Permanent income, heterogeneity
and the error correction mechanism

by
Mario Forni *
Marco Lippi **

Ottobre 1995

* Dipartimento di Economia Politica
Viale Berengario, 51
41100 Modena (Italia)
e-mail: forni@merlino.unimo.it

** Dipartimento di Scienze Economiche
Via Cesalpino, 14
00161 Roma (Italia)
e-mail: lippi@giannutri.caspur.it
Abstract

Recent papers have shown that allowing for \\n\textit{heterogeneity and incomplete information} the Permanent Income Hypothesis can be partially reconciled with evidence. In particular, autocorrelation of first differences, excess sensitivity and smoothness can be obtained as features of aggregate consumption. However, due to insufficient heterogeneity, the models proposed are usually singular and therefore not suitable for discussing and modeling cointegration. In this paper we show that allowing for more complex and realistic heterogeneity: (1) Autocorrelation, excess sensitivity and smoothness can still be obtained for aggregate consumption; (2) Aggregate consumption and total income follow a non-singular and general error correction model. Thus, contrary to the results obtained in the representative agent case, the permanent income hypothesis, with heterogeneous agents, is not at odds with the line of research based on cointegration and error correction mechanisms. A simple exercise, based on two groups of consumers leads to an estimable overidentified model. The latter is estimated using U.S. data from 1964:1 to 1991:4. Over this period total income and consumption are cointegrated, as predicted by the theory, and the overidentifying restrictions are not rejected.
PERMANENT INCOME, HETEROGENEITY AND THE ERROR CORRECTION MECHANISM

Introduction

In the last fifteen years an impressive macroeconomic literature has been devoted both to theoretical development and empirical testing of the permanent income hypothesis under rational expectations (PIH) as it was worked out by Hall in 1978. Hall’s result, in its simplest version, implies that aggregate consumption is a random walk whose changes are proportional to labor income innovations. Important consequences are: (i) Consumption changes are independent of past labor income changes; (ii) If labor income is I(1) and its measure of persistence is greater than unity, then consumption should be more volatile than income (Deaton (1987)); (iii) If labor income is I(1) then, as shown in Campbell (1987), total income and consumption are cointegrated by the vector (1 − 1); however, the implied error correction mechanism has a very special form: since consumption change is a pure innovation, all the coefficients in the consumption equation are zero.

The implications of Hall’s theory have been tested employing both macro and panel data by many authors. A partial list includes Flavin (1981), Davidson and Hendry (1981), Hall and Mishkin (1982), Hayashi (1985), Campbell (1987), Engle and Granger (1987), Stock and West (1988), West (1988), Campbell and Deaton (1989), Zeldes (1989), Runkle (1991), Attanasio and Weber (1992). The PIH has generally been rejected by the data, but the micro evidence is quite different from the macro one, indicating substantial aggregation effects. Firstly, at the micro level the evidence is somewhat mixed, while the rejection at the macro level is quite strong. Secondly, the ‘excess sensitivity’ has opposite signs: aggregate consumption change is positively correlated with past aggregate income change, while the correlation is negative (and smaller) in the micro data. Moreover, individual income appears to be negatively autocorrelated, while aggregate income is positively autocorrelated (Pischke 1991, Deaton 1992).

These findings have stimulated a literature aimed at reconciling micro and macro evidence. In this literature, consumers are assumed to behave according
to the PIH, but the representative agent hypothesis is abandoned, allowing for individual heterogeneity. Clarida (1991), Gali (1990) and Attanasio and Weber (1991) are concerned with overlapping generations, whereas Lippi (1990), Pischke (1991) and Goodfriend (1992) assume infinitely lived agents, with aggregation effects arising from incomplete information and different income processes.

This paper is a contribution in the latter line of research. We argue that models in which heterogeneity of incomes is modeled as an equal-for-all macroeconomic component plus an idiosyncratic term, like in Pischke and Goodfriend, though sufficient to obtain autocorrelation, excess sensitivity and smoothness, still, like most of representative-agent versions of Hall's model, yield singular vectors, i.e. consumption and income are driven by the same innovation. This is a quite unappealing implication, since, besides being in sharp contradiction with empirical evidence, it prevents us from analyzing cointegration, which is trivially verified in the singular case.

By contrast, modelling the macroeconomic component with more than one shock and different responses by individual agents, we get a model in which:
(i) autocorrelation, excess sensitivity and smoothness are obtained; (ii) the consumption-income vector is non-singular; (iii) total income and consumption are cointegrated; (iv) the implied error-correction model, unlike Campbell's, is general, i.e. lagged income and consumption appear in the consumption equation, as well as in the income equation. Thus we show that the vast literature on ECM and cointegration, spurred by Davidson et al. (1978) and Engle and Granger (1987), is not incompatible with the PIH, provided that sufficient heterogeneity of consumers' incomes is allowed.

Our empirical analysis with U.S. data shows that stationarity of saving, as obtained from an accurate estimation of total income and consumption, is not rejected over the period 1964:1 up to 1991:4. Moreover, a simple tentative model, based on two PIH consumer groups, whose labor incomes are heterogeneous, fits the data reasonably well, with the overidentifying restrictions not rejected by the comparison with a free error correction model.

Coming to the general import of this paper, we wish to put across the idea that when heterogeneity is seriously taken into account the relationship between theory and data becomes much more complex than in the usual representative agent approach. Looking back at the debate about the PIH, we find that data features like autocorrelation of consumption, excess sensitivity and smoothness have been taken as being in contrast with the theory. In fact they only contradict the micro consequences of the theory, not its aggregate implications. In the same way, a general ECM for total income and consumption, far from being in contrast with the PIH, may be derived from it assuming sufficient heterogeneity.
On the other hand, other reasons to reject the PIH remain untouched. The first moments implications of the theory are rejected by the data, at least when infinitely lived agents are assumed (Campbell 1987). Moreover, considering the period from 1947:1 to 1991:4 savings are not stationary.

Lastly, an important general consequence of introducing heterogeneity in the framework must be pointed out. Since the dynamic shape may change dramatically from micro to macro equations, the same testable implications are likely to be common to quite different theoretical explanations (Lippi 1988, Lippi and Forni 1990). In our case, a general error-correction mechanism, which has been originally motivated by a two-step optimization of a representative agent (Nickell 1985, Salmon 1982), may be re-obtained either within the model we discuss in this paper, or even in a model where a fraction of the population follows the PIH, whereas another fraction follows the 'rule of thumb' of consuming the whole current income, as recently suggested by Campbell and Mankiw (1989).

The paper is organized as follows. The first two Sections summarize the results of the representative-agent literature and specify the basic micro model. Section 3 describes the structure of the micro incomes and derives the aggregate relations. Section 4 and 5 illustrate the main results. In Section 6 empirical results are reported and commented on. Summary and conclusions are left to the last Section.

1. The micro model

We shall refer to the following formulation of Hall's model. An infinitely lived consumer maximizes the expected utility $\sum_{k=0}^{\infty} \theta^k E_t u(c_{t+k})$, subject to the sequence of budget constraints

$$\beta A_{t+1} = A_t + y_t - c_t,$$  \hspace{1cm} (1)

where $\beta = 1/(1 + r)$, $r$ is a constant interest rate, $\theta$ is an individual utility discount factor, $c_t$ is consumption, $A_t$ is assets and $y_t$ labor income, taken as exogenous. Lastly, $E_t$ stands for projection on the linear space spanned by the information set $I_t$, available to the agent at time $t$. We do not need to make specific assumptions on $I_t$ for the moment.

In the simplest case where a quadratic utility and $\theta = \beta$ are assumed, the first order conditions for this problem are

$$E_t c_{t+1} = c_t.$$

3
Equation (2) implies that optimal consumption is a martingale. The change in consumption at time $t$ is independent of agent's information at time $t-k$, $k > 1$; in particular, it is independent of its own past and the past of labor income. A significant correlation between consumption change and past income change has been referred to in the literature as 'excess sensitivity' of consumption.

Solving forward the budget constraint (1) gives

$$A_t = \frac{1}{1 - \beta F} c_t - \frac{1}{1 - \beta F} y_t,$$

(3)

where $F$ is the forward operator. Taking expectations and using $E_t c_{t+k} = c_t$, $k \geq 0$, we get

$$c_t = (1 - \beta) \left[ A_t + E_t \frac{1}{1 - \beta F} y_t \right].$$

(4)

Equation (4) states that consumption equals permanent income, defined as the flow of rental income from total wealth. Total wealth includes both assets and human wealth, which in turn is defined as the present value of the expected labor income stream.

Taking the first differences of (3), multiplying by $\beta F$ and applying expectations gives

$$s_t = -E_t \frac{\beta F}{1 - \beta F} \Delta y_t,$$

(5)

where $s_t = \beta \Delta A_{t+1}$ is saving. Equation (5) is Campbell's 'saving for a rainy day' equation: savings anticipate future income falls. An important implication of (5) is that if the change in labor income is stationary then saving is stationary. Notice that in this case, since consumption is I(1), total income $x_t = (1 - \beta) A_t + y_t = c_t + s_t$ is also I(1). Therefore, stationarity of saving may be rephrased by saying that total income and consumption are cointegrated with cointegrating vector $(1 \ -1).$\footnote{This implication holds true even if labor income is trend-stationary; in this case the order of cointegration is two.}

To conclude this Section, we derive two further relations which will prove useful in the sequel. Taking the first difference of (4), noting that $(1 - \beta) \Delta A_t = [(1 - \beta)/\beta] s_{t-1}$ and substituting from (5) we obtain

$$\Delta c_t = (1 - \beta)(E_t - E_{t-1}) \frac{1}{1 - \beta F} y_t,$$

(6)
which shows that the changes in consumption depend on consumer's revisions in expected human wealth. Rearranging terms gives, after some manipulations,

\[ \Delta c_t = (E_t - E_{t-1}) \frac{1}{1 - \beta F} \Delta y_t. \]  

(7)

2. Income expectations

To proceed further we must specify the variables contained in the information set \( I_t \) of the agent and their model. The usual assumption is that the agent only observes labor income \( y_t \) and that the latter evolves according to:

\[ \Delta y_t = a(L)e_t, \]  

(8)

where \( e_t \) is white noise, \( a(L) \) is a rational function with no poles of modulus less or equal to unity, no roots of modulus less than unity. Since we are only interested in the second moments of the processes, let us assume that \( e_t \), and therefore \( \Delta y_t \), has zero mean. We also suppose that \( a(1) \neq 0 \), so that \( y_t \) is I(1), even though many of the results below do not depend on this assumption.

Equation (8) can be used to obtain explicit expressions for consumption, saving and total income. Consumption is easily obtained by noting that (8) implies \( (E_t - E_{t-1}) \beta^k \Delta y_{t+k} = a_k \beta^k e_t \). Hence from (7) we get

\[ \Delta c_t = a(\beta)e_t. \]  

(9)

Combining (5) and (8) gives saving:

\[ s_t = \beta \frac{a(L) - a(\beta)}{\beta - L} e_t. \]  

(10)

Notice that the expression on the right is a linear combination of present and past of income innovation, even if the denominator vanishes for \( L = \beta < 1 \). In fact the numerator also vanishes for \( L = \beta \), so that the expression has no essential poles of modulus less than one.

Lastly, taking the first difference of (10) and adding (9) the change in total income \( x_t = s_t + c_t \) is found to be

\[ \Delta x_t = h(L)e_t, \]  

(11)
where
\[ h(L) = \frac{\beta a(L)(1 - L) - (1 - \beta)a(\beta)L}{\beta - L}. \]

Equations (8) and (9) show that both consumption and labor income are driven by the same white noise \( e_t \). This is a very unappealing feature of Hall's micro theory when the labor income is the only variable observed by the agent: the vectors of labor income and consumption, total income and consumption, saving and total income, are all singular. This implies that a regression of consumption on present past and future of labor income should produce no residual, which is completely at odds with any empirical evidence: for instance, employing U.S. data, the \( R^2 \) resulting from such a regression does nor exceed 0.45. Moreover, consumption should be cointegrated with both labor income and total income. In fact, I(1) variables driven by the same shock are trivially cointegrated.

Such paradoxical implications can be avoided by assuming that individuals possess a richer information than past income alone, like for instance in Quah (1990), where two distinct components are observed by a representative agent to predict future values of labor income. Different components of labor income will be introduced in the next Section. However, we will not follow Quah's way. Rather, we will show that if sufficient heterogeneity is assumed for individual labor incomes, the singular micro model considered so far aggregates into a non-singular model, in which non-trivial cointegration holds for total income and consumption but not for labor income and consumption.

3. Aggregation

In order to derive the macro model, we must specify a structure for consumers' labor incomes. We assume that individual income is driven by 'common' shocks and 'individual' shocks. The individual shocks of consumer \( i \) are orthogonal both to the common shocks and to the individual shocks of consumer \( j \). For simplicity, we assume that there are only two common shocks; the generalization to the case of \( n \) common shocks is straightforward. Moreover, we allow for different individual responses to common shocks. Formally, individual labor income is given by

\[ \Delta y_{it} = \alpha_i(L)U_t + \gamma_i(L)V_t + \omega_i(L)\chi_{it}, \]

where \( U_t, V_t \) and \( \chi_{it} \) are mutually orthogonal white noises, while \( \alpha_i(L), \beta_i(L) \) and \( \omega_i(L) \) are rational function of \( L \). To ensure that both individual and aggregate labor income are I(1) it will be convenient to assume that \( \alpha(1) \neq 0 \) for
any \( i \). This implies that at least \( U_t \) has a permanent effect on all individual incomes.

The shocks \( U_t \) and \( V_t \) are common to all agents, whereas \( \chi_{it} \) is idiosyncratic, namely \( \chi_{it} \) is orthogonal to \( \chi_{jt} \), for \( i \neq j \), at all leads and lags. In addition to representation (12) we shall need the univariate Wold representation of \( \Delta y_{it} \):

\[
\Delta y_{it} = a_i(L)e_{it}. \tag{13}
\]

The latter is obtained by solving:

\[
a_i(L)e_{it} = \alpha_i(L)U_t + \gamma_i(L)V_t + \omega(L)\chi_{it}
\]

in the unknowns \( a_i(L) \) and \( e_{it} \). This is done by: (1) Taking the spectral density of the RHS of the equation; (2) Splitting it as \( a_i(L)a_i(F)\sigma_i^2 \), where \( a_i(L) \) fulfills the usual conditions on roots and poles; (3) Setting:

\[
e_{it} = \frac{\alpha_i(L)U_t + \gamma_i(L)V_t + \omega_i(L)\chi_{it}}{a_i(L)}. \tag{14}
\]

Let us now come to the aggregation of equations (12). Denoting with \( N \) the number of agents in the economy, per-capita income is:

\[
\Delta Y_t = N^{-1} \sum_{i=1}^{N} \Delta y_{it} = \left[ N^{-1} \sum_{i=1}^{N} \alpha_i(L) \right] U_t + \left[ N^{-1} \sum_{i=1}^{N} \gamma_i(L) \right] V_t + N^{-1} \sum_{i=1}^{M} \omega_i(L)\chi_{it}.
\]

As \( N \) tends to infinity, if the coefficients of \( \alpha_i(L) \), \( \beta_i(L) \) and \( \omega_i(L) \) are bounded over the population, a simple argument based on mutual orthogonality of the shocks \( \chi_{it} \) leads to the conclusion that the idiosyncratic component vanishes in variance (see Granger (1987)). Thus, for huge numbers of agents, the idiosyncratic component, although possibly important for each single agent, becomes negligible in the aggregate, so that we may write to a good approximation:

\[
\Delta Y_t = \alpha(L)U_t + \gamma(L)V_t, \tag{15}
\]

where \( \alpha(L) = N^{-1} \sum_i \alpha_i(L) \) and \( \gamma(L) = N^{-1} \sum_i \gamma_i(L) \).

Comparison between equations (12) and (15) gives an important insight into the observed differences in autocorrelation structures of micro and micro
data. Suppose the idiosyncratic shock dominates in variance the common shocks in (12). This implies that the function $\alpha_i(L)$ is almost equal to $\omega_i(L)$. On the other hand the $\omega_i(L)$ have no influence on the aggregate $\Delta Y_t$. Therefore, if $\omega_i(L)$ implies a negative autocorrelation whereas $\alpha_i(L)$ and $\gamma_i(L)$ imply a positive autocorrelation, the latter will be a feature of aggregate income whereas the former will characterize the corresponding micro data (on this point see Section 4).

Let us now deal with consumption. Unlike in Section 2, here individual incomes are explicitly modeled by adding components which may be different across different agents. Therefore we must choose amongst different assumptions regarding the information set employed by agents. For instance, if agents observe the components of their individual income separately, consumption change is:

$$\Delta c_{it} = \alpha_i(\beta)U_{it} + \gamma_i(\beta)V_{it} + \omega_i(\beta)x_{it}.$$  

As a variant, they might observe the idiosyncratic term and the sum of the other two components, but not the latter separately. Such assumptions lead to non-singular models in which excess smoothness can find an explanation (as in Quah (1990)). Interesting models arise if we assume that agents observe their individual incomes and the aggregate income, possibly with some lag (Goodfriend (1992)).

Here we shall assume that agents observe only their individual income. Indeed, the use of aggregate variables in predicting individual future incomes is rather implausible. The value of aggregate information on labor income is typically insignificant at the micro level (Altorji and Ashenfelter 1980, Pischke 1991); hence, the best thing that consumers can do is not to acquire it even if it is available at a low cost. This is tantamount to assuming that the variance of the common terms $\alpha_i(L)U_t$ and $\gamma_i(L)V_t$ is very small in comparison with that of $\Delta y_{it}$, so that predictions based on individual information perform almost as well as predictions based both on individual and aggregate information.

Under our assumption consumption changes for agent $i$ are obtained by straightforward application of the micromodel to (13):

$$\Delta c_{it} = \alpha_i(\beta)e_{it},$$  \hspace{1cm} (9')

while changes in individual total income are

$$\Delta x_{it} = h_i(L)e_{it},$$  \hspace{1cm} (11')

where

$$h_i(L) = \frac{\beta a_i(L)(1 - L) - (1 - \beta)a_i(\beta)L}{\beta - L}.$$
Moreover, substituting (14) into (9'), summing and dividing by $N$:

\[
\Delta C_t = \delta(L)U_t + \theta(L)V_t,
\]

(16)

where

\[
\delta(L) = N^{-1} \sum_i a_i(\beta)\alpha_i(L)/a_i(L)
\]

\[
\theta(L) = N^{-1} \sum_i a_i(\beta)\gamma_i(L)/a_i(L),
\]

and $C_t$ is per-capita consumption.

In a similar way, from equation (11') we get

\[
\Delta X_t = \lambda(L)U_t + \mu(L)V_t,
\]

(17)

where $X_t$ is per-capita total income, and

\[
\lambda(L) = N^{-1} \sum_i h_i(L)\alpha_i(L)/a_i(L)
\]

\[
\mu(L) = N^{-1} \sum_i h_i(L)\gamma_i(L)/a_i(L).
\]

(18)

4. Sensitivity and smoothness

It is quite clear from equation (16) that aggregate consumption is no longer a random walk. A simple example will be sufficient to show how excess sensitivity and smoothness can arise through aggregation. Assume that $V_t = 0$ and that $\alpha_i(L) = (1 + \alpha L)$, with $0 < \alpha < 1$ for all $i$. In this case changes in aggregate income are:

\[
\Delta Y_t = (1 + \alpha L)U_t,
\]

so that $\text{cov}(\Delta Y_t, \Delta Y_{t-1}) > 0$, consistently with observed data. Moreover, assume that $\omega_i(L) = (1 - \omega L)$, with $0 < \omega < 1$ for all $i$, and that $\Delta y_{it}$ is dominated by the idiosyncratic component. This implies that $a_i(L) = (1 - a L)$, with $0 < a < 1$. Under such assumptions, which are consistent with empirical findings (Pischke (1991)), consumption changes are:

\[
\Delta C_t = (1 - a\beta)\frac{1 + \alpha L}{1 - a L} U_t.
\]

(19)
Clearly, \( \text{cov} (\Delta C_t, \Delta U_{t-k}) > 0 \) for \( k > 0 \), so that \( \text{cov} (\Delta C_t, \Delta Y_{t-k}) > 0 \), i.e. there is positive ‘excess sensitivity’.

To get an economic intuition of the result, consider that consumers’ human wealth is positively correlated with the common shock \( U_t \). Since this correlation is small, consumers ignore it in predicting their permanent income, so that, when a positive common shock occurs, they fix consumption at a slightly lower than the ‘optimal’ level. This ‘error’, while insignificant at the micro level, becomes important in the aggregate.\(^2\) But a positive common shock induces (little) positive changes in future univariate income innovations at the micro level;\(^3\) when these innovations are observed, agents correct their consumption upward. Once again, the correction is small at the individual level but induce large positive changes in aggregate consumption. This is why changes in aggregate consumption are positively correlated with the past of the common shock (and hence with the past of aggregate income change).

The example above can be used to show that consumption is correctly predicted to be less volatile than income. Clearly, consumption is smooth at the micro level, since \( \text{var} (\Delta c_{it}) = (1 - \alpha \beta)^2 \sigma^2_{\varepsilon_t} \), while \( \text{var} (\Delta y_{it}) = (1 + \alpha^2) \sigma^2_{\varepsilon_t} \). The intuition is simple. As we have seen, consumption is more volatile than income if and only if the change in income is positively correlated with its own past. But at the micro level the correlation is negative; therefore the PIH predicts individual consumption to be smooth, even if consumers make univariate forecasts. This property is preserved when aggregate consumption is considered. To see this, consider equation (19): \( \Delta C_t \) is obtained from \( \Delta Y_t \) by applying the filter \( (1 - \alpha \beta)/(1 - aL) \). Therefore the spectral density of \( \Delta C_t \) is equal to that of \( \Delta Y_t \) multiplied by \( (1 - \alpha \beta)^2 /[1 - ae^{-i\lambda t}]^2 \). This expression is less than one at all frequencies other than \( \lambda = 0 \) and \( 2\pi \) if \( \beta = 1 \); therefore the variance of \( \Delta C_t \) is less than the variance of \( \Delta Y_t \).

5. Singularity and cointegration

Coupling equations (16) and (15) we get a bivariate MA representation of consumption and labor income changes, which result from aggregation of the micro equations (9') and (13) respectively. The matrix of this representation is

\[
N^{-1} \left( \frac{\sum a_i(\beta) a_i(L)}{\sum a_i(L)} \bigg\{ \sum a_i(\beta) \gamma_i(L) / a_i(L) \bigg\} \bigg\} \right). \tag{20}
\]

\(^2\) In our example, if \( U_t \) were observed, we would have \( \Delta C_t = (1 + \alpha \beta) U_t \), while the impact multiplier of (19) is \( (1 - \alpha \beta) \), which is considerably smaller if \( \alpha \) and \( a \) are not close to zero.

\(^3\) In our example, \( c_{it} = [(1 + aL)/(1 - aL)] U_t + [(1 - \omega L)/(1 - aL)] \chi_{it} \) (see (14)).
A simple inspection of matrix (20) shows that, although the micro model is singular, the aggregate model is non-singular if the common components of individual incomes are sufficiently heterogeneous. More precisely, given equation (12) for individual incomes non-singularity of (20) requires that the ratio $\alpha_i(L)/\gamma_i(L)$ is not uniform across the population. This is equivalent to requiring that the shocks $U_i$ and $V_i$ are not redundant in equation (12): in fact, if
\[
\frac{\alpha_i(L)}{\gamma_i(L)} = \tau(L) \text{ for all } i
\] (21)
then
\[
\alpha_i(L)U_i + \gamma_i(L)V_i = \gamma_i(L)[\tau(L)U_t + V_i] = \gamma_i(L)\rho(L)W_t,
\]
where $W_t$ is independent of $i$, i.e. one single shock is sufficient to generate the common component of all the agents in the population.

Setting $\alpha_i(L) = \alpha(L)$ and $\gamma_i(L) = \gamma(L)$ for all $i$ in (12), which is a particular case of (21), one gets the standard model employed in micro-data literature, as well as in the papers by Pischke (1991) and Goodfriend (1992). In this case the heterogeneity of individual incomes is completely confined in the idiosyncratic component.

Naturally, as already observed in Section 2, singularity of (20) is strongly rejected by empirical evidence. Therefore, if model (12) is assumed for individual data, then under PIH restriction (21) is rejected as well.

Summing up, non-singularity of the consumption-income vector is a crucial feature of data that many PIH models so far proposed fail to reproduce. Non-singularity is achieved in representative-agent models in which more than one shock is observed by the agent (e.g. Quah 1988). Our model contains different macroeconomic shocks, but agents do not observe them separately. Therefore our micromodel is singular. Non-singularity in our macromodel stems from aggregation of heterogeneous agents. More precisely, non-singularity is achieved if the common component is driven by at least two shocks, while restriction (21) does not occur.  \footnote{Indeed, exclusion of (21), although necessary, is not sufficient for non-singularity of (20); however, cases in which (20) is singular but (21) does not hold do not seem to have any economic interest.}

Let us come now to cointegration. We can distinguish two cases. Firstly, $\alpha_i(1) \neq 0$, $\gamma_i(1) = 0$, for all $i$, so that only $U_t$ has a permanent effect on incomes. In this case, the determinant of (20) vanishes for $L = 1$, implying that aggregate consumption and labor income are cointegrated. This is because a unique permanent shock to income and consumption implies that there is only one common trend in the macro variables. Notice however that a different result

4
emerges when consumers can distinguish permanent from transitory shocks. In this case both $U_t$ and $V_t$ are permanent for the consumption process, even if $V_t$ is transitory for the income process (see Quah 1990); so that there are two distinct common trends in aggregate labor income and consumption.

Secondly, both $U_t$ and $V_t$ have a permanent effect. In this case aggregate consumption and labor income are not cointegrated, unless the parameters of individual incomes take on very particular values.\(^5\)

Unlike consumption and labor income, consumption and total income are cointegrated independently of the number of common permanent shocks. Indeed, equations (16) and (17) give the bivariate MA representation:

$$
\begin{pmatrix}
\Delta C_t \\
\Delta X_t
\end{pmatrix} =
\begin{pmatrix}
\delta(L) & \theta(L) \\
\lambda(L) & \mu(L)
\end{pmatrix}
\begin{pmatrix}
U_t \\
V_t
\end{pmatrix}.
$$

Equations (18) and (11') imply that

$$
\delta(1) = \lambda(1) = N^{-1} \sum_i a_i(\beta) a_i(1)/a_i(1) \\
\theta(1) = \mu(1) = N^{-1} \sum_i a_i(\beta) \gamma_i(1)/a_i(1).
$$

Hence, $\Delta X_t$ and $\Delta C_t$ are cointegrated with cointegrating vector $(1 \ -1)$.\(^6\)

Notice that cointegration of aggregate total income ad consumption could have been obtained as a consequence of a general result. Since the cointegrating vector is the same for all consumers, cointegration is preserved by aggregation (see Lippi 1988, Gonzalo 1989). Put another way, aggregation preserves stationarity of saving, along with difference-stationarity of consumption and total income, so that cointegration is maintained at the macro level.

---

\(^5\) Notice that the determinant of (20) vanishes for $L = \beta < 1$, so that representation (15)-(16) is not the Wold representation. This result is similar to that of Quah (1990). In the representative agent model of Quah, nonfundamentalness arises from the fact that, unlike the econometrician, the representative agent can distinguish permanent from transitory shocks. Here consumers cannot distinguish permanent from transitory shocks. However, they still know something that the macro econometrician does not observe, namely their individual incomes.

\(^6\) Notice that the determinant of the matrix in (22) does not vanish for $L = \beta$. Indeed, the determinant is $\beta(1-L)/(\beta-L)$ times the determinant of (20), so that it has exactly the same roots except for $\beta$, which is replaced by 1. It follows that if $U_t$ is permanent and $V_t$ is transitory, then the determinant of (22) has two unitary roots instead of only one, so that a VAR model linking consumption and total income does not exist either in differences or in levels.
6. Empirical results

For our empirical analysis we use the U.S.A. national income and product accounts (NIPA) quarterly data from 1947:1 to 1991:4. All the data are seasonally adjusted at quarterly rates, taken in per-capita terms and expressed in thousands of 1987 U.S. dollars. We analyze total disposable income ($X$) and total consumption expenditures ($C$). Following Blinder and Deaton (1985), we make some adjustments to NIPA definitions. Firstly, the 1975 tax rebate is removed from the income data; the numbers for this correction are taken from Blinder (1981). Secondly, personal 'nontax' payments to state and local governments are added to both consumption and income. Lastly, interest paid by consumers to business is subtracted from disposable income. Savings ($S$) are defined as the difference between the resulting series $X$ and $C$.

Total consumption expenditures might not be the most appropriate series to use in our context. Indeed, for durable goods it is the service flows rather than purchases that should be included in the aggregate. We therefore use both consumer spending $C$ and an alternative measure of total consumption ($CC$) that includes an estimate of the imputed rent on the stock of consumer durables. The stock of durables is calculated by accumulating the spending flow, starting with the NIPA net stock of consumer durables for the end of 1946, and assuming a depreciation rate of 5 per cent per quarter. To calculate the imputed rent, we assumed a user cost of 6 per cent per quarter. Consumption on clothing and shoes is treated in a similar way, but assuming a depreciation rate of 20 per cent, a zero initial value and a user cost of 21 per cent per quarter. Consumption $CC$ is the imputed rent on durables, clothing and shoes plus expenditures on services and nondurables other than clothing and shoes. In relation with $CC$ we use an estimate of disposable income, $XX$, which is the sum of $X$ and net interests on durables, clothing and shoes. Saving $SS$ is defined as the difference between $XX$ and $CC$.

As shown in Table 1, there is no evidence against the null hypothesis of unit root in the four series $X$, $C$, $XX$ and $CC$ when compared with the trend-stationary alternative. Hence the data are consistent with the assumption $\theta = r$, i.e. that the rate of time preference is equal to the interest rate.

Figure 1 shows the two measures of saving $S$ and $SS$. Neither of these seems stationary. As regards $S$, the unit root hypothesis is rejected by the Dickey-Fuller test at the 5 per cent level but cannot be rejected by the augmented Dickey-Fuller test with one lag or more. $SS$ is I(1) according to both the DF and the ADF test with any number of lags (see Table 1). Hence the aggregate

---

7 Empirical results do not change substantially when the 1975 tax rebate is not removed.
implications of the theory are not supported by the data over the whole period 1947:1 1991:4.

However, both the series in Figure 1 look much more stationary if we concentrate attention on the subperiod starting in the mid-sixties. Indeed, within the period 1964:1 1991:4 the hypothesis of a unit root in $S$ is rejected at the 5 per cent level by the DF and the ADF(1) tests, while $SS$ is stationary according to all the tests (see Table 2).
Table 1. Unit root tests 1947:1 1991:4

<table>
<thead>
<tr>
<th>variable</th>
<th>DF</th>
<th>ADF(1)</th>
<th>ADF(2)</th>
<th>ADF(3)</th>
<th>95% critical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>with trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>-2.441</td>
<td>-2.296</td>
<td>-2.373</td>
<td>-2.294</td>
<td>-3.436</td>
</tr>
<tr>
<td>CC</td>
<td>-1.771</td>
<td>-2.344</td>
<td>-2.294</td>
<td>-2.432</td>
<td>-3.436</td>
</tr>
<tr>
<td>without trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>-3.015</td>
<td>-2.817</td>
<td>-2.422</td>
<td>-2.594</td>
<td>-2.878</td>
</tr>
<tr>
<td>SS</td>
<td>-2.363</td>
<td>-2.293</td>
<td>-1.966</td>
<td>-2.061</td>
<td>-2.878</td>
</tr>
</tbody>
</table>

Table 2. Unit root tests 1964:1 1991:4

<table>
<thead>
<tr>
<th>variable</th>
<th>DF</th>
<th>ADF(1)</th>
<th>ADF(2)</th>
<th>ADF(3)</th>
<th>95% critical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>with trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>-1.315</td>
<td>-1.877</td>
<td>-2.362</td>
<td>-3.231</td>
<td>-3.450</td>
</tr>
<tr>
<td>CC</td>
<td>-1.278</td>
<td>-2.206</td>
<td>-2.493</td>
<td>-3.505</td>
<td>-3.450</td>
</tr>
<tr>
<td>XX</td>
<td>-2.329</td>
<td>-2.742</td>
<td>-2.450</td>
<td>-2.914</td>
<td>-3.450</td>
</tr>
<tr>
<td>without trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td>-3.312</td>
<td>-3.348</td>
<td>-2.935</td>
<td>-2.977</td>
<td>-2.887</td>
</tr>
</tbody>
</table>

Having obtained cointegration for total income and consumption, as predicted, one could be tempted to go further and estimate model (22) on the period 1964:1 1991:4. However, application of (22) to empirical data calls for a specification of the joint distribution of the microparameters amongst agents. Unfortunately, existing empirical studies on micro data are useless for our purpose, due to their poor specification of heterogeneity and dynamics. For these reasons we shall limit ourselves to a highly artificial exercise in which we simply assume two groups of agents. Precisely, let us assume that there are two types of workers (for instance self employed and employees), namely type 1 and type 2. The income of type 1 workers responds only to $U_t$, whereas the income of
type 2 workers responds only to $V_t$:

$$\Delta y_{1t} = \frac{1}{1-\alpha L} U_t + \omega_1(L) \chi_{1t}$$

$$\Delta y_{2t} = \frac{1}{1-\gamma L} V_t + \omega_2(L) \chi_{2t}.$$ 

The idiosyncratic components are assumed to be such that the univariate Wold representations are:

$$\Delta y_{1t} = (1-\alpha L)\epsilon_{1t}$$

$$\Delta y_{2t} = (1-\gamma L)\epsilon_{2t},$$

where $\alpha$, $a$, $\gamma$ and $c$ are, according to empirical findings, all positive and less than unity.

The resulting joint model of consumption and total income is:

$$
\begin{pmatrix}
\Delta C_t \\
\Delta X_t
\end{pmatrix} = 
\begin{pmatrix}
\frac{1-a\beta}{(1-a\beta)(1-a\alpha)} & \frac{1-c\beta}{(1-c\beta)(1-\gamma L)} \\
\frac{1-a\beta L}{(1-a\beta)(1-a\alpha)} & \frac{1-c\beta L}{(1-c\beta)(1-\gamma L)}
\end{pmatrix}
\begin{pmatrix}
P_t \\
Q_t
\end{pmatrix},
$$

(23)

where $P_t = (M/N)U_t$, $Q_t = [(N - M)/N]V_t$, $M$ being the number of type 1 agents. The implied ECM is

$$
\begin{pmatrix}
1 - \delta_1 L - \delta_2 L^2 \\
-\phi_1 L - \phi_2 L^2
\end{pmatrix}
\begin{pmatrix}
\Delta C_t \\
\Delta X_t
\end{pmatrix} = 
\begin{pmatrix}
\mu \\
\nu
\end{pmatrix} S_{t-1} + 
\begin{pmatrix}
W_t \\
Z_t
\end{pmatrix},
$$

(24),

where

$$\delta_1 = \frac{(1-a\beta)(a\alpha \beta + ac\beta + a\alpha - acc\alpha \beta)}{\beta(c-a)} - \frac{(1-c\beta)(a\gamma \beta + ac\beta + c\gamma - ac\gamma \beta)}{\beta(c-a)}$$

$$\delta_2 = \frac{(1-c\beta)ac\gamma - (1-a\beta)acc\alpha}{c-a}$$

$$\theta = \frac{(1-a\beta)(1-c\beta)(a\alpha - c\gamma)}{\beta(c-a)}$$

$$\phi_1 = -\psi + \frac{ac - a\gamma}{c-a}$$

$$\phi_2 = \frac{ac(\gamma - \alpha)}{c-a}$$
\[
\psi = \frac{(1 - a\beta) c\gamma - (1 - c\beta)a\alpha}{\beta(c - a)}
\]
\[
\mu = \frac{(1 - a\beta)(1 - c\beta)[(1 - a)(1 - \alpha) - (1 - c)(1 - \gamma)]}{\beta(c - a)}
\]
\[
\nu = \frac{(1 - c\beta)(1 - a)(1 - \alpha) - (1 - a\beta)(1 - c)(1 - \gamma)}{\beta(c - a)}
\]

and
\[
\begin{pmatrix}
W_t \\
Z_t
\end{pmatrix} = \begin{pmatrix}
(1 - a\beta)P_t + (1 - c\beta)Q_t \\
Pt + Qt
\end{pmatrix}.
\]

We estimated model (23), which is equivalent to estimating (24) under the restrictions implied by the functional relationships listed above. Moreover, we estimated model (24) without restrictions and model (24) with the restrictions implied by the permanent-income hypothesis in its representative-agent version, i.e. \(\delta_1 = \delta_2 = \theta = \mu = 0\). The parameter \(\beta\) has been arbitrarily fixed to .99, which implies a quarterly real interest rate near to 1 per cent. Data are measured in deviations from their means. The results are reported in Table 3.

The first column shows the estimates for the parameters of model (23) obtained when using \(C\) and \(X\), along with the implied parameters of model (24). The second column shows the estimates for the free model (24). The third column shows the estimates for the Hall-Campbell’s representative-agent ECM. Columns fourth, fifth and sixth replicate the first three with reference to \(CC\) and \(XX\).

Excess sensitivity is clearly confirmed: the restrictions implied by the representative-agent version of the PIH model are rejected by both data sets. Rejection is very strong when using the second (and more interesting) data set: the LR test gives a probability value less than .001.

Model (23) fits the data remarkably better, especially with \(CC\) and \(XX\). The only shortcoming is the significantly negative estimate of the parameter \(a\) both using \(C\), \(X\), and using \(CC\), \(XX\), so that the negative autocorrelation for micro incomes, observed in empirical data, holds for only one of the groups (notice however that with \(CC\), \(XX\) parameter \(a\) is considerably smaller in modulus). Model (23) cannot be rejected against the free model by the LR test at the 5 per cent level with the second data set, at the 2.5 per cent level with the first.

---

8 The programs for maximum likelihood estimations and calculation of the information matrices have been written by the authors in MATLAB.
Table 3. Estimates of (23), free ECM and Campbell’s Representative-Agent ECM*

<table>
<thead>
<tr>
<th></th>
<th>with C, X</th>
<th>with CC, XX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>model (23)</td>
<td>free ECM</td>
</tr>
<tr>
<td>α</td>
<td>.62 (.09)</td>
<td>.65 (.09)</td>
</tr>
<tr>
<td>a</td>
<td>-.47 (.09)</td>
<td>-.28 (.11)</td>
</tr>
<tr>
<td>γ</td>
<td>-.08 (.10)</td>
<td>.00 (.10)</td>
</tr>
<tr>
<td>c</td>
<td>.93 (.04)</td>
<td>.95 (.03)</td>
</tr>
<tr>
<td>$10^2 \sigma^2_P$</td>
<td>.16 (.03)</td>
<td>.06 (.01)</td>
</tr>
<tr>
<td>$10^2 \sigma^2_Q$</td>
<td>.53 (.07)</td>
<td>.55 (.08)</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>.15 .18 (.10)</td>
<td>0 .36 .41 (.10)</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>.28 .14 (.10)</td>
<td>0 .18 .10 (.10)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>.02 .08 (.08)</td>
<td>0 .01 .01 (.04)</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>.45 .42 (.13)</td>
<td>.33 (.13)</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>.22 .02 (.13)</td>
<td>-.05 (.12)</td>
</tr>
<tr>
<td>$\psi$</td>
<td>-.07 .04 (.10)</td>
<td>.00 (.09)</td>
</tr>
<tr>
<td>$\mu$</td>
<td>.04 .02 (.04)</td>
<td>0 .03 .01 (.02)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>-.05 -.15 (.06)</td>
<td>-.16 (.05)</td>
</tr>
<tr>
<td>$10^2 \sigma^2_Y$</td>
<td>.33 .31 (.04)</td>
<td>.35 (.05)</td>
</tr>
<tr>
<td>$10^2 \sigma^2_{\nu}$</td>
<td>.68 .53 (.07)</td>
<td>.53 .61</td>
</tr>
<tr>
<td>$10^2 \sigma_{Y \nu}$</td>
<td>.27 .16 (.04)</td>
<td>.17 (.05)</td>
</tr>
<tr>
<td>LLikelihood</td>
<td>765.6 771.6</td>
<td>765.6 833.6</td>
</tr>
<tr>
<td>LR Statistic</td>
<td>12.2 12.0</td>
<td>10.8 28.7</td>
</tr>
<tr>
<td># Restr.</td>
<td>5 0</td>
<td>4 5</td>
</tr>
<tr>
<td>$\chi^2_{.05}$</td>
<td>11.1 9.5</td>
<td>11.1 9.5</td>
</tr>
</tbody>
</table>

* Standard errors in brackets.

Summary and conclusions

Autocorrelation of consumption under the PIH and different forms of limited information has been shown in Lippi (1990), Pischke (1991), Goodfriend (1992). The latter two authors also motivate excess sensitivity and smoothness. However, due to their standard common plus idiosyncratic model of heterogeneity, they get a singular macro model. In this paper we have shown that allowing for more complex heterogeneity one gets a non-singular macro model. Moreover, total income and consumption are predicted to be cointegrated with cointegration vector $(1 - 1)$ and to follow a general error correction mechanism. Thus, once sufficient heterogeneity is allowed, PIH and ECM models of consumption appear to be no longer in contrast.

U.S. data do not reject cointegration for the period from 1964:1 to 1991:4. A highly artificial specification of our heterogeneity model, containing only two
types of agents, has been estimated for this period and not rejected when compared to its free parameters version by the likelihood ratio test. Although quite naïve, our empirical exercise can be considered as a useful prototype for the introduction of heterogeneity in macro analysis.

References


rics, Fifth World Congress, Volume II. Cambridge: Cambridge University Press, 121-46.


27. Giovanna Procacci [1988] “The State and Social Control in Italy During the First World War”, pp. 18
43. Giovanna Procacci [1989] “State coercion and worker solidarity in Italy (1915-1918): the moral and political content of social unrest”, pp. 41
44. Carlo Alberto Magni [1989] “Reputazione e credibilità di una minaccia in un gioco bargaining”, pp. 56
55. Paolo Silvestri [1990] “Sull’autonomia finanziaria dell’università”, pp. 11
64. Mario Forni [1990] “Incertezza, informazione e mercati assicurativi: una rassegna”, pp. 37
74. Enrico Giovannetti [1990] “Illusioni ottiche negli andamenti delle Grandezze distributive: la scala mobile e l’’appiattimento’ delle retribuzioni in una ricerca”, pp. 120
78. Antonietta Bassetti e Costanza Torricelli [1990] “Una riqualificazione dell’approccio bargaining alla selezione di portafoglio”, pp. 4
77. Antonietta Bassetti e Costanza Torricelli [1990] “Il portafoglio ottimo come soluzione di un gioco bargaining”, pp. 15
101. Barbara Pistersi [1994] “Using a VECM to characterise the relative impotence of permanent and transitority components”, pp. 28
102. Gian Paolo Caselli and Gabriele Pastrello [1994] “Polish recovery form the slump to an old dilemma”, pp. 20
105. Giuseppe Marotta [1994] “Credit view and trade credit: evidence from Italy”, pp. 20


