A nonparametric analysis of welfare and the economic shocks

Nicoletta Rosati
CEMAPRE and ISEG - Lisbon, Portugal*

July 2, 2007

Abstract

The behaviour of the permanent and transitory economic shocks for different levels of household's welfare is studied using both consumption and income measures. After testing for heteroskedasticity of the economic shocks, we use local polynomial regression models to estimate the variance of the shocks conditional on welfare level. Italian data covering the period 1980-2004 show evidence of heteroskedasticity of both the transitory and the permanent economic shocks, with the poor experiencing higher variances. The permanent shocks seem to have a more uniform effect at all welfare levels.

Keywords: Heteroskedasticity; income and consumption welfare measures; local polynomial regression; permanent and transitory shocks.

*Corresponding address: Department of Mathematics, ISEG, Technical University of Lisbon. Rua do Quelhas, 6. 1200-781 Lisbon. E-mail: nicoletta@iseg.utl.pt

1 Introduction

The welfare of a family is usually defined according to the level of income or consumption; in particular, the poverty state is normally identified through comparison to a threshold (poverty line). Intuitively, for those in the poverty state there is more than a low-income or low-consumption condition, as they also suffer from the consequences of such low levels, especially in the most difficult times for the economy. The aim of this study is to investigate the impact of the economic shocks along the distribution of welfare, in particular trying to understand whether, in addition to a disadvantaged starting point, the poor are hit harder than the rest of the population.

The study of welfare needs the careful distinction of its temporary and permanent aspects. For example, in a policy decision process we might be dealing with temporary fluctuations of income, which can hide the real, long-run, level of welfare of a household, and therefore mis-identify the target population. Given the objective of this study, we wish to be able to distinguish the permanent economic shocks from their transitory counterpart, and study separately their effects along the welfare distribution.

Some studies have considered (unobservable) permanent income as a measure of permanent welfare; for instance, Abul-Naga and Burgess (1997) develop a model to predict household permanent income making use of both income and expenditure plus an extra set of causal variables for permanent income, such as family size, housing asset ownership, educational and occupational status of the household. Other approaches instead, like the studies by Deaton and Paxson (1994) and Blundell and Preston (1998), combine a specification of the permanent income process with an intertemporal model of optimization over the life cycle. In particular, the last paper highlights the value of the joint use of income and consumption measures in the identification of the temporary and permanent components of *inequality*. Following this line, the present study investigates further the joint role of income and consumption in the identification of temporary and permanent economic shocks.

An assumption that is fairly common in the literature is that the economic shocks are homoskedastic, at least within certain homogeneous groups. This is usually a quite mild assumption, since heteroskedasticity-robust inference can be done for most standard econometric techniques. However, a non-constant variance of the economic shocks is not only a statistical issue but also an interesting aspect of the economic study of welfare. As noticed earlier, it is likely that those in the lower percentiles of the welfare distribution experience a stronger impact of the economic shocks. Therefore, in the present study, rather than assuming homoskedasticity, the behaviour of the variances of the economic shocks is investigated, to see whether they vary along the population.

The paper is organised as follows. Section 2 presents an intertemporal model of consumer's behaviour, and identifies the economic shocks. Section 3 develops a formal model to test for the presence of heteroskedasticity. Section 4 presents a framework for the estimation of the variance function of the shocks, based on local polynomial regression. Section 5 presents an application to Italian data, and Section 6 concludes.

2 An economic model for income and consumption

This section presents an economic model of consumer's behaviour based on the joint use of income and consumption, following the lines of Blundell and Preston (1998). In order to catch any generation effect, the population is split into birth cohorts. Suppose income can be decomposed into a permanent component y^p plus a transitory shock u, namely:

$$y_{it} = y^p_{it} + u_{it} \tag{1}$$

for individual i in cohort j at time t, where y^p_{it} and u_{it} are uncorrelated. The u_t's have zero mean and are independent over time. Assume that:

- y^p_{it} evolves according to: $\Delta y^p_{it} = v_{it}$, where v_{it} is a permanent shock uncorrelated with y^p_{it-1} and orthogonal to u_{it}. The v_t's have zero mean and are independent over time;
- $\text{Cov}_j(u_t, y_{t-l}) = \text{Cov}_j(v_t, y_{t-l}) = 0 \quad \forall l \geq 1$

The stochastic component of permanent income, v_{it}, includes all the economic shocks that can be considered permanent, such as an increase in the salary, or an increase in the cost of housing, and so on. It does not include returns on assets, since their riskiness has a more transitory than permanent nature. That would be part of the transitory error u_{it} in the income process.

All-together:

$$\Delta y_{it} = \Delta u_{it} + v_{it}$$

i.e. the changes in income are affected by both permanent and transitory shocks.

Add now to the above income process a life-cycle model for consumption. Assuming quadratic preferences, with the discount rate r equal to the fixed real interest rate, we obtain the random walk property of consumption as in Hall (1978); for small r and large $T - t$ the following approximate relation holds:

$$\Delta c_{it} = v_{it}$$

where T is the year of retirement. This relation implies that the consumption decisions are mainly driven by the permanent shocks if we are sufficiently far from retirement.

The model can be summarised with the following three equations:

$$y_{it} = y^p_{it} + u_{it} \tag{2}$$

$$y^p_{it} = y^p_{i\ t-1} + v_{it} \quad (3)$$

$$c_{it} = c_{i\ t-1} + v_{it} \quad (4)$$

In this economic model the transitory and permanent shocks are assumed to be uncorrelated with permanent income or consumption, but their variance is left unspecified. As noticed earlier, it is likely that the shocks have different effects along the distribution of permanent income or consumption. In particular, the shocks are expected to be more volatile in the bottom and top of the distributions. In fact, the poorer households have - for instance - more difficult access to financial services, or do not have stable job positions, facts that might increase the effects of the fluctuations of the economy. On the other hand, the richer families have more choice about the use of their income, and may show different attitudes towards savings; this in turn determines different reactions to economic shocks, and hence more variability. Therefore we shall assume that the variances of the shocks are not constant for a given time and a given cohort.

2.1 Variance of the economic shocks

How can the variance of the shocks be studied, distinguishing between permanent and transitory? The basic idea is to start from equations (2) and (4) to obtain the two errors, and then use them to study the behaviour of their variance. For this purpose, the functions of interest are the two conditional variances $\mathrm{Var}(u_{it}|y^p_{it})$ and $\mathrm{Var}(v_{it}|c_{it-1})$. For this last case (permanent shocks), v_{it} can be simply obtained as $v_{it} = c_{it} - c_{it-1}$, and then its conditional variance can be studied.

We know from the Permanent Income Hypothesis that $\Delta y^p_t = \Delta c_t = v_{it}$, i.e. that y^p_t and c_t are parallel time processes. Moreover, consumption over the life-time $(\sum_t c_t)$ should approximately equal total permanent income over the life-cycle $(\sum_t y^p_t)$, unless we are going to leave a large debt or inheritance to our heirs. This, together with the PIH, implies that $y_t^p \approx c_t$, therefore for the transitory shocks we can look at the simplest case where $y_t^p = c_t$, and obtain $u_t = y_t - c_t$, which corresponds to savings. The regression of the squared savings on consumption, therefore, gives an estimate of the conditional variance of the transitory shocks, given consumption, $\mathrm{Var}(u_{it}|c_{it})$.

In both cases, what is studied is the effect of the economic shocks at different points of the consumption distribution. This implies that here consumption plays the role of a welfare measure.

3 Testing for heteroskedasticity of the shocks

To study the conditional variances $\mathrm{Var}(u_{it}|c_{it})$ and $\mathrm{Var}(v_{it}|c_{it-1})$ consider the following additive models[1]:

$$u_{it}^2 = m(c_{it}) + \varepsilon_{it} \tag{5}$$

$$v_{it}^2 = g(c_{it-1}) + \eta_{it} \tag{6}$$

where the regression functions m and g are assumed to be general functions of c_t and c_{t-1}. To illustrate the testing procedures, we focus now on equation (5) for the transitory shocks, but analogous reasoning can be followed for the case of permanent shocks. Since one fixed time period at a time is considered, the time index is dropped to simplify notation.

For a fixed time period t, the error ε is uncorrelated with c, and $E(\varepsilon) = 0$. Let us start from the standard case where $\mathrm{Var}(\varepsilon) = \sigma_\varepsilon^2$ is a constant; this point will be relaxed later. An F-ratio type test can be derived to test the null hypothesis of no effect, i.e.

$$H_0 \; : \; E(u_i^2) = \theta$$

$$H_1 \; : \; E(u_i^2) = m(c_i)$$

[1]See, for instance, Hastie and Tibshirani (1990).

Let $RSS_0 = \sum_{i=1}^{n}\{u_i^2 - k\}^2$ be the residual sum of squares under the null, where k is the sample average of the u_i^2; let also $RSS_1 = \sum_{i=1}^{n}\{u_i^2 - \hat{m}(c_i)\}^2$ be the residual sum of squares under the alternative. Then we can define the following F-statistic

$$F = \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} \tag{7}$$

where $df_0 = \text{tr}(I - L)$ and $df_1 = \text{tr}(I - S)$ denote the degrees of freedom for error under each hypothesis; here L is an $n \times n$ matrix filled with the value $1/n$, and S is the matrix such that $\hat{m} = Su_i^2$. S is analogous to the projection matrix $H = X(X'X)^{-1}X'$ in a standard linear model of the form $y = X\beta + \nu$. Due to the ratio form, the test statistic is independent of the unknown error variance σ_ε^2.

The simplest way to perform such test is to use the Ordinary Least Squares (OLS) procedure, wich allows us to test H_0 using heteroskedasticity-robust standard errors. Given the structure of the economic problem, we can reasonably expect an U-shaped $m(.)$ function, therefore the OLS procedure might happen to fit an horizontal line through the "U" and not reject constancy of the $m(.)$ function, when in reality it is non-constant. To avoid this problem, a quadratic term should be added in the OLS regression, so that under the alternative hypothesis the regression function would be of the form $m(c_i) = \beta_0 + \beta_1 c_i + \beta_2 c_i^2$. The F test of zero slopes associated to this OLS regression is equivalent to (7) , and can be performed using heteroskedasticity-robust standard errors.

A more flexible alternative to OLS is local polynomial regression[2]. If $m(c_i)$ is specified locally as a polynomial of order p in c, local polynomial regression procedures can be used to perform the test. In this case, the test statistic (7) turns out to be a pseudo-likelihood ratio, and therefore shall be referred to as the PLRT (pseudo-likelihood ratio test). The degrees of freedom are approximate, and the distribution of the test statistic can be obtained using results on quadratic forms in Normal variables, as illustrated

[2]See, for instance, Fan and Gijbels (1996).

in Azzalini and Bowman (1997). Standard procedures within this framework assume that $\text{Var}(\varepsilon) = \sigma_\varepsilon^2$ is constant. This would be equivalent to testing for $m(c_{it})$ constant, given $\text{Var}(\varepsilon) = \sigma_\varepsilon^2$. In this case, the stronger hypothesis of u_{it} independent of c_{it} can be tested, which implies both $m(c_{it})$ constant and $\text{Var}(\varepsilon) = \sigma_\varepsilon^2$. Under the null, the standard model can be used to test directly for constancy of $m(c_{it})$.

4 Estimation of the conditional variance

Once verified that the $m(.)$ function is not constant, estimation can be performed. Local polynomial regression models constitute a semi-parametric framework that is embedded in the classical weighted least squares problem, providing a full set of inferential tools, and therefore is an ideal setting for our analysis. This framework assumes that the regression functions m and g are (locally) polynomials of order p in c, and estimates them using a smooth function. As before, let us focus on the $m(.)$ function for the transitory shocks specified in (5).

Remember that it was assumed that for a fixed time period t, the error ε is uncorrelated with c, and $E(\varepsilon) = 0$. Consider first the case $\text{Var}(\varepsilon) = \sigma_\varepsilon^2$ constant. The variance function for the transitory shocks can be modelled as locally quadratic or cubic, or even using a higher-order polynomial; however, the local linear approach is sufficiently flexible for most objectives, so first consider the case $p = 1$. For a given c, solving the least squares problem

$$\min_{\alpha_c,\beta_c} \sum_{i=1}^{n} \{u_i^2 - \alpha_c - \beta_c(c_i - c)\}^2 w(c_i - c; h) \tag{8}$$

and taking as the estimate at c the value of $\hat{\alpha}_c$ gives a local linear regression. The subscript c in the parameters is to stress the local nature of the minimisation problem, i.e. the dependence of the parameters on c. The smoother $w(.)$ is here chosen to be the standard normal density $\phi(\frac{c_i - c}{h})$. In this case h is the standard deviation of the normal kernel, i.e. we use observations approximately within a window of width $4h$; this implies that it is enough to consider $0 < h < \frac{\text{range(c)}}{2}$; in fact, choosing an $h > \frac{\text{range(c)}}{2}$ widens the window only to include areas with no observations.

For $p > 1$ the problem can be defined in a similar fashion. For instance, in the quadratic case the minimisation problem would be

$$\min_{\alpha_c,\beta_c,\gamma_c} \sum_{i=1}^{n} \{u_i^2 - \alpha_c - \beta_c(c_i - c) - \gamma_c(c_i - c)^2\}^2 w(c_i - c; h) \tag{9}$$

and so on.

This approach can be viewed as a generalisation of the usual linear regression model. As the smoothing parameter h becomes very large, the curve estimate approaches the fitted least squares regression line. This happens when all observations have the same weight, i.e. when $\phi(\min|c_i - c|) \approx \phi(\max|c_i - c|)$. Given that $\min|c_i - c| \geq 0$ and that $\max|c_i - c| \leq \text{range}(c)$, then the minimum h that grants equivalence between local linear regression and ordinary least squares is such that $\phi(0) \approx \phi(\frac{\text{range(c)}}{h})$, which gives $h \approx 40\ \text{range}(c)$. In fact, with this value $\phi(\frac{\text{range(c)}}{h}) = \phi(0.025) = 0.3988$, while $\phi(0) = 0.3989$. It follows that, for instance, the OLS-based test and the PLRT obtained setting $h = 40\ \text{range}(c)$ coincide.

The estimate $\hat{m}(c)$ has an explicit formula of the type $\hat{m}(c) = \sum \delta_i u_i^2$, where the weights δ_i are function of the kernel weights $w(c; h)$. We can therefore express the mean and variance as follows:

$$\begin{aligned} E\{\hat{m}(c)\} &= \sum \delta_i\, m(c_i) \\ \text{Var}\{\hat{m}(c)\} &= \sigma_\varepsilon^2 \sum \delta_i^2 \end{aligned}$$

Asymptotic expressions reported in Ruppert and Wand (1994) show that the bias increases with large h, while the variance increases with small h. An optimal choice of the smoothing parameter h can be obtained via cross-validation. In this case, the optimal value is that which minimises the function

$$CV(h) = \frac{1}{n}\sum_{i=1}^{n}\{u_i^2 - \hat{m}_{-i}(c_i)\}^2 \qquad (10)$$

where $\hat{m}_{-i}(c_i)$ denotes the estimate for m based on all observation excluding u_i.

Mild assumptions allow a Normal approximation for $\hat{m}(c)$ using the Central Limit Theorem, even if ε is not Normal, namely

$$\frac{\hat{m}(c) - m(c) - b(c)}{\sqrt{\hat{v}(c)}} \sim N(0, 1) \qquad (11)$$

where $\hat{v}(c)$ is the estimated variance of $\hat{m}(c)$, and $b(c)$ is the bias of the estimator $\hat{m}(c)$. If the bias $b(c)$ is known, the approximate result above can be used directly to construct a confidence interval. When $b(c)$ is not known, the approximation can still be used, after substituting into (11) an estimate for $b(c)$. Unfortunately, estimation of $b(c)$ is rather complex, as highlighted in Eubank and Speckman (1993), since it involves the second derivative of $m(c)$, therefore some authors, among which Hastie and Tibshirani (1990), compute confidence bands for the entire curve based on the maximum deviation of the estimate $\hat{m}(c)$ from $m(c)$.

A widely accepted alernative is the computation of variability bands, i.e. pointwise confidence intervals for $E[\hat{m}(c)]$ rather than $m(c)$. They are computed as

$$\hat{m}(c) \pm 2 \times \sqrt{\hat{v}(c)}$$

Such bands indicate the level of variability involved in the regression estimator, without attempting to adjust for the presence of bias. Here all that is required is an estimate of the variance $v(c)$ of $\hat{m}(c)$, which is much easier to compute.

5 An empirical illustration

To investigate the behaviour of the economic shocks along the welfare distribution, the model developed above is now applied to Italian data covering the period 1980-2004.

5.1 Data description

The empirical analyses are run on data from the Survey on Household Income and Wealth (SHIW) from the Bank of Italy.

The SHIW is available from 1977 to 2004, data being collected annually from 1977 to 1987 except in 1985, then every second year from 1989 to 1995, and again biannually from 1998 to 2004. The design is substantially a series of cross-sections, but from 1989 a portion of the sample has been followed over time, producing panel data on a restricted number of households. Unfortunately the first three waves (1977 to 1979) cannot be used since consumption is not available in these years. Therefore all the cross-sectional analyses cover the period from 1980 to 2004. For the last study panel information over two periods is necessary, hence all the two-wave panels from 1989 to 2004 are used.

Since the economic model is valid far from the year of retirement, in principle it would be advisable to exclude all families whose head is retired. However, as shown in Table 1, among the families with head retired, about one fourth has incomes from employment and/or self-employment. This might be due to other family members who are still working. Since the income measure that is used is family disposable income, the behaviour of those family members who are not yet retired has to be taken into account. For this reason the only families who are excluded from the analysis are those where the head is retired and no income from employment and/or self-employment is present.

Households whose consumption or income fell below the 1^{st} percentile or above the 99^{th} were dropped. The sizes of the selected samples are reported in Tables 2–3. The analyses are conducted within six large cohorts, corresponding to heads of the family born in the 1920s, 1930s, 1940s, 1950s, 1960s and 1970s. Obviously in some cohorts the head is in retirement age, but there are working members in the family. The cohorts of families whose head was born before 1919 and after 1980 are excluded from the cohort analysis due to small sample size (less than 99 observations) in the cross-sectional and/or in the panel data. Families belonging to other cohorts, but still in groups with less than 99 observations, were excluded too from the cohort analysis.

As regards the variables used in the analysis, note that income is average monthly family disposable income, and the consumption measure is average monthly total expenditure, where the average is computed over the year of observation. Both income and consumption are rescaled according to an equivalence scale giving weight 1 to the head of the household, 0.7 to extra adults and 0.5 to any household member aged less than 18[3]. The analyses are run in the logarithmic scale; the use of the logarithmic scale is common practice, and can be found in several other studies on welfare.

5.2 Italy between 1980 and 2004

Before we start the empirical analysis, we report here some main facts and empirical findings that can give a general picture of the Italian economy between 1980 and 2004.

Brandolini *et al.* (2001) report several inequality and poverty indices based on income measures, such as the Gini index, the 90-10 decile ratio, the share of low-paid workers and the head count poverty ratio. All indices show a decrease starting from 1987 for a few years, going back to the 1987 values only in 1995.

Other factors that characterised the Italian economy in this period are the wage-setting mechanism "Scala Mobile" (literally "escalator"), which had a strong equalising effect on wages until 1992, when it was abolished. Moreover, in the early 1990s Italy experienced the deepest economic recession since the Second World War, which probably put all families in difficult economic situation, with an equalising effect.

Brandolini *et al.* (2004) report steady growth of housing prices from 1987 to 1993, and a decline of wealth inequality from 1989 to 1991, with a following rise in the rest

[3] A sensitivity analysis of the equivalence scale is available upon request.

of the decade "driven by large gains at the very top of the distribution". Bertola and Ichino (1995) show decreasing GDP rates in the period 1987-1994, and decreasing unemployment growth rate from 1989 to 1993. Real wage growth also decreased in the period 1989-1993, as the country transitioned from rigid to flexible employment.

5.3 Transitory shocks

Consider now the transitory shocks u_{it}. Before estimating the regression in equation (5), the test for no effect must be performed; only if the test gives indication of non-constancy of the regression function, the actual variance function will then be estimated. Both the PLR test and the OLS-based test were performed, and gave very similar results. The p-values of the nonparametric test are always smaller than 0.05, in most cases equal to zero, for any values of the smoothing parameter within a reasonable range, as we can see in Figure 1. Likewise, the OLS-based test presents all zero p-values.

When the sample is split into birth-cohorts, the OLS test very often (82% of the cases) rejects the hypothesis of constancy of the variance. The p-values are shown in Table 4. Since the behaviour of the PLR is extremely similar (as in the annual case), it was not performed in the analysis by cohort.

To estimate now the actual regression function, the standard local linear regression approach is used. In fact, the assumption of $\mathrm{Var}(\varepsilon) = \sigma_\varepsilon^2$ only affects the variability (hence the results of the test), but not the estimates. The smoothing parameter has been chosen case by case by cross-validation. For simplicity, the annual results jointly for all cohorts are shown. The optimal values for the smoothing parameter are reported in Figure 2. In two cases (1984 and 1995) the cross-validation procedure did not reach convergence, since the objective function (10) changed very little as the smoothing parameter varied. For these cases the parameter has been set equal to a fixed value.

The regression results are shown in Figure 2, together with the variability bands, and the logarithm of the poverty line (from Table 5). A local quadratic regression was also estimated and is shown in Figure 2; the curve is almost overlapping with the local linear estimate, and always lying within the variability bands. For this reason, we analyse here the results from the local linear model.

As expected, the results show that most heteroskedasticity is concentrated at the bottom of the consumption distribution, in most cases a cut-off point being values around the logarithm of the poverty line, superimposed in Figure 2, showing that the families that suffer a more volatile effect of the transitory economic shocks are those at the bottom of the consumption distribution.

In the central years of observation, in particular from 1987 to 1995, the shocks show a more homoskedastic behaviour, and then go back to higher variances at the low consumption levels after 1995. The more equal behaviour between 1987 and 1995 can be due to several factors that affected the Italian economy in those years, recalled in Section 5.2. In particular, since the transitory shocks are part of the income process, their more equal behaviour can be a reflection of the reduction in income inequality and poverty observed in this period in Brandolini *et al.* (2001).

In the analyses by cohort, the same smoothing parameter value was used for all cohorts in the same year, value corresponding to the optimal cross-validation estimate obtained in the analysis by year, as in Figure 2. This guarantees comparability of the annual results with the analyses by cohort. The results obtained separately for the six cohorts[4] did not show any cohort effect, and are extremely similar to the annual ones.

[4]Available upon request.

5.4 Permanent shocks

The analysis of the permanent shocks requires the use of panel information, since the v_{it} were obtained as $c_{it} - c_{it-1}$. The dataset used so far is mainly cross-sectional, but starting from 1989 it includes panel information too. Therefore this analysis can only be conducted for the last eight waves of the survey, collected in 1989, 1991, 1993, 1995, 1998, 2000, 2002 and 2004. Since the data are observed only every other year (and in 1995-1998 only after three years), economic shocks over two- (or three-) year periods are analysed. Therefore define $v_{it} = c_{it} - c_{it-2}$, and similarly for the three-year case.

Once obtained the values for the permanent shocks, we study regression (6). Again, due to biannual observations, such equation cannot be used directly; consider then the following regression:

$$v_{it}^2 = g(c_{it-2}) + \xi_{it}$$

The form of function g shows whether the impact of the permanent shocks is uniform along the distribution of consumption.

Similarly to what done in the case of the transitory shocks, a test of constancy of the g function is performed, using the two different techniques. Both in the OLS-based and in the nonparametric test, the hypothesis of constancy of the variance function is always rejected. The results of the PLRT are shown in figure 3, while in the OLS case all p-values are zero. The g function itself is then estimated, using local linear and local quadratic regression.

The results are reported in Figure 4, together with the variability bands and logarithm of the poverty lines, the local quadratic estimate and the optimal smoothing parameters.

We see again higher levels of variance in the bottom part of the distribution of consumption, i.e. the families with lower levels of consumption experience stronger impact of the economic shocks, but an overall more equal effect along the consumption distribution is evident, if compared to the transitory shocks. In particular, the variance levels at the top and bottom of the distribution are very similar, which did not happen for the transitory shocks where variances for the poor were around 1.5 to 2 times the variances of the rich.

Generally, the variances of the permanent shocks are smaller than those of the transitory shocks, showing that all families are affected more equally and that the wefare level makes little difference in dealing with permanent changes.

The behaviour of the shocks does not show large differences across time, levels being slightly flat only in the first wave (1989-1991), and then settling around a U-shaped pattern from 1993 onwards. The local quadratic regression yields results almost identical to the local linear case, as is evident from Figure 4.

In the analysis by cohort, only the OLS test was performed, as in the case of the transitory shocks. The test frequently rejected the hypothesis of constancy of the variance; p-values are shown in Table 6. We can notice that in general rejection is slightly lower than in the case of transitory shocks (73% of the times against 82% in the case of the transitory shocks). This confirms the more equal behaviour of the permanent shocks with respect to the transitory ones.

The estimates of the variance function by cohort[5] show no evident cohort pattern, the $g(c)$ function varying very little from cohort to cohort within the same year of observation.

5.5 Discussion

Putting together this last information with the results on transitory shocks, we can observe that the permanent shocks seem to be affecting all families more or less the

[5] Available upon request.

same way, while the transitory shocks appear to have a more uneven effect along the distribution of consumption, with the families with lower consumption levels experiencing higher volatility. What seems to be happening is that the transitory shocks can be absorbed through access to financial and credit system (which is not available to the poorer families), while the permanent shocks are more difficult to compensate, and therefore affect everyone more uniformly.

The more homoskedastic behaviour of transitory shocks in the central years could be interpreted as a signal of the more critical situation of the Italian economy during and after the recession of the early 1990s: even the richest families lost their ability to compensate for transitory shocks in the economy, so that such shocks affect all families similarly.

6 Conclusions

Starting from an economic model of intertemporal optimization, identification of permanent and transitory shocks of the economy is obtained through joint use of income and consumption measures.

The model allows us to derive two tests for the assumption of common variance of the shocks among all individuals in the same birth cohort, for a given time period, the first based on OLS results, the second on local polynomial modelling. Estimation of the conditional variance function of the shocks given the welfare level is performed using local polynomial regression models.

In an application with Italian data, strong heteroskedasticity is detected for the transitory shocks, while the permanent shocks seem to be affecting all families more equally. Higher variances are observed in the lower part of the distribution of consumption; this might be an index of reduced access to the credit system, or a more unstable job situation. The fact that permanent shocks are more "equal" might be showing that these shocks are more difficult to compensate, and therefore cannot be absorbed simply through access to financial and credit facilities.

Acknowledgments

This paper benefited from useful discussion with Montezuma Dumangane, Ian Preston, João Santos Silva, and from comments by audiences at ESEM04 and UCP. Thanks also go to the Bank of Italy for availability of data. This study was supported in part by the Fundação para Ciência e a Tecnologia, program POCTI, partially funded by FEDER. The usual disclaimer applies.

References

Abul-Naga, R. and Burgess, R. (1997). Prediction and determination of household permanent income. *Cahiers de Recherches Économiques*, No. 9705, Université de Lausanne.

Azzalini, A. and Bowman, A. W. (1997). *Applied Smoothing Techniques for Data Analysis*. Oxford University Press.

Bertola, G. and Ichino, A. (1995). Crossing the river: a comparative perspective on Italian employment dynamics. *Economic Policy*, **21**, 359-420.

Blundell, R. and Preston, I. (1998). Consumption inequality and income uncertainty. *The Quarterly Journal of Economics*, **113**, 603-640.

Brandolini, A., Cannari, L., D'Alessio, G. and Faiella, I. (2004). Household wealth distribution in Italy in the 1990s. Banca d'Italia, *Temi di Discussione*, n. 530. Fortcoming in *International Perspectives on Household Wealth* (ed E. N. Wolff), Edward Elgar, Cheltenham.

Brandolini, A., Cipollone, P. and Sestito, P. (2001). Earnings dispersion, low pay and household poverty in Italy, 1977-1998. In *The New Economics of Rising Inequalities* (eds D. Cohen, T. Piketty and G. Saint-Paul), pp. 225-264, Oxford University Press.

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

Eubank, R. L. and Speckman, P. L. (1993). Confidence Bands in Nonparametric Regression. *Journal of the American Statistical Association*, **88**, 1287-1301.

Fan, J. and Gijbels, I. (1996). *Local polynomial modelling and its applications.* Chapman and Hall.

Hall, R. E. (1978). Stochastic implications of the life-cycle permanent income hypothesis: theory and evidence. *Journal of Political Economy*, **86**, 971-987.

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized additive models.* London: Chapman and Hall.

Ruppert, D. and Wand, M. P. (1994). Multivariate locally weighted least squares regression. *Annals of Statistics*, **22**, 1346-1370.

Table 1: Composition of the sample (%) according to the occupational state of the head of the household. Cross-sectional data from SHIW, 1980-2004. (1) = Head of the household not retired (2a) = Head of the household retired, with income from employment and/or self-employment in the family (2b) = Head of the household retired, with no income from employment and/or self-employment in the family

Year	**(1)**	**(2a)**	**(2b)**	**Total**
1980	71.79	9.25	18.97	100
1981	70.97	9.18	19.85	100
1982	75.11	8.02	16.86	100
1983	71.60	8.81	19.60	100
1984	70.43	8.00	21.58	100
1986	68.71	7.32	23.97	100
1987	70.47	7.42	22.11	100
1989	66.69	8.82	24.49	100
1991	61.60	10.82	27.58	100
1993	59.20	11.25	29.55	100
1995	58.30	12.15	29.55	100
1998	61.35	11.05	27.59	100
2000	58.77	11.70	29.53	100
2002	54.85	11.50	33.65	100
2004	55.83	9.91	34.26	100

Table 2: Sample sizes of cross-sectional data from SHIW, 1980-2004, by birth cohort of the head of the household. Cells in parentheses not used in the analyses by cohort.

	Cohort									
Year	**<1910**	**1910s**	**1920s**	**1930s**	**1940s**	**1950s**	**1960s**	**1970s**	**1980s**	**Total**
1980	(0)	(168)	774	633	524	195	(0)	(0)	(0)	2294
1981	(0)	(225)	954	807	758	364	(0)	(0)	(0)	3108
1982	(0)	(209)	1038	838	720	323	(0)	(0)	(0)	3128
1983	(0)	(210)	1010	902	782	246	(0)	(0)	(0)	3150
1984	(54)	(178)	593	894	874	496	(33)	(0)	(0)	3122
1986	(56)	(197)	902	1647	1660	1192	174	(0)	(0)	5828
1987	(60)	(223)	814	1628	1771	1290	200	(0)	(0)	5986
1989	(27)	(147)	631	1487	1715	1448	540	(4)	(0)	5999
1991	(22)	(115)	505	1378	1700	1401	567	(9)	(0)	5697
1993	(11)	(92)	442	1147	1583	1443	736	(24)	(0)	5478
1995	(4)	(65)	360	1020	1620	1437	914	(85)	(0)	5505
1998	(3)	(37)	209	688	1405	1441	1029	163	(2)	4977
2000	(2)	(41)	227	627	1356	1568	1264	325	(19)	5429
2002	(0)	(26)	195	533	1190	1488	1243	408	(26)	5109
2004	(0)	(20)	137	390	1050	1486	1344	571	(65)	5063
Total	(239)	(1953)	8791	14619	18708	15818	8044	1589	(112)	69873

Table 3: Sample sizes of utilised two-wave panel data from SHIW, 1989-2004, by birth cohort of the head of the household. Cells in parentheses not used in the analyses by cohort.

	Cohort									
Year	**<1910**	**1910s**	**1920s**	**1930s**	**1940s**	**1950s**	**1960s**	**1970s**	**1980s**	**Total**
89-91	(3)	(36)	129	356	443	379	117	(1)	(0)	1464
91-93	(6)	(42)	166	478	768	610	203	(3)	(0)	2276
93-95	(1)	(31)	154	440	747	676	288	(8)	(0)	2345
95-98	(0)	(13)	(85)	254	533	523	268	(21)	(0)	1697
98-00	(0)	(10)	(85)	282	700	774	492	(60)	(0)	2403
00-02	(0)	(7)	(69)	229	558	706	498	99	(7)	2173
02-04	(0)	(8)	(56)	178	489	675	517	130	(7)	2060
Total	(10)	(147)	744	2217	4238	4343	2383	322	(14)	14418

Table 4: P-values of the OLS-based test of heteroskedasticity of the transitory shocks. Cross-sectional data from SHIW, 1980-2004, by birth cohort of the head of the household.

	Cohort						
Year	**1920s**	**1930s**	**1940s**	**1950s**	**1960s**	**1970s**	**All data**
1980	0,0000	0,0000	0,0000	0,0002	-	-	0,0000
1981	0,0000	0,0000	0,0000	0,0024	-	-	0,0000
1982	0,0000	0,0000	0,0000	0,0000	-	-	0,0000
1983	0,0000	0,0000	0,0055	0,0053	-	-	0,0000
1984	0,0000	0,0000	0,0000	0,0000	-	-	0,0000
1986	0,0000	0,0000	0,0000	0,0000	0,1753	-	0,0000
1987	0,2205	0,0531	0,0215	0,0026	0,2518	-	0,0030
1989	0,0146	0,0732	0,0001	0,0001	0,0012	-	0,0000
1991	0,0014	0,0000	0,0000	0,0015	0,5436	-	0,0000
1993	0,0004	0,0000	0,0003	0,1047	0,6472	-	0,0000
1995	0,0377	0,0077	0,9513	0,7136	0,2442	-	0,0038
1998	0,0000	0,0000	0,0000	0,0000	0,0000	0,0777	0,0000
2000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0013	0,0000
2002	0,0000	0,0000	0,0000	0,0000	0,0000	0,1071	0,0000
2004	0,0023	0,0000	0,0000	0,0000	0,0000	0,1750	0,0000

Table 5: Poverty line equivalent to 50% of mean consumption, for a 2-component family. Thousand Liras.

Year	**1980**	**1981**	**1982**	**1983**	**1984**	**1986**	**1987**	**1989**
Line	154	194	219	249	283	316	419	475
log(Line)	5.04	5.27	5.39	5.52	5.65	5.76	6.04	6.16

Year	**1991**	**1993**	**1995**	**1998**	**2000**	**2002**	**2004**
Line	538	584	667	703	788	837	967
log(Line)	6.29	6.37	6.50	6.56	6.67	6.73	6.87

Table 6: P-values of the OLS-based test of heteroskedasticity of the permanent shocks. Two-wave panel data from SHIW, 1989-2004, by birth cohort of the head of the household.

	Cohort						
Year	**1920s**	**1930s**	**1940s**	**1950s**	**1960s**	**1970s**	**All data**
89-91	0,3657	0,0409	0,0074	0,3970	0,1779	-	0,0032
91-93	0,0389	0,0963	0,0066	0,0000	0,0264	-	0,0000
93-95	0,0116	0,0001	0,0000	0,0398	0,0670	-	0,0000
95-98	-	0,2590	0,0084	0,0124	0,0325	-	0,0000
98-00	-	0,0006	0,0000	0,0001	0,0002	-	0,0000
00-02	-	0,0131	0,1102	0,0001	0,0053	0,1327	0,0000
02-04	-	0,1192	0,0001	0,0006	0,0001	0,3460	0,0000

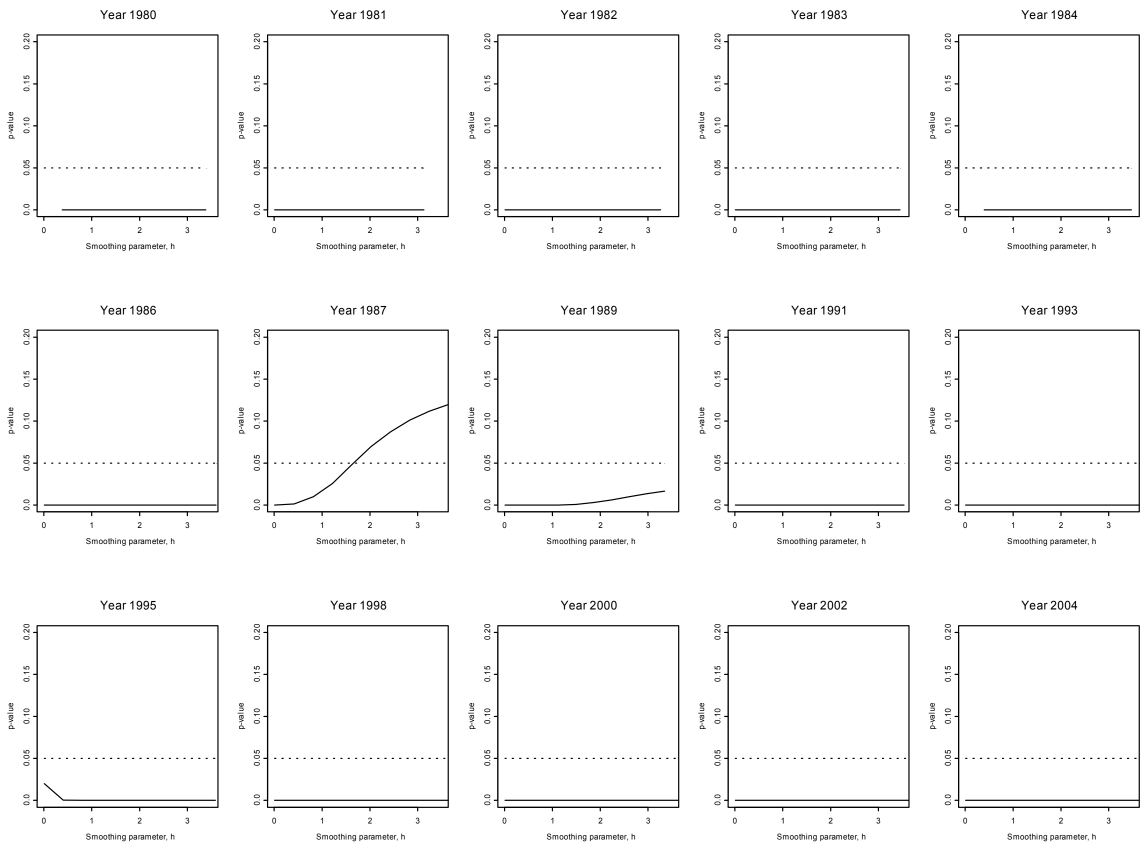

Figure 1: Test for no effect for the transitory shocks. P-values as function of smoothing paramenter h. SHIW data, 1980-2004.

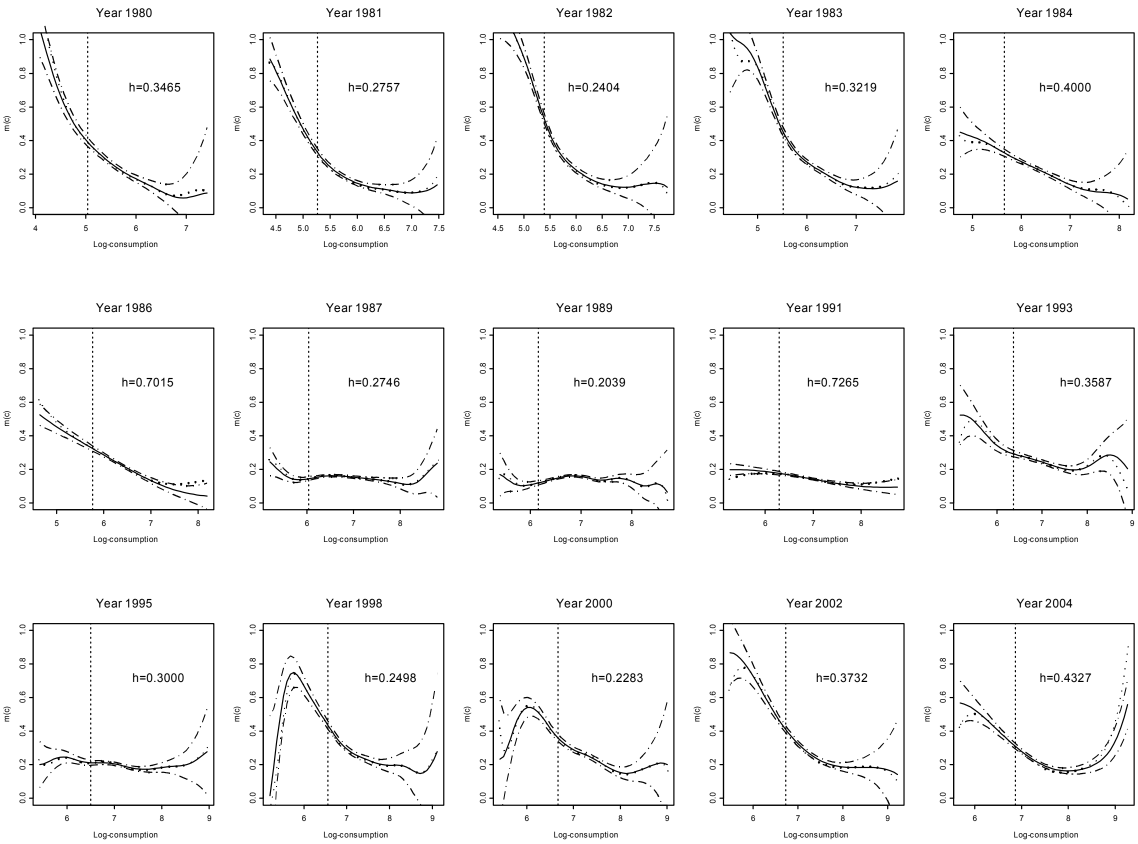

Figure 2: Nonparametric estimates of the variance function for the transitory shocks, as function of log-consumption. Local linear regression (—), local quadratic regression $(\cdots)$, variability bands $(- \cdot - \cdot -)$. Log of poverty line superimposed (vertical dashed line). Optimal cross-validation smoothing parameter. SHIW data, 1980-2004.

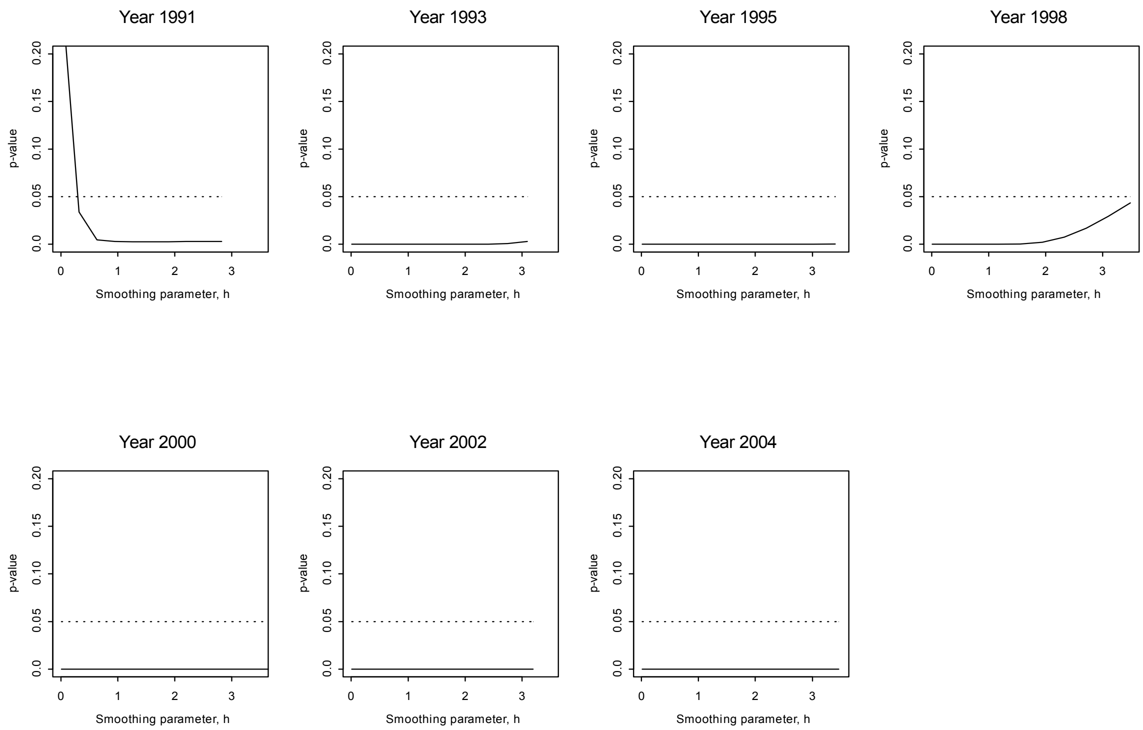

Figure 3: Test for no effect for the permanent shocks. P-values as function of smoothing paramenter h. SHIW data, 1980-2004.

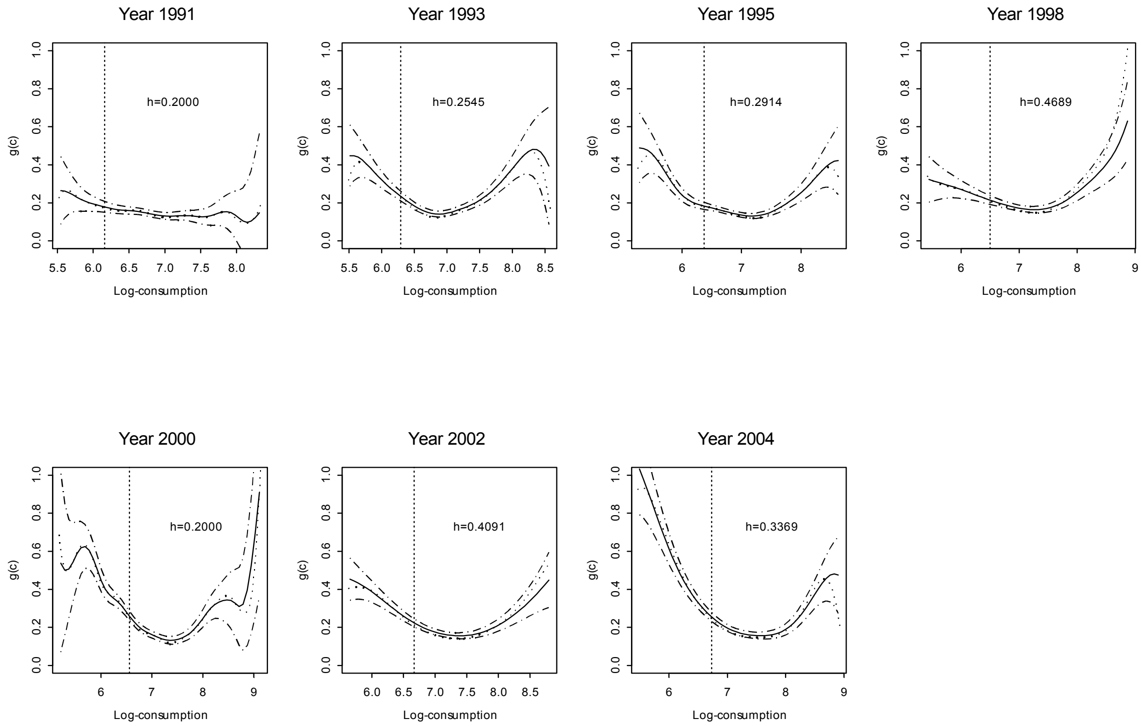

Figure 4: Nonparametric estimate of the variance function for the permanent shocks, as function of log-consumption. Local linear regression (—), local quadratic regression (· · ·), variability bands (— · — · —). Log of poverty line superimposed (vertical dashed line). Optimal cross-validation smoothing parameter. SHIW panel data, 1989-2004.