_t(x)\rho_t(x)dx \ .$$

Assumption (2) implies

$$\frac{1}{\pi_t}c_t(x) = \frac{1}{\pi_s}c_s(\frac{\pi_s}{\pi_t}x) \ .$$

Hence

$$\bar{c}_t = \frac{\pi_t}{\pi_s} \int c_s \left(\frac{\pi_s}{\pi_t} x \right) \rho_t(x) dx$$

Substituting $x = \frac{\bar{x}_t}{\bar{x}_s}\xi$ leads to

$$\bar{c}_t = \frac{\pi_t}{\pi_s} \int c_s \left(\frac{\pi_s}{\pi_t} \frac{\bar{x}_t}{\bar{x}_s} \xi \right) \frac{\bar{x}_t}{\bar{x}_s} \rho_t \left(\frac{\bar{x}_t}{\bar{x}_s} \xi \right) d\xi \ .$$

Assumption (1) implies

$$\rho_s(\xi) = \frac{\bar{x}_t}{\bar{x}_s} \rho_t \left(\frac{\bar{x}_t}{\bar{x}_s} \xi \right) \ .$$

Thus we obtain claim (ā), i.e.,

$$\bar{c}_t = \frac{\pi_t}{\pi_s} \int c_s \left(\frac{\pi_s}{\pi_t} \frac{\bar{x}_t}{\bar{x}_s} \xi \right) \rho_s(\xi) d\xi \ , \tag{ā}$$

which implies claim (ã), since by definition $\tilde{c}_t = \frac{1}{\pi_t}\bar{c}_t$ and $\tilde{x}_t = \frac{1}{\pi_t}\bar{x}_t$.

By assumption (3) one obtains

$$\begin{aligned} c_s \left(\frac{\pi_s}{\pi_s} \frac{\bar{x}_t}{\bar{x}_s} \xi \right) &= \\ &= \alpha_s \frac{\pi_s \bar{x}_t}{\pi_t \bar{x}_s} \xi + \beta_s \frac{\pi_s \bar{x}_t}{\pi_t \bar{x}_s} \xi \log \left(\frac{\pi_s \bar{x}_t}{\pi_t \bar{x}_s} \xi \right) \\ &= \frac{\pi_s \bar{x}_t}{\pi_t \bar{x}_s} \left(\alpha_s \xi + \beta_s \xi \log \xi \right) + \beta_s \xi \frac{\pi_s \bar{x}_t}{\pi_t \bar{x}_s} \log \frac{\pi_s \bar{x}_t}{\pi_t \bar{x}_s} \end{aligned}$$

Therefore (ā) implies

$$\bar{c}_t = \frac{\bar{x}_t}{\bar{x}_s} \int (\alpha_s \xi + \beta_s \xi \log \xi) \rho_s(\xi) d\xi + \beta_s \frac{\bar{x}_t}{\bar{x}_s} \log \frac{\pi_s \bar{x}_t}{\pi_t \bar{x}_s} \int \xi \rho_s(\xi) d\xi \ .$$

Since by assumption (3)

$$\int (\alpha_s \xi + \beta_s \xi \log \xi) \rho_s(\xi) d\xi = \bar{c}_s$$

and since $\int \xi \rho_s(\xi) d\xi = \bar{x}_s$ we obtain

$$\bar{c}_t = \frac{\bar{c}_s}{\bar{x}_s} \bar{x}_t + \beta_s \bar{x}_t \log \frac{\pi_s \bar{x}_t}{\pi_t \bar{x}_s} \tag{b̄}$$

and

$$\tilde{c}_t = \left(\frac{\bar{c}_s}{\bar{x}_s} - \beta_s \log \tilde{x}_s \right) \tilde{x}_t + \beta_s \tilde{x}_t \log \tilde{x}_t \ . \tag{$\tilde{\mathrm{b}}$}$$

Finally, $(\tilde{\mathrm{b}})$ implies

$$\begin{aligned} \tilde{c}_t &= \frac{\bar{c}_s}{\bar{x}_s} \tilde{x}_t + \beta_s \tilde{x}_s \cdot \frac{\tilde{x}_t}{\tilde{x}_s} \log \frac{\tilde{x}_t}{\tilde{x}_s} \\ &= \frac{\bar{c}_s}{\bar{x}_s} \tilde{x}_t + \beta_s \tilde{x}_s \left(\frac{\tilde{x}_t}{\tilde{x}_s} - 1 \right) + \beta_s \tilde{x}_s g \left(\frac{\tilde{x}_t}{\tilde{x}_s} \right) \end{aligned}$$

where the function g is defined by $g(z) = z - 1 - z \log z$. Obviously, $g(1) = 0$ and $g'(1) = 0$. The function g is tabulated in Table 2. Consequently,

$$\tilde{c}_t = \left(\frac{\bar{c}_s}{\bar{x}_s} + \beta_s \right) \tilde{x}_t - \beta_s \tilde{x}_s + \beta_s \tilde{x}_s g \left(\frac{\tilde{x}_t}{\tilde{x}_s} \right)$$

$$\tilde{c}_t = a_s \tilde{x}_t + b_s + \beta_s \tilde{x}_s g \left(\frac{\tilde{x}_t}{\tilde{x}_s} \right) \tag{c}$$

Since the Engel consumers' expenditure curve is convex, i.e., $\beta_s < 0$, one obtains $a_s < 1$ and $b_s > 0$. It is somewhat surprising that the coefficients a_s and b_s do not depend on higher moments of the income density ρ_s; they only depend on the first moment $\bar{x}_s$. This is a consequence of the simple functional form (4) of the Engelcurve; for details see the Appendix of this section.

Q.E.D.

The consequence of the Proposition for a structurally stable evolution $(\pi_t, \rho_t, c_t)_{t \in T}$

Choose any period $s \in T$ and consider all periods t which are nearby, for example $T(s) = \{s, s \pm 1, s \pm 2, \ldots\}$. The hypothesis of structural stability allows to apply the Proposition to the periods s and t with $t \in T(s)$. Hence it follows from assertion (c) of the Proposition that the vectors $(\tilde{x}_t, \tilde{c}_t)$, $t \in T(s)$, lie approximately on a straight line which is given by the equation

$$\tilde{c} = a_s \tilde{x} + b_s$$

with $a_s = \dfrac{\bar{c}_s}{\bar{x}_s} + \beta_s < 1$ and $b_s = -\beta_s \tilde{x}_s > 0$, that is to say,

$$\tilde{c}_t \approx a_s \tilde{x}_t + b_s$$

, for all $t \in T(s)$.

Consequently, *the hypothesis of structural stability implies for short- or medium-run time series the well-known empirical finding* A, that is to say, a strong linear association between $\tilde{x}_t$ and $\tilde{c}_t$, $t \in T(s)$.

For a strictly structurally stable evolution $(\pi_t, \rho_t, c_t)_{t \in T(s)}$ one obtains from assertion $(\tilde{a})$ that $\tilde{c}_t$ is a function in $\tilde{x}$ given π_s, ρ_s and c_s. Thus, the derivative of this function with respect to $\tilde{x}_t$ evaluated at $\tilde{x}_t = \tilde{x}_s$ is the *marginal propensity to consume* $MPC(s)$ in period s, i.e.,

$$\begin{aligned} MPC(s) &:= \partial_{\tilde{x}_t} \frac{1}{\pi_s} \int c_s \left(\frac{\tilde{x}_t}{\tilde{x}_s} \xi \right) \rho_s(\xi) d\xi |_{\tilde{x}_t = \tilde{x}_s} \\ &= \frac{1}{\pi_s} \int \partial_\xi c_s(\xi) \frac{\xi}{\tilde{x}_s} \rho_s(\xi) d\xi \ . \end{aligned}$$

Assumption (4) on the functional form of the Engel curve c_s therefore leads to

$$MPC(s) = a_s \ .$$

Of course, one obtains the same conclusion if one chooses instead of the period s a later period s' and all periods t, which are nearby s', say $T(s')$. Again the vectors $(\tilde{x}_t, \tilde{c}_t)$, $t \in T(s')$ lie approximately on a straight line, which is given by the equation

$$\tilde{c} = a_{s'} \tilde{x} + b_{s'}$$

with $a_{s'} < 1$ and $b_{s'} > 0$. However, *the two straight lines, in general, do not coincide* even if the evolution is strictly structurally stable. It might well be the case that the marginal propensity to consume is quite stable, that is to say, $a_s \approx a_{s'}$. Indeed, the empirical finding B suggests that the average propensity to consume $\bar{c}_s / \bar{x}_s$ is stable in the long run, and one can expect that the coefficient β_s does not change much; for example, for a strictly structurally stable evolution one obtains $\beta_s = \beta_{s'}$. Hence one might expect that

$$a_s = \frac{\bar{c}_s}{\bar{x}_s} + \beta_s \approx \frac{\bar{c}_{s'}}{\bar{x}_{s'}} + \beta_{s'} = a_{s'} \ .$$

However, the coefficients b_s and $b_{s'}$ are different. By definition $b_s = -\beta_s \tilde{x}_s$ and $b_{s'} = -\beta_{s'} \tilde{x}_{s'}$. Since real mean income $\tilde{x}_t$ changes overtime, see Table 1,

– there is typically an increasing trend – the straight line

$$\tilde{c} = a_s \tilde{x} + b_s$$

is not stable but has a tendency to move upward over time. This is the so-called "ratchet" effect of Samuelson (1943, p.34-5).

In summary, for a structurally stable evolution the straight line

$$\tilde{c} = a_s \tilde{x} + b_s$$

– which might be called the "linear Keynesian consumption function"– is an *approximate relation* between real consumption and real income for a short- or *medium-run time series.* However this straight line has to be adjusted as time goes on since it is not a stable (time invariant) relation.

A remark on MPC

In the literaturee one often finds the claim that "cross-section" MPC are smaller than "time-series" MPC. The validity of this claim obviously depends on the definition of these concepts. For a structurally stable evolution we defined the MPC by $a_s = \frac{\bar{c}_s}{\bar{x}_s} + \beta_s$, and the property (c) of the Proposition suggests that the coefficient a_s should be interpreted as a "time-series" MPC in period s. Now, if one defines the "cross-section" MPC in period s by the expression $\partial_x c_s(p_s, \bar{x}_s)$, that is to say, by the derivative of the Engel consumers' expenditure curve $c_s(p_s, \cdot)$ in period s evaluated at mean income $\bar{x}_s$, then the above claim is valid yet by an invalid concept of "cross-section" MPC. Indeed,

$$\begin{aligned} \partial_x c_s(p_s, \bar{x}_s) &= (\alpha_s \bar{x}_s + \beta_s \bar{x}_s \log \bar{x}_s)' \\ &= \beta_s + \frac{1}{\bar{x}_s}(\alpha_s \bar{x}_s + \beta_s \bar{x}_s \log \bar{x}_s) \ . \end{aligned}$$

Since the Engel curve $c_s(p_s, \cdot)$ is convex, i.e., $\beta_s < 0$, one obtains

$$c_s(p_s, \bar{x}_s) < \int c_s(p_s, x) \rho_s(x) dx = \bar{c}_s \ .$$

Hence, $\partial_x c_s(p_s, \bar{x}_s) < a_s$. In order to define the "cross-section" MPC one has to specify how the income density ρ_s changes, if mean income $\bar{x}_s$ changes *in period* s and one has also to specify whether with such a change of income in period s the Engel curve c_s is considered to be fixed or how the curve changes. Under the assumption that the *normalized* income densities and the Engel curve c_s remain invariant, and since

$$\bar{c}_s = \int c_s(x)\rho_s(x)dx = \int c_s(\bar{x}_s\xi)\rho_s^*(\xi)d\xi$$

one obtains that the "cross-section" MPC is defined by the expression

$$\partial_{\bar{x}_s} \int c_s(\bar{x}_s\xi)\rho_s^*(\xi)d\xi = \frac{1}{\bar{x}_s}\int \partial_x c_s(\rho_s x)x\rho_s(x)dx \ ,$$

which is greater than $\partial_x c_s(p_s, \bar{x}_s)$ and is actually equal to the MPC in period s as defined above.

Appendix to section 4

In Proposition 1 we assumed that the Engel consumers' expenditure curve c_s is well approximated by the function

$$c_s(x) = \alpha_s x + \beta_s x \log x \ .$$

If one considers expenditure on all consumption goods and services this seems to be an acceptable assumption. However, if one considers expenditure on food or housing then one needs a more general functional form. How would the conclusion (c) of Proposition 1 change if we generalize (4) to

$$c_s(x) = \alpha_s + \beta_s x \log x + \gamma_s x(\log x)^2 \qquad (4')$$

It is not hard to show that conclusion (c) becomes

$$\tilde{c}_t = a_s\tilde{x}_t + b_s + \left(\beta_s\tilde{x}_s + \frac{2}{\pi_s}\gamma_s \int \xi \log \xi \rho_s(\xi)d\xi\right) g\left(\frac{\tilde{x}_t}{\tilde{x}_s}\right) + \gamma_s\tilde{x}_s h\left(\frac{\tilde{x}_t}{\tilde{x}_s}\right) \quad (c')$$

$$\begin{aligned} \text{with} \quad a_s &= \frac{\bar{c}_s}{\bar{x}_s} + \beta_s + \frac{2}{\bar{x}_s}\gamma_s \int \xi \log \xi \rho_s(\xi)d\xi \\ b_s &= -\beta_s\tilde{x}_s - \frac{2}{\pi_s}\gamma_s \int \xi \log \xi \rho_s(\xi)d\xi \\ g(z) &= z - 1 - z\log z \\ h(z) &= z(\log x)^2 \end{aligned}$$

The functions $g(z)$ and $h(z)$ are tabulated in Table 2 and Table 3. Note that the marginal propensity to consume a_s now depends on the income distribution ρ.

x	0.60	0.70	0.80	0.90	0.95	1	1.05	1.10	1.20	1.30	1.40
$h(x)$	0.156	0.089	0.039	0.009	0.002	0	0.002	0.009	0.039	0.089	0.158

Table 3:

One might ask how conclusion (c) of Proposition 1 would change if one uses a different functional form for the Engel curve, for example,

$$c_s(x) = \alpha x + \beta x^2 \ , \tag{4''}$$

with $\alpha > 0$ and $\beta < 0$.

In this case one obtains

$$\tilde{c}_t = \left(\frac{\bar{c}_s}{\bar{x}_s} + \frac{m^2(\rho_s)}{\bar{x}_s} \beta_s \right) \tilde{x}_t - \frac{1}{\pi_s} \beta_s m^2(\rho_s) + \frac{1}{\pi_s} \beta_s m^2(\rho_s) h \left(\frac{\tilde{x}_t}{\tilde{x}_s} \right)$$

where $m^2(\rho_s) = \int x^2 \rho_s(x) dx$ and $h(z) = z^2 - 2z$. The function h is tabulated in Table 4.

x	0.60	0.70	0.80	0.90	0.95	1	1.05	1.10	1.20	1.30	1.40
$h(x)$	0.160	0.090	0.040	0.010	0.002	0	0.002	0.010	0.040	0.090	0.160

Table 4:

Thus,

$$\begin{array}{rrcl} & \tilde{c}_t & \approx & a_s \tilde{x}_t + b_s \\ \text{with} & a_s & = & \dfrac{\bar{c}_s}{\bar{x}_s} + \beta_s m^2(\rho_s) \dfrac{1}{\bar{x}_s} < 1 \quad \text{and} \\ & b_s & = & -\dfrac{1}{\pi_s} \beta_s m^2(\rho_s) > 0 \ . \end{array}$$

The coefficients a_s and b_s now depend on the second moment of the income density ρ_s.

Shortcomings of the simple dynamic model

First, in computing mean consumers' expenditure in current prices in period t, i.e.,

$$\bar{c}_t = \int c_t(x)\rho_t(x)dx$$

we substituted for the density $\rho_t(x)$ the density $\frac{\bar{x}_s}{\bar{x}_t}\rho_s\left(\frac{\bar{x}_s}{\bar{x}_t}x\right)$, which has the same mean – equality would follow from assumption (1), $\rho_t^* = \rho_s^*$ – and we substituted for the Engel curve $c_t(x)$ the function $\frac{\pi_t}{\pi_s}c_s\left(\frac{\pi_s}{\pi_t}x\right)$ – equality would follow from assumption (2), $c_t^\pi = c_s^\pi$. Consequently, if the evolution is not strictly structurally stable between the periods s and t then we make an error $d_{t,s}$ in computing $\bar{c}_t$ with the above substitutions. The hypothesis of structural stability was used for arguing that this error term $d_{t,s}$ is small if the periods t and s are nearby. This argument is based on a "visual inspection " of the curves ρ_t^* and c_t^π, $t \in T$; see Figures 8 and 11-13. For the purpose of explaining the empirical fact A this visual inspection might be sufficient. If, however, one wants to use the simple dynamic model for prediction then one has to be more careful, since the error might be biased. The following two examples illustrat this point.

1. Assume that the normalized densities ρ_t^* and ρ_s^* are *exactly equal* and that the $\pi -$ *scaled* Engel curves c_t^π and c_s^π are *approximately equal*, but $c_t(x) > \frac{\pi_t}{\pi_s}c_s\left(\frac{\pi_t}{\pi_s}x\right)$. It then follows that $d_{t,s} > 0$, that is to say, one would systematically underpredict mean consumers' expenditure in period t conditioned on $\bar{x}_t$, π_t and the data in period s.

2. Assume that the π-scaled Engel curves c_t^π and c_s^π are *exactly equal* and the densities $\rho_t(x)$ and $\frac{\bar{x}_s}{\bar{x}_t}\rho_s\left(\frac{\bar{x}_s}{\bar{x}_t}x\right)$ are *approximately equal* but variance $(\rho_t) <$ variance $\frac{\bar{x}_s}{\bar{x}_t}\rho_s\left(\frac{\bar{x}_s}{\bar{x}_t}x\right)$. Assumption (3) on the form of the Engel curves then implies that $d_{t,s} < 0$, that is to say, one would systematically overpredict mean consumers' expenditure in period t.

We certainly agree that the evolution of the densities ρ_t and the Engel curves

c_t has to be analysed more thoroughly. It might well be that one can do better than simply assuming that the sequence $(\rho_t^*, c_t^\pi)_{t \in T}$ is approximately invariant. Indeed, one has to analyse whether there is still some undiscovered "short-run dynamics" in the evolution of (ρ_t^*, c_t^π) which is not covered by the transformation of (ρ_t, c_t) that we used here. Or, is it possible to model the remaining error $d_{t,s}$ as a random variable with a variance which is increasing in $t - s$?

Second, the Engel consumers' expenditure curves c_t have not been derived from a microeconomic model of household demand. For this reason we called the model of this section 'simple'. An explicit microeconomic foundation for the Engel curves not only is required by today's scientific standards in the professional journals, but more importantly, might further the formulation of hypothesis on the evolution of (ρ_t^*, c_t^π) as just discussed. The above Remark on the π-scaled Engel curves gives an indication that a purely theoretical microeconomic foundation can indeed be useful.

It should be clear that a satisfactory microeconomic formulation of the simple model has to start from the assumption of a large and hetorogenous population of households. A "representative" household model cannot serve here as a microeconomic foundation. In a forthcoming paper we shall remedy the simplistic short cut which was used in the present section. This however requires a certain amount of mathematical techniques that we thought should be avoided in this paper.

5 Compatibility of the simple dynamic model with the data: Predictions

In the last section we derived for a structurally stable evolution of $(\rho_t, c_t)_{t \in T}$, that satisfies assumption IV, the relation

$$\tilde{c}_t = \frac{\bar{c}_s}{\bar{x}_s}\tilde{x}_t + \beta_s \tilde{x}_t \log\left(\frac{\tilde{x}_t}{\tilde{x}_s}\right) + d_{t,s} \ . \tag{5}$$

The hypothesis of structural stability implies that the error term $d_{t,s}$ is small for nearby periods t and s.

One can use relation (??) to predict real consumption $\tilde{c}_t$ conditional on $\tilde{x}_t$ from the data $\bar{c}_s$ and $\bar{x}_s$ of period s provided there is a way to estimate the parameter β_s.

By defintion, β_s is the parameter that determines the structure of the Engel consumers' expenditure curve

$$c_s(x) = \alpha_s x + \beta_s x \log x$$

which is defined as regression function; $c_s(x)$ is the average expenditure of all households with the level of income x.

For every period s an estimate $\hat{\beta}_s$ of the parameter β_s can thus be obtained by standard regression methods from cross-section data on income and expenditure of a representative sample of households.[7]

Prediction can then be based on the empirical model

$$\tilde{c}_t = \frac{\bar{c}_s}{\bar{x}_s}\tilde{x}_t + \hat{\beta}_s \tilde{x}_t \log\left(\frac{\tilde{x}_t}{\tilde{x}_s}\right) + d_{t,s} \ , \tag{6}$$

where $\hat{\beta}_s$ denots an estimate of the parameter β_s.

[7]It is well-known that the variance of the conditional distribution of expenditure given an income level is increasing in income (see, for example, HILDENBRAND (1994), Chpater 3, Figure 3.6 and 3.7). Instead of the Engelcurve, we therefore consider the income share curve $\omega_s(x) = c_s(x)/x$. This leads to a regression problem that is approximately homoscedastic. The parameters α_s and β_s are then estimated by the method of least squares.

In applications it will be advantageous to average the predictions of $\tilde{c}_t$ based on several nearby periods s, since averaging will reduce some of the variability inherent in the error terms $d_{t,s}$. For predicting real consumption $\tilde{c}_t$ in period $t \in T$ from the data in period $s \in T$ we might either average over *all nearby* periods,

$$s \in T(t,k) = \{t-k, \ldots, t-1, t+1, \ldots, t+k\} \cap T \ ,$$

or average only over *periods prior* to t,

$$s \in S(t,k) = \{t-k, \ldots, t-1\} \cap T \ .$$

The second case corresponds to predictions in the usual sense; the first case is used to show the compatibility of the model with the data.

For every integer $k = 1, \ldots, 5$ we consider the following two predictions for $\tilde{c}_t$:

$$\hat{c}_{T(t,k)} := \frac{1}{\#T(t,k)} \sum_{s \in T(t,k)} \frac{\bar{c}_s}{\bar{x}_s} \tilde{x}_t + \hat{\beta}_s \tilde{x}_t \log\left(\frac{\tilde{x}_t}{\tilde{x}_s}\right)$$

and

$$\hat{c}_{S(t,k)} := \frac{1}{\#S(t,k)} \sum_{s \in S(t,k)} \frac{\bar{c}_s}{\bar{x}_s} \tilde{x}_t + \hat{\beta}_s \tilde{x}_t \log\left(\frac{\tilde{x}_t}{\tilde{x}_s}\right) \ .$$

Two different criterions are used to measure the error of the different predictions $\hat{c}_t$ of $\tilde{c}_t$:

- the *average squared error*

$$ASE = \frac{1}{\#T} \sum_{t \in T} (\hat{c}_t - \tilde{c}_t)^2$$

- the *average percentage of deviation*

$$APD = \frac{1}{\#T} \sum_{t \in T} 100 \cdot \frac{|\hat{c}_t - \tilde{c}_t|}{\tilde{c}_t}$$

The results for the FES-data with $T = \{1968, \ldots, 1984\}$ are shown in Table 5. The data are normalized such that $\tilde{x}_{1968} = 100$.

	$\hat{c}_{T(t,k)}$		$\hat{c}_{S(t,k)}$	
k	ASE	APD	ASE	APD
1	2.37	1.18	2.75	1.20
2	1.26	0.83	1.56	0.97
3	1.42	0.91	1.79	0.97
4	1.66	0.97	2.02	1.01
5	1.45	1.20	1.92	0.99

Table 5:

Figure .5.16: Time path of disposable income $\tilde{x}_t$, consumers' expenditure $\tilde{c}_t$ and predicted $\hat{c}_{T(t,2)}$ expenditure (dotted line) in constant prices.

As one can see from Figure 16 the "predictions" $\hat{c}_{T(t,2)}$ of $\tilde{c}_t$ are very satisfactory up to the year 1976; the average percentage deviation for these years is less than 0.8%. However for later years the percentage deviation is larger (see Table 8). This is due to the fact that for later years real income $\tilde{x}_t$ and real consumption $\tilde{c}_t$ do not always move in the same direction (for example, from 76 to 77 and 79 to 80).

In order to obtain some further insight, we also analysed some subpopulations. Table 6 contains the results for the subpopulation of all households classified as "workers" by the FES, 1968-84.

Can one conclude from the empirical results reported in Tables 5 and 6 that the methodology of our approach and, in particular, that the simple dynamic model of section 4 is satisfactory? Are the predictions sufficiently accurate and how do they compare with alternative models?

The FES-data are *survey* data and therefore, in evaluating the above

	$\hat{c}_{T(t,k)}$		$\hat{c}_{S(t,k)}$	
k	ASE	APD	ASE	APD
2	1.09	0.78	1.42	0.85
3	1.05	0.76	1.48	0.87

Table 6:

empirical results, one has to take into account the sampling variance. Let (x_{it}, c_{it}), $i = 1, \ldots, n_t$, denote the sample in period t, where the sample size n_t is approximately 7000 households. The estimate $\hat{\bar{c}}_t$ of the true value $\bar{c}_t$ is $\frac{1}{n_t}\sum_i c_{it}$. If the random sample (c_{it}) is a representative sample from the whole population of households (i.e., c_{it}, $i = 1, \ldots, n_t$ are independent and identically distributed random variables) one obtains that variance $(\hat{\bar{c}}_t) = 1/n_t$ variance (c_{it}). The variance of the random variable c_{it} can be estimated. Consequently, the standard deviation, SD_t, from the true value $\bar{c}_t$ that is due to sampling variance is

$$SD_t = \left(\frac{1}{n_t}\ \mathrm{variance}(c_{it})\right)^{1/2} .$$

The average standard deviation, $\overline{SD} = 1/T \sum_{t \in T} SD_t$.

For the FES-data one obtains in the case of

$$\begin{aligned} &\text{all households} : \overline{SD} = 0.86\% \\ &\text{workers} \qquad\quad\ : \overline{SD} = 1.02\% . \end{aligned}$$

Given the size of these average standard deviations that are only due to sampling variance one can not expect essentially sharper predictions than those reported in tables 5 and 6.

It might be of interest to compare our results with the standard linear regression model. Of course, we are aware that the standard assumption for the linear regression model might not be satisfied since the time series might be highly autocorrelated.

One can "predict" real consumption $\tilde{c}_t$ in period t by the linear regression model using the data for all periods s in T, $s \neq t$ or only for periods $s \in$

linear regression model	ASE	APD
$T\backslash\{t\}$	1.90	1.08
$T(t,2)$	1.80	1.00
$T(t,3)$	1.31	0.86
$T(t,4)$	1.90	1.03
$T(t,5)$	1.81	1.07

Table 7:

$T(t,k) = \{t-k, \ldots, t-1, t+1, \ldots, t+k\}$. Table 7 gives the results for the FES-data.

Finally we predict real consumption $\tilde{c}_t$ for the periods $t \in \{1973, \ldots, 1984\}$ by using only data that are prior to period t. The following estimates are considered:

1)
$$\hat{c}_{S(t,k)} = \frac{1}{k}\sum_{s=1}^{k} \frac{\bar{c}_{t-s}}{\bar{x}_{t-s}}\tilde{x}_t + \hat{\beta}_{t-s}\tilde{x}_t \log\left(\frac{\tilde{x}_t}{\tilde{x}_{t-s}}\right) ,$$

   where the parameters $\hat{\beta}_{t-s}$ are estimated as explained above

2) linear model
$$\hat{c}_t = \hat{a}\tilde{x}_t + \hat{b}$$

   where the parametes $\hat{a}$ and $\hat{b}$ are estimated by least squares with the data from periods 1968 to $t-1$

3) local linear model
$$\hat{c}_t = \hat{a}\tilde{x}_t + \hat{b} ,$$

   where the parameters $\hat{a}$ and $\hat{b}$ are estimated by least squares with the data from the periods in $S(t,k) = \{t-k, \ldots, t-1\}$

4) Hall's model
$$\hat{c}_t = \hat{\gamma} \cdot \tilde{c}_{t-1}$$

   where the parameter $\hat{\gamma}$ is estimated by least squares with the data from periods $1968, \ldots, t-1$.

Table 8 gives the results of the average square error and the average percentage deviation for the whole population and the subpopulation of those households that are classified as workers by FES.

	all households		workers	
estimate	ASE	APD	ASE	APD
$\hat{c}_{S(t,2)}$	2.14	1.23	1.70	0.88
$\hat{c}_{S(t,3)}$	2.44	1.22	1.75	0.87
linear model, 1968,..., $t-1$	3.14	1.36	3.32	1.32
linear model, $S(t,4)$	4.33	1.70	2.95	1.25
linear model $S(t,5)$	3.11	1.34	3.07	1.28
Hall's model	7.97	2.03	13.23	2.99

Table 8:

References

Atkinson, A. (1976) : *The Personal Distribution of Incomes;* London. George Allen & Unwin Ltd.

Atkinson, A. (1983) : *The Economics of Inequality;* 2nd Edition, Clarendon Press, London.

Atkinson, A. and J. Micklewright (1983) : On the Reliability of Income Data in the Family Expenditure Survey 1970-1977; *The Journal of the Royal Statistical Society 146, 1: pp. 33-61.*

Davidson, J., Hendry, D., Srba, F. and S. Yeo (1978) : Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom; *Economic Journal, 88, 661-92.*

Dornbusch, R. and S. Fischer (1987) : *Macroeconomics;* 4th Edition, New York, McGraw-Hill.

Duesenberry, J. (1949) : *Income, Saving and the Theory of Consumer Behavior;* Cambridge, Mass. Harvard University Press.

EVANS, M. (1969) : *Macroeconomic Activity;* Harper International Edition.

ENGEL, J. and A. KNEIP (1994) : Recent Approachens to Estimating Income Distributions, Engel Curves and Related Functions; *Disc. Paper No. A-442, Sonderforschungsbereich 303, University of Bonn.*

FREEDMAN D., PISANI R. and R. PURVES (1978) : *Statistics;* Norton & Company.

FRIEDMAN, M. (1957) : *A Theory of the Consumption Function;* Princeton. Princeton University Press.

GOLDSMITH, R. (1955) : *A Study of Saving in the United States;* Vol. 1 (Princeton, N.J.: Princeton University Press for NBER.

GRANGER, C. and P. NEWBOLD (1974) : Spurious Regressions in Econometrics; *Journal of Econometrics, 2, pp. 111-20.*

HÄRDLE, W. and M. JERISON (1990) : Cross-Section Engel Curves over Time; *Recherches Economiques de Louvain 57, 1525-1549.*

HILDENBRAND, W. (1994) : *Market Demand: Theory and Empirical Evidence;* Princeton University Press, Princeton, New Jersey.

KEMSLEY, W., REDPATH, R. and M.HOLMES, eds (1980) : *Family Expenditure Survey Handbook;* London: Her Majesty's Stationary Office.

KUZNETS, S. (1942) : *Uses of National Income in Peace and War;* NBER, Occasional Paper 6.

MADDISON, A. (1992) : A long-run perspective on saving; *Scandinavian Journal of Economics, 94, 181-96.*

MALINVAUD, E. (1981) : *Théorie macro-économique;* Dunod, Paris.

MALINVAUD, E. (1991) : *Voies de la Recherche Macroéconomique;* Editions Odile Jacob, Paris.

MODIGLIANI, F. and R. BRUMBERG (1954) : 'Utility Analysis and the Consumption Function: an Interpretation of Cross-Section Data', in Kenneth K. Kurihara (ed.); *Post-Keynesian Economics,*New Brunswick, N.J. Rutgers University Press, pp. 388-436.

PARKIN, M. and R. BADE (1982) : *Modern Macroeconomics;* Deddington, Philip Allan.

THOMAS, J. (1989) : The early Econometric History of the Consumption Function; *Oxford Economic Papers, 41, pp. 131-149.*

SAMUELSON, P. A. (1941) : 'Appendix to Chapter 11: A Statistical Analysis of the Consumption Function', in Hansen, A.H.; *Fiscal Policy and Business Cycles,* Allen & Unwin,*250-61.*

SPANOS, A. (1989) : Early Empirical Findings on the Consumption Function, Stylized Facts or Fiction: a Retrospective View; *Oxford Economic Papers, 41, pp. 150-169.*