Poverty and Permanent Income: A Methodology for Cross-Section Data

Ramses H. Abul Naga and Enrico Bolzani*
DEEP, Université de Lausanne, CH-1015 Lausanne, Switzerland.

November 24, 2000

Abstract

If the set of households which are income poor does not fully overlap with the set of the consumption poor, it could well be that income and consumption expenditure convey different information regarding an unobserved variable on the basis of which families allocate their resources intertemporally. This paper presents a methodology for predicting the unobserved permanent incomes of households using multiple welfare indicators typically available in cross-section data. The methods are illustrated using data from the Swiss Consumption Survey of 1990.

Keywords: Poverty, permanent income, latent variables, prediction, Switzerland.

JEL codes: C_2, C_3, D_6, I_3

*This research was supported by a grant no. 1217-53567.98 from the Swiss National Science Foundation. We wish to thank Alberto Holly and Christophe Kolodziejczyk for discussions. We are grateful to Robin Burgess for several suggestions which were developed in the present paper. We are solely responsible for any mistakes or omissions.

1. Introduction

Whether an economy is growing or stagnating, the extent of poverty and its incidence ought to remain causes for concern for a government at all times. The days when growth was believed to provide a miracle solution for poverty alleviation are now gone and, as a result, policy makers and researchers alike have increasingly been drawing a distinction between transient and chronic poverty.

The identification of the long-term poor, at a given moment in time, needless to say raises some conceptual problems related to the measurement of resources. Cross-section data, the environment discussed in the present paper, provide a single measurement on a household's income and consumption expenditure, together with other demographic and socio-economic information, on the basis of which a researcher is to attempt to draw lessons about the extent and incidence of chronic poverty. A common practice in the area consists in defining a benchmark of concept of resources[1], and to examine, in light of this concept, how competing welfare indicators perform in identifying the poor. Thus Glewwe and Van der Gaag (1990) conclude that alternative welfare indicators identify different populations as being in poverty, and that the choice of income concept does matter indeed. Chaudhuri and Ravaillon (1994) arrive at similar conclusions on the basis of multiple measurements on the household's income and consumption expenditure.

Of the two most commonly used welfare indicators, namely family income and consumption expenditure, the latter has perhaps received more support on the grounds that it comes closer to an understanding of poverty as being associated with low levels of living[2]. However, the approach of selecting a unique, in some ways superior, welfare indicator does not address fully the question as to what to do in practice when different indicators identify different groups of individuals as being in poverty. Furthermore, if it is the case that alternative welfare indicators convey different information about a household's long-term income, working with a single welfare indicator is invariably associated with a loss of statistical information. Proposing methods of identifying the poor, which exhaust more of the information typically available in cross-section surveys regarding the long term incomes of households, may provide a common ground for reconciling conflicting evidence derived from income and consumption based measures of resources. More importantly, multi-dimensional approaches may provide researchers with

[1] See for instance the proposition of Anand and Harris (1990).

[2] There is also an economic relation between consumption expenditure and life-time resources (that is in an intertemporal context) which will be discussed in section 2 of the paper.

new tools for identifying the long-term poor, a necessary step for understanding the determinants of chronic poverty, and subsequently for designing appropriate welfare support policies.

In the present paper household permanent income is defined as the expected value of life-time resources, averaged over adult life (more on this below). Permanent income being typically unobservable we discuss prediction of this variable using (1) multiple welfare indicators, (2) multiple determinants and (3) multiple indicators and determinants (simultaneously). The weights given to the various measurements are dictated from a class of minimization problems of mean-square error of prediction.

The plan of the paper is the following. Section 2 considers an economic model of the intertemporal allocation of consumption subject to income uncertainty (hence the earlier definition of permanent income as being a function of expected life-time resources). The purpose of this section is to make explicit the assumptions under which the econometric framework (section 3) for the joint modelling of income and consumption expenditure is derived. In section 3 we review the multiple indicator formulations of Abul Naga (1994) and Mercader-Prats (1997). Then, we present the multiple indicators and multiple causes (MIMIC) framework which Abul Naga and Burgess (1997) have used to derive three related predictors of household permanent income.

The multiple indicator approach has its roots in the statistical model of factor analysis. The MIMIC model proposed by Zellner (1970) and Jöreskog and Goldberger (1975) has been used by Muellbauer (1983) in order to analyze living standards in Sri Lanka. Both the multiple indicator and MIMIC frameworks are generalized in the present paper by introducing separate components of variation for household demographic variables. We also discuss specification testing after estimation and prior to prediction, a point largely ignored by Abul Naga (1994) Mercader-Prats (1997) and Abul Naga and Burgess (1997). In section 4 we present our Swiss household data extracted from the 1990 consumption survey, the Enquête sur la Consommation des Ménages. The applications are used in section 5 to illustrate how a multi-dimensional approach may help the researcher in resolving some conflicting evidence regarding the incidence of poverty, derived from income and consumption-based definitions of resources. Section 6 concludes with a summary and discussion of various limitations of the proposed approaches, which may form the basis of an agenda for a further research in the area.

As indicated in the title of the paper, the methods proposed here are intended for application to cross-section data. As such, they do not allow for a rich modelling of the dynamics of permanent income as in the panel data context (for instance, see Hall and Mishkin, 1982), or in the repeated cross-section environment (for example Blundell and Preston, 1998). Nonetheless, they are less demanding in terms of data requirements, since they necessitate only data from a unique wave of a cross-section survey. It is also the case that they exhaust information on household permanent income from cross-section data in a way not considered in other approaches. Finally, the emphasis here is different, primarily being placed on the identification of the poor.

2. Consumption, Income and Permanent Income

In this section we consider a resource allocation problem for a household assumed to live for two periods: today, for which the data analyst observes information from a cross-section survey, and tomorrow, the unknown future. Though highly simplified, the example helps to motivate the discussion that follows in section 3 on the joint modelling of income and consumption expenditure. A discussion on the joint modelling of income and consumption expenditure within the life-cycle context may also be of some help in understanding some apparently contradictory findings regarding the well-being of Swiss households. Finally, and most importantly, understanding the assumptions underlying a linear model of income and expenditure (see below) helps to identify potential weaknesses and drawbacks of the proposed approach, and to draw an agenda for further research.

Assume then that a household maximizes a utility function $\Psi(C_1; C_2)$ taken to be additively separable over time and states of nature. If within-period utility is quadratic, and we consider two states for tomorrow's income: L (low) and H (high), we may write the objective function as:

$$\Psi(C_1; C_2) = \varphi(C_1) + \frac{1}{1+\rho}\left[\pi\varphi(C_{2L}) + (1-\pi)\varphi(C_{2H})\right] \tag{1}$$

where π denotes the probability of the low income state, ρ is the discount factor and $\varphi(C) = C - C^2/2$. Resources are allocated subject to a budget constraint, which in the present setting takes the form:

$$C_1 + \frac{1}{1+r}\left[\pi C_{2L} + (1-\pi) C_{2H}\right] = A + m_1 + \frac{1}{1+r}\left[\pi m_{2L} + (1-\pi) m_{2H}\right] \tag{2}$$

The interest rate r is taken to be a constant (known) quantity, A is the initial stock of assets and m denotes labor income.

Let λ denote the Lagrange multiplier on the resource constraint. First order conditions entail the following:

$$\frac{\partial L}{\partial C_1} = 1 - C_1 - \lambda = 0 \tag{3}$$

$$\frac{\partial L}{\partial C_2} = \frac{\pi(1 - C_{2L}) + (1 - \pi)(1 - C_{2H})}{1 + \rho} - \frac{\lambda}{1 + r} = 0 \tag{4}$$

The above equations show that the marginal utility of consumption at a given time period is a function of the Lagrange multiplier on the resource constraint λ and of the price of consumption, a function of the interest rate. It is also the case that the marginal utility of consumption at time t is independent of the chosen level of consumption for period s.

If we assume that the rate of time preference is constant in the population and equal to the interest rate ($\rho = r$), and we let $E(\cdot)$ denote the expectations operator, we may obtain at once

$$C_1 = E(C_2) \tag{5}$$

Equation (5) tells us that, in mathematical expectation, consumption is constant over the life cycle. Underlying this martingale property is the assumption that the within period utility function is quadratic, so that the marginal utility of consumption becomes linear. The assumption of quadratic utility provides a formal justification for Friedman's (1957) formulation of the Permanent Income Hypothesis (PIH). While convenient for empirical work, the above assumption of quadratic utility drowns the rationale for household saving for precautionary motives. Other hidden assumptions in the story include the hypothesis that households may be allowed to borrow and lend at the same interest rate r and that labor supply is exogenous at the family level. While these assumptions may be more easily justified at the aggregate level, they may pose more serious conceptual problems at the micro-level, especially when one seeks to identify the long term poor. We shall return to some of these points in our final section where we discuss directions for further work.

Substituting (4) and (5) in the budget constraint we have that

$$C_1 \left(\frac{2 + r}{1 + r} \right) = A + m_1 + \frac{E(m_2)}{1 + r} \tag{6}$$

The above equation shows that, in theory, if a researcher seeks to identify the long run poor, then period 1 consumption expenditure (which s/he may observe on the basis of a cross-section survey) contains all the relevant information concerning the household's lifetime resources. Alternatively, in the words of Slesnick (1998):

"Total expenditure (along with the marginal utility of wealth) serves as a sufficient statistic that links the within period model to the intertemporal allocation of consumption. The use of income in static demand models, which is very common, results in a potentially serious misspecification because of its systematic understatement of expenditure for low-income households and the reverse for those with high incomes. This turns out to be very important when welfare measures are used to examine distributional issues such as inequality or poverty."

In practice, however, consumption expenditure is unlikely to be measured without error. On such grounds, period 1 outlay, $m_1 + A(1+r)$, may also contain information about the households' lifetime resources. More specifically, define permanent income as an annual average of the households lifetime resources: $\mu = [A + m_1 + E(m_2)/(1+r)]/2$, then we may write the following system of equations for period 1 observed income, $I_1 = rA + m_1$, and expenditure:

$$I_1 = \mu + e_{I1} \tag{7}$$

$$C_1 = \beta\mu + e_{C1} \tag{8}$$

where e_{I1} and e_{C1} are respectively transitory variations in observed income and expenditure from their long-run components.

Equation (7) mirrors the Friedman and Kuznets (1945) decomposition of observed income into permanent and transitory components. Equation (8), subsequently introduced by Friedman (1957), also decomposes observed consumption expenditure into permanent and transitory component with β being the marginal propensity to spend (out of permanent income). The formulation (8) is in fact more general than may appear to be the case by looking at our simplified example. In this respect, see Gorman's (1959) discussion on the general preference structure underlying Friedman's consumption function.

It may appear from the discussion above that consumption expenditure should primarily be used for the analysis of levels of living, and that a case can only be made for the use of income concepts to the extent that the former variable is subject to measurement error. This is certainly not the case. A researcher may still want to use an income concept in order to identify the poor, though the motivation underlying this choice is no longer based on a standard of living approach; but on an entitlements, or rights, notion. Atkinson (1989, p. 12) writes:

"The second conception is that of poverty as concerned with the right to a minimum level of resources. On this basis, families are entitled, as citizens, to a minimum income, the disposal of which is a matter for them. This approach may be more appealing to those who see concern for poverty as based on a notion as to what constitutes a good society".

Finally, note that within the standards of living approach a case can be made for identifying the poor on the basis of a threshold consumption of specific goods, which may be essential to the functioning of individuals in society. If certain goods (of which housing is a frequently cited example) are subject to rationing or other forms of market failures, one may want to pursue a multiple deprivations approach to the identification of the poor, rather than to use an aggregate concept such as total consumption expenditure[3].

3. Methods

Given data on a household's income and consumption expenditure in a given year, what can we hope to infer about its permanent income? Starting from equations (7) and (8), this section reviews various approaches to the prediction of household permanent income. In sub-section 3.1 we review the multiple indicator approaches proposed by Abul Naga (1994) and Mercader-Prats (1997). In sub-section 3.2 we summarize the multiple indicator and multiple causes (MIMIC) methodology adopted by Abul Naga and Burgess (1997) for this same problem. For both approaches we propose generalizations which provide, what we believe to be, more flexible ways of controlling for household demographic structure.

The common ground between the various approaches discussed here is that they are multiple equation systems. The multiple indicator approach of sub-section 3.1 can be read as a factor analysis where income and expenditure are assumed to correlate through their common dependence on the (unobserved) permanent income of the family. While the last three decades have seen tremendous progress in understanding the statistical foundations of factor analysis (see for instance the monographs of Lawley and Maxwell, 1971; and Bartholomew, 1987),

[3] See Atkinson (1989), ch.1, for a discussion on the parallel between the multiple deprivation approach and Tobin's (1970) concept of specific egalitarism.

such methods are received with a fair amount of skepticism in the economics profession. For this reason, in sub-section 3.3 we follow Chamberlain (1977) in providing an instrumental variables interpretation of estimation in the multiple indicator approach. The MIMIC model (Zellner, 1970; Jöreskog and Goldberger, 1975) is easily embedded within the econometrics literature on simultaneous equations systems. The instrumental variables avenue to estimation of the MIMIC model is also reviewed. Finally, as a means of checking the plausibility of ones estimates, we also suggest in sub-section 3.3 various specification tests that the researcher may wish to undertake after estimation.

3.1. Predicting permanent income using multiple indicators

Going back to equations (7) and (8) we shall let Y_1 denote household income $(Y_1 = I_1)$, Y_2 denote consumption expenditure $(Y_2 = C_1)$, and u_j denote the transitory component associated with Y_j. Assume that, all in all, the researcher has data on $j = 1, ..., p$ indicators of permanent income on a sample of $i = 1, ..., n$ observations:

$$\begin{aligned} y_{i1} &= \eta_i + u_{i1} \\ y_{i2} &= \beta_2 \eta_i + u_{i2} \\ &\vdots \\ y_{ip} &= \beta_p \eta_i + u_{ip} \end{aligned}$$

Define $Y_i = [y_{i1}, y_{i2}, ..., y_{ip}]'$, $\beta = \left[1, \beta_2, ..., \beta_p\right]'$ and $U_i' = [u_{i1}, u_{i2}, ..., u_{ip}]'$. The vector notation for the system of p equations related to household i takes the form

$$Y_i = \beta \eta_i + U_i \tag{9}$$

where η_i is an unobserved random variable, β is a $p \times 1$ vector of unknown structural parameters and Y_i and p_i are p-dimensional random vectors. It is assumed throughout that $E[U_i | \eta_i] = 0$.

Counting sample moments and unknowns

One natural question to ask is how many indicators of permanent income one must observe in order to be able to predict η. Given that we have set $\beta_1 = 1$, we could clearly predict η using the household's observed income Y_1. More generally however, if one is to make some claims about the statistical properties of a predictor (see below) some knowledge regarding the structural parameters of (9) is required. Let Ω denote the covariance matrix of U_i and let $\sigma_{\eta\eta}$ denote the variance of η_i. There is a total of $p(p+1)/2$ unknown parameters in Ω, $p-1$ unknowns in β and assuming η_i has a zero mean[4], a further unknown parameter being the variance of η_i. In its general form then, (9) necessitates the estimation of $p+p(p+1)/2$ parameters, on the basis of $p(p+1)/2$ sample moments available from the p indicators on η. In general therefore, model (9) cannot be identified without imposing some restrictions on the vector β or the matrix Ω. The assumption underlying the model of factor analysis is that Ω is a diagonal matrix. Letting Σ_F denote the covariance matrix of Y, we have

$$\Sigma_F = \sigma_{\eta\eta}\beta\beta' + \Omega \tag{10}$$

Σ_F is the sum of a unit rank matrix $\sigma_{\eta\eta}\beta\beta'$ arising from the common dependence of the p indicators on η, and a full rank diagonal matrix Ω pertaining to the transitory, specific, variance components of Y. Under the factor analytic covariance structure Ω possesses p non-zero elements, so that the total number of unknowns sums to $2p$ structural parameters. A necessary condition that must be met for identification is that the total number of unknowns does not exceed the number of sample moments. In the present context this condition takes the form $2p \leq p(p+1)/2$. The bottom line then is that a minimum of $p = 3$ indicators is required in order to identify (9). When $p > 3$ the model is potentially over-identified, thus allowing the researcher to test the plausibility of her/his specification. This point will be discussed further in sub-section 3.3.

Mercader-Prats' framework

As pointed out by Deaton and Muellbauer (1980, pp.103-105), there are different versions of the PIH. Permanent income may take a different meaning than life-time resources (or their expectation) and β_2 also takes a variety of interpretations. Mercader-Prats (1997) considers a case where β_2, the proportion of permanent income allocated to family consumption, is specific (depending on demographics such as family size and composition), and the only available indicators of η_i are household income and consumption expenditure:

$$y_{i1} = \eta_i + u_{i1} \tag{11}$$

[4] This amounts to measuring $y_{i1}, \ldots, y_{ip}$ in deviation from their respective sample means $\overline{y_1}, \ldots, \overline{y_p}$.

$$y_{i2} = \beta_{i2}\eta_i + u_{i2} \tag{12}$$

where the subscript i is introduced to highlight that the marginal propensity to consume is family specific. Mercader-Prats works with the assumption that household equivalence scales can be used to approximate β_{i2}, and that these may be constructed from the data. Under such circumstances β_{i2} is no longer an unknown structural parameter, and the remaining unknowns are the variances of the transitory income and consumption components, ω_{11} and ω_{22}, together with $\sigma_{\eta\eta}$. Income and consumption expenditure provide 3 sample moments: two variance terms and a covariance. On such basis, the system (11-12) may be identified provided β_{i2} is approximated by an equivalence scale.

Prediction

Going back to the general model (9), the next point we wish to consider is what may be inferred about household permanent income once we have observed the indicators $y_{i1}, \ldots, y_{ip}$[5]. Given the linearity assumption underlying (9), it is natural to focus our discussion on the class of linear predictors, i.e. statistics of the form $\hat{\eta}_Y = a'Y_i$. To the extent that observations are collected from random samples, it should be the case that data on family j should be uninformative about the permanent income of family i. For this reason, below we suppress the subscript i, and we simply write

$$\hat{\eta}_Y = a'Y \tag{13}$$

for the class of linear predictors of η. The optimal linear predictor in a mean-square error sense [MSE] chooses a in order to minimize the criterion

$$E\,(a'Y - \eta)^2 \tag{14}$$

First order conditions for the above problem yield

$$2E\,[(a'Y - \eta)\,Y'] = 0 \tag{15}$$

since $E\,(YY') = \Sigma_F$ (cf. equation 10), and $E\,(\eta Y') = \sigma_{\eta\eta}\beta'$ we obtain at once that $a' = \sigma_{\eta\eta}\beta'\Sigma_F^{-1}$; that is, the optimal MSE predictor takes the form

[5] As Mercader-Prats normalizes income and expenditure in her data by appropriate equivalence scales, predictors for her system (11-12) may be obtained simply by setting $\beta = \begin{bmatrix} 1 & 1 \end{bmatrix}'$ in the formulae that follow.

$$\hat{\eta}_Y = \sigma_{\eta\eta}\beta'\Sigma_F^{-1}Y \tag{16}$$

The above predictor is also known as the Best Linear Predictor (BLP) of η. Writing $\eta = \hat{\eta}_Y + \xi$ we have

$$\xi = \eta - \hat{\eta}_Y = \eta - \sigma_{\eta\eta}\beta'\Sigma_F^{-1}\beta\eta - \sigma_{\eta\eta}\beta'\Sigma_F^{-1}U \tag{17}$$

The BLP is unbiased in the sense that $E(Y\xi) = 0$. That is, the prediction error ξ is uncorrelated with the data Y. Noting furthermore that U and η have zero means, it also follows that $E(\xi) = 0$.

In what follows however we shall work with a different class of predictors than the BLP. Members of this class may be shown to be unbiased in a certain sense, though the main motivation for working with this family of predictors will become clearer as we examine the more general MIMIC framework.

Consider then the class of linear predictors $\tilde{\eta}_Y = b'Y$ such that the vector b satisfies the condition $b'\beta = 1$:

$$C_Y = \{\tilde{\eta}_Y = b'Y \mid b'\beta = 1\} \tag{18}$$

Then, suppressing the subscript i in (9), we have

$$\eta^*_Y = \eta + b'U \tag{19}$$

In the social sciences members of the class (18) are said to be unbiased in the sense that $E(\eta^*_Y \mid \eta) = \eta$ (see Lawley and Maxwell, 1971, pp. 109-111). It may be noted that the BLP, $\hat{\eta}_Y$, is not a member of the class (18), since $a'\beta = \sigma_{\eta\eta}\beta'\Sigma_F^{-1}\beta$, which need no equal unity.

We do not wish to debate which of the two approaches is better. The relation between (16) and members of the class C_Y is best considered by examining the minimum mean-square error predictor in the class (18). This latter predictor is solution to the following problem:

$$\min E(b'Y - \eta)^2 - \mu(b'\beta - 1) \tag{20}$$

where μ is the Lagrange multiplier on the unbiasedness constraint. As shown by Lawley and Maxwell, the above problem yields for solution the predictor

$$\eta^*_Y = \left(\beta'\Sigma_F^{-1}\beta\right)^{-1}\beta'\Sigma_F^{-1}Y \tag{21}$$

By noting that $\beta'\Sigma_F^{-1}\beta$ is a scalar, it may be observed that η^*_Y, originally derived by Bartlett (1937), only differs from the BLP $\hat{\eta}_Y$ by a constant. For the purpose of predicting permanent incomes, both predictors will rank households in the same order.

It is possible to show with some algebra that (21) may be written in the form

$$\eta^*_Y = \left(\beta'\Omega^{-1}\beta\right)^{-1}\beta'\Omega^{-1}Y \tag{22a}$$

or more simply, for $\kappa = (\beta'\Omega^{-1}\beta)^{-1}$,

$$\eta^*_Y = \kappa \cdot \left(\frac{1}{\omega_{11}}y_1 + \frac{\beta_2}{\omega_{22}}y_2 + \ldots + \frac{\beta_p}{\omega_{pp}}y_p\right) \tag{22b}$$

That is, for a given vector β, the larger the transitory variance ω_{jj} of a given indicator j, the less weight this variable will be assigned in the prediction of the household's permanent income.

Controlling for household structure

We have noted earlier in section 2 that our intertemporal consumption allocation problem was lacking in realism. One simplifying assumption was that the marginal utility of consumption was unrelated to household composition. Because labor supply was taken to be exogenous, no story could be told about how an increase in family size, the presence of under schooling age children, the number of adults living in the household, etc., could influence the income receipt of the family.

For the purpose of controlling for household demographic structure, we therefore propose to generalize (9) by introducing a set D_j of l_j socio-demographic variables for equation j:

$$\begin{aligned} y_{i1} &= \eta_i + D'_{i1}\delta_1 + u_{i1} \\ y_{i2} &= \beta_2\eta_i + D'_{i2}\delta_2 + u_{i2} \\ &\vdots \\ y_{ip} &= \beta_p\eta_i + D'_{ip}\delta_p + u_{ip} \end{aligned}$$

Letting D_i denote the block-matrix

$$D_i = \begin{bmatrix} D'_{i1} & 0 & \cdots & & 0 \\ 0 & D'_{i2} & 0 & & \vdots \\ \vdots & 0 & \ddots & & \\ & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & D'_{ip} \end{bmatrix}$$

of dimensions $p \times l$, where $l = \sum_j l_j$, and let $\delta = \left[\delta'_1; \ldots; \delta'_p\right]'$, our modified p-equation system now becomes

$$Y_i = \beta\eta_i + D_i\delta + U_i \tag{23}$$

which is no longer a factor-analytic structure, but a system of Seemingly Unrelated Regression Equations (the SURE model of Zellner, 1962). By noting that η_i is an unobservable, we may prepare the ground for our transition to the more general MIMIC framework by writing (23) as

$$Y_i = D_i\delta + V_i \tag{24}$$

with $V_i = \beta\eta_i + U_i$. V_i, the disturbance vector of (24), now exhibits a factor analytic covariance structure. This may be seen by writing the covariance matrix Σ_S underlying the model (24):

$$\Sigma_S = \Delta + \sigma_{\eta\eta}\beta\beta' + \Omega \tag{25}$$

where Δ is the matrix

$$\Delta = \begin{bmatrix} \delta'_1 E\left(D_1 D'_1\right)\delta_1 & \delta'_1 E\left(D_1 D'_2\right)\delta_2 & \cdots & \delta'_1 E\left(D_1 D'_p\right)\delta_p \\ \vdots & \delta'_2 E\left(D_2 D'_2\right)\delta_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ \delta'_p E\left(D_p D'_1\right)\delta_1 & \cdots & & \delta'_p E\left(D_p D'_p\right)\delta_p \end{bmatrix} \tag{26}$$

so that the covariance matrix of the vector V_i now exhibits a factor-analytic structure given by the components $\sigma_{\eta\eta}\beta\beta' + \Omega$. Under the present model, disturbances v_{ij} are correlated across equations as in Zellner's (1962) formulation, with the particular feature here that the covariance structure of the disturbances is parametrized.

By noting that (23) in translated form, i.e. $Y_i - D_i\delta = \beta\eta_i + u_i$, has once again a factor-analytic structure, we may (suppressing the subscript i) consider predicting η using the class of predictors

$$C^0_Y = \{\hat{\eta}_Y = b'(Y - D\delta) \mid b'\beta = 1\} \tag{27}$$

for which the minimum square error predictor takes the form:

$$\hat{\eta}^*_Y = \left(\beta'\Sigma_S^{-1}\beta\right)^{-1}\beta'\Sigma_S^{-1}(Y - D\delta) \tag{28}$$

Demographic variables here act in a way as to translate income, expenditure, etc., from the origin. This approach may of course present the inequality analyst with some difficulty, since translation may result in some individuals having negative demographically adjusted permanent incomes. This is not however an inherent error of formulation, but rather a consequence of the way we may want to account for differences in family composition. Demographic translating is common practice in the analysis of demand systems (Pollack and Wales, 1992, ch.3), so is the alternative form of demographic scaling. To avoid having to work with negative values of $\hat{\eta}^*_Y$, the researcher may wish to work with the resources scaled using some equivalence scales (for instance the class of Buhmann et al., 1988).

In practice, the introduction of demographic variables suggest a simple procedure for estimating the parameters of (10) and predicting permanent incomes:

1) regress y_{ij} on D_{ij} for each of the p equations $(j = 1, \ldots, p)$

2) fit the model (25) to the $p \times p$ covariance matrix of the residuals with element $s_{jk} = \frac{1}{n}\sum_{i=1}^{n} \hat{v}_{ij}\hat{v}_{ik}$.

3) predict η_i using $b'\hat{v}_i$ where $\hat{v}_i = [\hat{v}_{i1}, \ldots, \hat{v}_{ip}]'$ is the vector of residuals for household i.

Thus, with the augmented model, the vector of residuals $\hat{v}_i$ performs the role of Y_i in the simpler model of factor analysis (9).

3.2. Predicting permanent income using multiple indicators and multiple causes

Building on Muellbauer's (1983) study of living standards in Sri Lanka, Abul Naga and Burgess (1997) propose to predict permanent incomes using the Multiple Indicators and Multiple Causes (MIMIC) model of Zellner (1970) and Jöreskog and Goldberger (1975). The MIMIC system can be seen as a generalization of the factor analysis framework, where together with the multiple indicators Y_i, it is assumed that the researcher possesses data Z on a set of determinants of η_i. The MIMIC system thus consists of (9) together with an equation for the multiple causes of the household's unobserved permanent income

$$\eta_i = \gamma' Z_i + \varepsilon_i \tag{29}$$

As the MIMIC model is identifiable with $p \geq 2$ indicators and $k \geq 1$ causes, in what follows, we shall examine the two-indicator model, where $Y_i = [y_{i1}; y_{i2}]'$ are respectively the income and consumption expenditure of the household. The MIMIC system can in fact be treated as a general simultaneous equations model where Y_i depends on η_i, and in turn η_i is a function of Z_i (a set of predetermined variables such as the human capital endowment and the stock of productive assets of the household). For the two-indicator case, the MIMIC model entails the following reduced form

$$y_{i1} = \gamma' Z_i + \varepsilon_i + u_{i1} \tag{30}$$

$$y_{i2} = \beta\gamma' Z_i + \beta\varepsilon_i + u_{i2} \tag{31}$$

which is obtained upon substituting $\gamma' Z_i + \varepsilon_i$ for η_i in the multiple indicators system (9). It is assumed that Z_i is orthogonal to ε_i and U_i, and also that $E(U_i \varepsilon_i) = 0$. If a given component z_{ij} $(j = 1; ...k)$ of Z_i were correlated with either of the errors ε_i or u_{ij}, then effectively it would be an endogenous variable, that is an indicator rather than a cause of η_i. Once again, to be ensured (at least to some extent) against this sort of specification error, we recommend testing the specification after estimation (see sub-section 3.3. below).

Letting Σ_M denote the covariance matrix of Y_i, we have

$$\Sigma_M = \beta\beta'\gamma' E(ZZ')\gamma + \sigma_{\varepsilon\varepsilon}\beta\beta' + \Omega \tag{32}$$

where it may be noted again, as in the multiple indicator framework augmented for the demographics, that the reduced form error vector $V_i = \beta\varepsilon_i + U_i$ has a factor analytic structure. In the absence of Z variables the component $\beta\beta'\gamma' E(ZZ')\gamma$ vanishes, and (29) reduces to the covariance structure (10) of the model of factor analysis.

Prediction

Unlike in the multiple indicators framework, the MIMIC model offers three possible routes to predicting permanent income:

(i) a predictor that uses the indicators Y, referred to as the Y-predictor

(ii) a predictor which uses the predetermined variables Z (the Z-predictor)

(iii) a predictor which combines the information contained in Y and Z. Let $W = \begin{pmatrix} Y \\ Z \end{pmatrix}$; the third predictor will be referred as to the W-predictor. Section 3 of Abul Naga and Burgess (and the appendix to their paper) discusses in some length prediction in the MIMIC model. The Y-predictor of the MIMIC model is the minimum MSE predictor of η in the class (18). It takes the form

$$\eta^*_Y = \left(\beta'\Sigma_M^{-1}\beta\right)^{-1}\beta'\Sigma_M^{-1}Y \tag{33}$$

which has a similar generalized least squares form to the predictor of the model of factor analysis. The predictors however differ in that the covariance structures Σ_M and Σ_F underlying the two models are distinct.

The Z-predictor is simply the regression function of η on Z:

$$\eta^*_Z = E(\eta|Z) = \gamma'Z \tag{34}$$

The predictor η^*_Y may be seen to be unbiased in the sense that

$$E[\eta^*_Y|Z] = E\left[\left(\beta'\Sigma_M^{-1}\beta\right)^{-1}\beta'\Sigma_M^{-1}(\beta\gamma'Z+\beta\varepsilon+U)|Z\right] = \gamma'Z \tag{35}$$

Since $\beta E(\varepsilon|Z) = E(U|Z) = 0$, η^*_Y (and any given member of the class C_Y) is unbiased in the sense that $E(\eta^*_Y|Z) = E(\eta|Z) = \gamma'Z$.

The W-predictor is a linear function $b_Y'Y + b_Z'Z$. That is, for a given predictor $\tilde{\eta}_W$ we would have

$$\tilde{\eta}_W = b_Y'(\beta\gamma'Z + \beta\varepsilon + U) + b_Z'Z \tag{36}$$

Now considers the class of predictors C_W

$$C_W = [b_Y'Y + b_Z'Z \mid b_Y'\beta\gamma' + b_Z' = \gamma'] \tag{37}$$

Then for any member of the class of predictors C_W we would have

$$E[\eta^*_W|Z] = (b_Y'\beta\gamma' + b_Z')Z = \gamma'Z \tag{38}$$

again, since $\beta E(\varepsilon|Z) = E(U|Z) = 0$). Members of the class C_W are therefore unbiased in the sense that $E(\eta^*_W|Z) = E(\eta|Z) = \gamma' Z$.

Abul Naga and Burgess show that the minimum MSE predictor in the class C_W is a weighted sum of η^*_Y and η^*_Z:

$$\eta^*_W = \theta\eta^*_Y + (1 - \theta)\eta^*_Z \text{ where } \theta = \sigma_{\eta\eta}\beta'(\beta\beta'\sigma_{\eta\eta} + \Omega)^{-1}\beta \tag{39}$$

Conditional on Z, the three predictors have therefore the same mean (but different variances). Restricting the predictors to be unbiased is therefore a means of putting some structure, or common ground, between the three approaches[6]. The predictor η^*_W pools the available information in Y and Z in order to predict household permanent income. Abul Naga and Burgess therefore are able to show that η^*_W has a lower mean-square error than η^*_Y and η^*_Z (while the latter two can not generally be ranked). Another way to argue for the superiority of the W-predictor over η^*_Y and η^*_Z is to appeal to the Bayesian concept of sufficiency. Writing the joint density of the variables as

$$f(\eta; Y; Z) = f(\eta|Y; Z) \cdot f(Y; Z) \tag{40}$$

we can use the results of Bartholomew (1987, ch.4) to argue that any sufficient statistic for η (a statistic that exhausts all the available information in the data) must be based on the distribution $f(\eta|Y; Z)$. Predictors which discard information on Z, or on Y, cannot meet the sufficiency requirement. Only a statistic based on Y and Z can be a candidate for sufficiency[7].

Controlling for demographic structure

In his study of living standards in Sri Lanka, Muellbauer (1983) lets his Z vector depend on variables pertaining to household composition, educational and occupational status, the stock of productive assets and, finally, community variables. Abul Naga and Burgess (1997) adopt a similar specification for their causes equation (29) in their study of living standards in rural China. One problem with the MIMIC framework is that the coefficients on all Z variables are constrained to be multiples of one another across the two equations. In other words, the matrix $\beta\gamma'$ of reduced form coefficients is constrained to possess a unit rank. One

[6] The predictors η^*_Y and η^*_Z together with the unobserved variable η also imply an inequality ordering, which is used in a companion paper, Abul Naga and Bolzani (2000), in order to rank Lorenz curves in the presence of measurement error.

[7] The exact form for the family of sufficient statistics depends on the distributional assumptions pertaining to $f(\eta; Y; Z)$.

way of rendering the MIMIC framework more flexible is to allow a subset of the explanatory variables for income and consumption to have unconstrained coefficients across the two equations. Let D_{i1} and D_{i2} denote a set of explanatory variables, typically demographic variables, which need not be subject to cross-equation constraints. Then, a more flexible MIMIC framework could take the form

$$y_{i1} = \gamma' Z_i + \delta_1' D_{i1} + \varepsilon_i + u_{i1} \tag{41}$$

$$y_{i2} = \beta_2 \gamma' Z_i + \delta_2' D_{i2} + \beta \varepsilon_i + u_{i2} \tag{42}$$

where D_{i1} and D_{i2} could be exactly identical, or may possess different (sub)sets of variates. The system reduces to the earlier MIMIC model (30-31) if δ_1 and δ_2 are both zero vectors. Below we shall refer to the system (41-42) as the MIMIC-D, or the demographics-augmented MIMIC model. The covariance structure underlying the MIMIC-D model has the form

$$\Sigma_D = \beta\beta'\gamma' E(ZZ')\gamma + M + \beta\beta'\sigma_{\varepsilon\varepsilon} + \Omega \tag{43}$$

where M is now a 2×2 version of (26). The Y-predictor for the MIMIC-D model is the minimum MSE predictor in the class C_Y' (26), that is

$$\eta_Y^* = \left(\beta'\Sigma_D^{-1}\beta\right)^{-1}\beta'\Sigma_D^{-1}(Y - D\delta) \tag{44}$$

which is identical in form to the predictor of the demographics-augmented multiple indicators framework (see equation 28), with the difference that Σ_D replaces Σ_S. The Z-predictor for the above model does not change. For prediction combining Y and Z variables the class C_W' replaces C_W:

$$C_W' = \{b_Y'(Y - D\delta) + b_Z' \mid b_Y'\beta\gamma' + b_Z' = \gamma'\} \tag{45}$$

for which the minimum MSE predictor takes the form

$$\eta_W^* = \theta\left[\left(\beta'\Sigma_D^{-1}\beta\right)^{-1}\beta'\Sigma_D^{-1}(Y - D\delta)\right] + (1 - \theta)\gamma' Z \tag{46}$$

and θ is as defined in (39).

The case $D_{i1} = D_{i2} = D_i$, where the same demographic variables enter both equations, produces an interesting relation between the MIMIC model augmented with the demographics and the earlier formulation (30-31). This may be shown by appealing to the Frisch-Lovell theorem of econometrics (Davidson and MacKinnon, 1993, pp.19-24).

Define Y_1, D, and Z

$$Y_1 = \begin{bmatrix} y_{11} \\ \vdots \\ y_{i1} \\ \vdots \\ y_{n1} \end{bmatrix} \quad D = \begin{bmatrix} D_1^{\prime} \\ \vdots \\ D_i^{\prime} \\ \vdots \\ D_n^{\prime} \end{bmatrix} \quad Z = \begin{bmatrix} Z_1^{\prime} \\ \vdots \\ Z_i^{\prime} \\ \vdots \\ Z_n^{\prime} \end{bmatrix} \tag{47}$$

as respectively an $n \times 1$ column vector, an $n \times l_1$ matrix and an $n \times k$ matrix. Writing (41) for the n individuals, we have

$$Y_1 = Z\gamma + D\delta_1 + \varepsilon + u_1 \tag{48}$$

where ε and u_1 are the $n \times 1$ vectors of disturbances pertaining to Y_1. Now, defining $Q = I - D(D^{\prime}D)^{-1}D^{\prime}$, it follows that $\tilde{Y}_1 = QY_1$ is the vector of residuals from a regression of Z on D. Note furthermore that $QD = 0$ by definition. Premultiplying (48) by Q we thus have

$$\tilde{Y}_1 = \tilde{Z}\gamma + \tilde{\varepsilon} + \tilde{u}_1 \tag{49}$$

where $\tilde{\varepsilon} = Q_1\varepsilon$ and $\tilde{u}_1 = Q_1u$. Now defining $\tilde{Y}_2 = QY_2$ and $\tilde{u}_2 = Qu_2$ a similar equation may be obtained for consumption expenditure:

$$\tilde{Y}_2 = \tilde{Z}\gamma\beta + \tilde{\varepsilon}\beta + \tilde{u}_2 \tag{50}$$

In the benchmark case where the same set of demographics is used in both equations, the demographics-augmented MIMIC model therefore reduces to the earlier MIMIC framework when y_{i1}, y_{i2} and Z_i are replaced by $\tilde{y}_1$, $\tilde{y}_2$ and $\tilde{Z}_i$, i.e. the residuals from prior regressions of y_{i1}, y_{i2} and Z_i on the demographic control variables. We shall proceed along these lines when we estimate MIMIC models for Swiss household data. Prior to this, however, we shall discuss two further points pertaining to identification and testing of the models considered so far.

3.3. A note on estimation and hypothesis testing

Models with latent variables are usually estimated by minimizing a distance $d[S; \Sigma(\theta)]$ between the sample covariance matrix and the theoretical model $\Sigma(\theta)$, where ϕ is a vector of unknown structural parameters. For instance, in the case of the model of factor analysis $\phi' = \left(\beta_2, \ldots, \beta_p, \omega_{11}, \ldots, \omega_{pp}\right)'$ and $\Sigma(\phi)$ is the matrix Σ_F. As pointed out by Chamberlain (1977), there are usually simpler, more transparent, estimation procedures for such models, based on the methods of instrumental variables (IV).

Consider again the situation where the researcher is working with two indicators y_{i1} and y_{i2} on permanent income η_i:

$$y_{i1} = \eta_i + u_{i1} \tag{51}$$

$$y_{i2} = \beta_2 \eta_i + u_{i2} \tag{52}$$

This is the simplest case to consider, though in what follows, the introduction of demographic variables does not change the nature of the estimation problem. If one were to proxy η_i using y_{i1} in (52), we would have a classical errors in variables problem; namely,

$$y_{i2} = \beta_2 y_{i1} + u_{i2} - \beta_2 u_{i1} \tag{53}$$

Because y_{i1} is correlated with u_{i1}, OLS estimation of (53) yields an inconsistent estimator of β_2. Now consider a situation where a third indicator is available on η_i:

$$y_{i3} = \beta_3 \eta_i + u_{i3} \tag{54}$$

For y_{i3} to be a valid instrument of y_{i1} in (53), it must be uncorrelated with the composite error term $u_{i2} - \beta_2 u_{i1}$. In other words, u_{i3} must be uncorrelated with both u_{i1} and u_{i2}. Under such circumstances an IV estimator for β_2 takes the form:

$$\hat{\beta}_2 = \frac{\sum_{i=1} y_{i2} y_{i3}}{\sum_{i=1} y_{i1} y_{i3}} \tag{55}$$

If Ω is a diagonal matrix one can construct an IV estimator for β_3 along the lines of $\hat{\beta}_2$, in which case it is straightforward to establish that with $p = 3$ indicators a model such as (9) is just-identified, and $p \geq 4$ presents an over-identified system which is testable (see below).

Now consider the alternative case (the MIMIC formulation) where together with (51) and (52), the researcher has a causal equation with a unique Z variate:

$$\eta_i = \gamma_1 z_{i1} + \varepsilon_i \tag{56}$$

with $E(\varepsilon_i | z_{i1}) = 0$. Then equation (53) can be instrumented using z_{i1} to produce an (other) IV estimator of β_2:

$$\tilde{\beta}_2 = \frac{\sum_{i=1} y_{i2} z_{i1}}{\sum_{i=1} y_{i1} z_{i1}} \tag{57}$$

while γ_1 may be consistently estimated via an OLS regression of y_{i1} on z_{i1}. Note again that identification will require that z_{i1} be uncorrelated with u_{i1} and u_{i2}.

The case $Z_i = [z_{i1}; \ldots; z_{ik}]$ presents an over-identified system (with $k - 1$ extra instruments) which may provide a basis for testing the underlying specification. Tests of over-identifying restrictions are well known in the econometrics literature; see for instance Godfrey (1988, pp. 168-174). In particular, they may be used to test whether an indicator y_{ij} in (55) is a valid instrument, or alternatively if some z_{il} in (57) can be used in the MIMIC context.

A general specification test which will be used in our empirical applications below is $(n-1) \cdot d\left[S; \Sigma\left(\hat{\theta}_{ML}\right)\right]$ where $\hat{\theta}_{ML}$ is the maximum likelihood estimator, and $d[S; \Sigma]$ is the formula (3.28) in Bartholomew (1987, p. 50). Then, if θ is a τ-dimensional vector, and there are q distinct sample moments (for example $q = \frac{p(p+1)}{2}$ and $\tau = 2p$ in the model factor analysis), then

$$(n-1)\, d\left[S; \Sigma\left(\hat{\theta}_{ML}\right)\right] \sim \chi^2_{q-\tau} \tag{58}$$

The null hypothesis that the model is correct is therefore rejected on a basis of a χ^2 test with increasing values of $(n-1)\, d\left[S; \Sigma\left(\hat{\theta}_{ML}\right)\right]$ being indicative of a poor fit for the data.

4. Data

In order to illustrate the various methods described above, we shall look at Swiss household data pertaining to the Enquête sur la Consommation des Ménages (ECM90) of 1990. ECM90 is a survey on the population residing in Switzerland on a permanent basis, conducted by the Swiss Federal Statistical Office[8]. Accordingly, those on temporary stays for visiting or employment purposes are not sampled. The household head, defined as the prime income earner, is asked to provide information on her/his sex, citizenship, education, marital and employment status. Of the 9000 initially selected households only 1994 families (that is 22%) agreed to take part in the survey. In order to ensure population representativeness, in particular with respect to the linguistic area, the data have been weighted using sample weights provided by the survey authorities. A related survey, recording the detailed demand patterns of a sample of families, was also conducted by the Federal Statistical Office during the same year.

Of the 1994 observations we have had to discard 321 records because of incompleteness or missreporting. There are 31 observations for families who simultaneously report moderate to high levels of consumption expenditure and substantially large negative incomes. Furthermore, the survey did not record information on the educational status of most households in retirement. On such grounds, we have lost another 290 observations. This is quite unfortunate since our results will not take into account the levels of living of the elderly population whose circumstances may vary a great deal. We were thus left with a total of 1673 observations.

Table 1 summarizes our data in terms of annual family income, total consumption expenditure as well as other demographic variables pertaining to the education and labor market status of the main income earner. Mean household income was equal to 59'000 Swiss Francs (CHF) while the average level of consumption expenditure amounted to 56'200 CHF. Consumption expenditure was however more equally distributed than household income, in accordance with many economic theories of consumption. For instance, the coefficient of variation (the standard deviation divided by the mean) for expenditure was equal to 0.457, whereas it amounted to a 0.526 figure for the latter variable. Regarding occupational status, 81.5% of our family heads were employees while 8.7% were independent workers (the remaining groups were divided between agricultural employment, 5.9%, retirement, 2.2%, and a residual 1% who were unemployed). 16% of our households are female headed and 78% reside in the German speaking area of the country. We note finally that foreign citizens may appear to be in some respect under-represented in our sample. We have 90% of Swiss household heads, when the nationwide figure for the foreign population at the time of the survey was in the order of 15%.

In order to summarize the essential features of our data, we run several re-

[8] For details, refer to "Office Fédéral de la Statistique", 1992.

gressions of the logarithm of household income on various socioeconomic characteristics of the family unit (see Table 2). Regression 1 is a standard speci…cation of the returns to education controlling for the age-income pro…le. The return to education is in the order of 9 to 10%. In regression 2 we further add variables pertaining to the sex, marital status and citizenship of the family head. There is an apparently negative premium associated to being a Swiss citizen. However, as pointed out by Leu, Buhmann and Frey (1986), when di¤erences in the socioeconomic characteristics of the Swiss and foreign populations are not controlled for, the group of foreigners may appear to enjoy higher levels of living than the Swiss population. On the other hand, foreign households are larger in size and have more individuals in employment (and they are also younger). For instance, we may note that in regression 5 the dummy on Swiss citizenship, while remaining negative, is no longer signi…cant once we adequately control for the number of workers and those in full-time employment. Finally, observe that the inclusion of variables pertaining to labor market status generally improves the explanatory power of the regressions. The adjusted R^2 for regressions I to III is in the order of 0.09, while for the latter two speci…cations R^2 is approximately 0.40.

5. Results

In this section, we provide applications of the multiple indicators and MIMIC frameworks on the sample of Swiss households described in the previous section. We then go on to compare the distributions of predicted permanent incomes for independent and employed workers.

5.1. Identifying the poor: multiple indicators

Referring back to our methodological section, we have estimated factor analytic speci…cations controlling for a string of demographic variables. These variables included age and age squared of the household head, household size, the number of children under the age of 10, the number of workers in the family, and dummy variables for full-time employment, self-employment, agricultural employment and whether the household head is in retirement.

Table 3 presents parameter estimates for three distinct speci…cations. The coe¢cient ¯$_j$ is the slope estimate pertaining to the jth indicator of ´. For example, ¯$_1$ is restricted to equal unity in all three speci…cations while Rcs, the coe¢cient of consumption expenditure, is estimated at 0.78 in the …rst application and at 0.77 for the next two models. The coefficient ω_{jj} is the estimated transitory variance pertaining to indicator j. Note from the estimates of the various models that the transitory variance of consumption expenditure is estimated at 0.05, while that of income is in the order of 0.15. The noise to total variance ratio, defined as $\omega_{jj}/(\beta_j^2\sigma_{\xi\xi} + \omega_{jj})$, takes a value of about 0.56 for income and 0.39 for consumption expenditure. Two remarks may be put forward on the basis of these findings. Firstly, there is a substantial amount of variation in income and consumption which remains unaccounted for even after we control for household characteristics. Secondly, the residual variation in consumption expenditure is smaller than in the case of the former indicator. In fact, it is the smallest for all indicators considered in Table 3. We may note in passing that of the remaining three indicators, the residual variation for education (as measured by the noise to total variance ratio) is in the order of 0.82 while the Swiss citizenship and male dummies have virtually nil explanatory power.

One natural question to ask is how plausible are these estimated covariance structures. As we are working with more than three indicators, the above models are overidentified, and hence testable. We may note therefore that model I has a lower Khi-square statistic than model II, and a P-value of 0.40 in comparison to 0.07 for the latter. The P-value for model III is equal to 0.02. On such grounds, this latter specification falls in the 5% critical region and accordingly must be rejected.

We may at this stage inquire about the consequences of controlling for household demographic composition. The Khi-square statistics corresponding to models I, II and III without demographic controls are respectively 8.89 (with a P value of 0.01), 20.8 (P = 0.00) and 9.38 (P = 0.01). In other words, the introduction of demographic controls adds realism to the model specifications, and must be recommended in empirical work. Another benefit from the introduction of demographic controls may be noted by examining the coefficient on the Swiss citizenship dummy. Unlike in the earlier regressions of Table 2, it may be noted that this coefficient is now positive (though only significant at the 10% level). Furthermore, in the factor analyses without demographic controls (results not shown) the coefficient on the Swiss dummy is negative for both models where it appears, and significant at the 95% level for the specification corresponding to model I of Table 3.

Turning now to the prediction results, we may observe that income and consumption expenditure are assigned the highest weights whereas the roles of the remaining variables are largely residual. As a general rule, consumption expenditure is assigned a weight of about 0.77, while income is allocated a weight of about 0.31. These results echo our earlier discussion of the residual variation pertaining to the various indicators. As was shown in equation (22b), the prediction weight given to indicator j is proportional to the ratio β_j / ω_{jj}. It comes as no surprise therefore that consumption expenditure receives the largest weight. Empirically therefore, our predictor of permanent income may be constructed from the data as $\eta_Y^* = 0.307y_1 + 0.775y_2 + 0.046y_3 + 0.022y_4$ for the variables pertaining to model I, and likewise for the remaining two specifications. As comes out from the theory of intertemporal allocation of consumption over the life-cycle, consumption expenditure comes closest to a definition of permanent income. However, our results suggest that total family income has also a non-negligeable informational content in predicting household permanent income.

As a final exercise, we classify households according to their consumption expenditures, incomes and their predicted permanent incomes. The income and consumption variables are the residuals from the initial regressions on household demographic variables. These data are summarized in terms of classification matrices in Tables 4a (for consumption and income), 4b (for income and predicted permanent income) and 4c (for consumption expenditure and η_Y^*). Data are ranked from the lowest (Q1) to the highest (Q5) quintile. For instance, in Table 4a there are 186 families falling simultaneously in the bottom income and consumption quintiles (the element Q_{11}) against 9 families in the bottom consumption quintile and in the top income quintile (the element Q_{15}).

Note first that in the case of the classification matrix between consumption and income, the main diagonal only contains 698 families out of the 1673 observations, that is 41.7% of the sum total. This is not the case for the classification matrix of table 4b, where income and the multiple indicator index rank 874 observations (that is 52.2% of all cases) along the main diagonal. It may further be noted that the off-diagonal cells become less populated as we move away from the main diagonal. For instance, there are 12 families falling simultaneously in the highest consumption and lowest income quintile, in comparison to 0 cases in the corresponding cell of table 4b. If we set the poverty line at the lowest income quintile, then by definition a given indicator identifies 335 households as being in poverty. For income and consumption, agreement is reached in 186 cases. Thus, in 55.5% of cases these two indicators identify the same families as being in poverty. In the case of income and predicted permanent income, agreement is now possible for 236 families, that is 70% of all possible cases. This pattern is further intensified when one compares the results of Tables 4a and 4c. The main diagonal of Table 4c (where consumption expenditure is classified along the line and the multiple indicator index along the column) contains now 1109 observations (66% of all cases). Note furthermore that agreement on the poor population is now reached in 269 cases, in comparison to the 236 families of table 4b and the 186 cases of Table 4a.

The results of the exercise of identifying the poor on the basis of the classification matrices of Tables 4a to 4c are summarized in Figure 1 by means of a Venn Diagram. When the poverty line is set at the bottom quintile of a given standard, each indicator will identify 335 families as being in poverty. As shown in the diagram, income, expenditure and the multiple indicator index jointly identify 181 families as being in poverty. The latter predictor commonly identifies with income another 55 observations. Together with consumption expenditure, η^*_Y ranks another 88 families as being in the bottom 20% group. On the other hand, only 5 households are jointly identified by income and consumption as being in poverty, but not by the multiple indicator index. Furthermore, we note that the permanent income measure has the largest intersection with the income and expenditure sets, with the residual 10 families identified as being in poverty by the predictor but not by the other two variables.

5.2. Mimic estimations

Next turn to the estimation of the MIMIC model for family income and consumption expenditure. The same set of demographics as in our earlier applications are used here to control for family composition. On the other hand, some variables are used to model the determination of household permanent income. In terms of the notation of our methodology section, these were the Z variates. We consider three distinct specifications for the joint model of income and consumption, results of which are reported in Table 5. The estimate of β_2 is approximately 0.76, which is in broad agreement with the estimates of Table 3. The return to schooling is estimated at 10% (the γ_1 coefficient)[9]. The coefficient on the Swiss citizenship dummy is positive though it is not statistically different from zero. Living in the Germanic part of Switzerland is associated with a positive and statistically significant premium as shown by the coefficient γ_3 on the variable Rger (columns 2 and 3). The premium associated with being a male household head is also positive

[9] Note that we had discarded 321 observations from the initial survey because of missing values on education. If the income-schooling relation amongst these missing values deferred from that of the sample at large, our resulting estimates, by failing to control for sample selection, may result in positive or negative biases.

though failing to be significant at the 10% level. Finally, our model III estimates suggest that house ownership (Rprop) has a positive effect on permanent income.

The residual variation as measured by the noise to total variance ratio is estimated to be in the order of 0.56 for income and 0.40 for expenditure. These results are broadly the same as our earlier findings pertaining to the models of Table 3. All three specifications failed to be rejected using the Khi-square specification tests. P-values for models I and II are respectively 0.22 and 0.49, while for model III this figure falls to 0.053. Overall then, specifications I and II adequately explain our data, whereas model III only fails to be rejected at the 5% level and is thus a borderline case.

Unlike in the multiple indicator framework, the MIMIC approach provides three different routes to the problem of predicting household permanent income. Permanent income may be predicted using left-hand variables as in the case of the model of factor analysis, using the determinants of permanent income (the Z variables), and finally using a combination of Y and Z variables which are denoted in Tables 6 as W. Regarding the Y-predictor, a similar pattern emerges as in our earlier results of Table 3: for models I and II, a predictor of permanent income attributes weights to consumption expenditure and income in the ratio of approximately 2.5:1. For model III, this ratio is equal to 2.15:1. The coefficients on the Z-predictor are simply the γ estimates of Table 5. The W-predictor, on the other hand, is a weighted sum of the former two predictors. It is for this reason that the weights on income and consumption expenditure in the W-predictor are exactly in the same ratio as in the Y-predictor. As argued in our methodology section, the W-predictor is to be preferred over the other two predictors on grounds of both mean square error and sufficiency. In results not shown, we repeat our earlier exercise of identifying the poor with the difference this time that the W-predictor replaces the former multiple indicator index. The classification matrices are very similar to the corresponding ones of Table 4 and the Venn Diagram of Figure 2 illustrates the same point, namely that a multi-dimensional approach establishes a greater consensus in identifying the poor.

5.3. Comparing the situations of employees and the self-employed

In order to illustrate what further insights may be obtained from a multi-dimensional approach to the identification of the poor, we consider a comparison of the distributions of employees and the self-employed. In figure 3 we plot the deficit curves for the incomes of these two groups. There, we notice that the deficit curve of the self-employed lies everywhere above that of the other group. On the other hand, the curves are in exactly reverse order when household consumption expenditure is used as an indicator of welfare (see …gure 4). That is, according to the income de…nition poverty is lower amongst employees at all poverty lines. This result holds true for all members of the Atkinson class of concave poverty measures (Atkinson, 1989; ch.2). Income and expenditure in …gures 3 and 4 have been adjusted by the square root of family size, which is a type of equivalence scale suggested by Buhmann et al. (1988). The same conclusion arises regarding the poverty ordering when total family income and expenditure are used without adjustment for family size (results not shown).

This result is to some extent disturbing since the conclusions reached about poverty amongst these two groups are not robust to the choice of welfare indicator. If independent workers include professionals with high level of training together with semi-skilled manual workers, while the group of employees is more homogeneous in this respect, we could expect the de…cit curves of these two distributions to cross for both indicators. In practice, the validity of this story depends on the extent to which consumption expenditure tracks observed household income.

In order to shed light on this problem, we propose to examine these two groups in terms of their distributions of predicted permanent incomes. For the purpose of measuring permanent income in the same scale as annual income, we estimate a MIMIC-D speci…cation (equations 41-42). Estimation results are not reported here, as they di¤er from those of table 5 only moderately. The use of the W-predictor (…gure 5) depicts a scenario of crossing poverty de…cit curves. When predicting permanent incomes by means of $´^{¤}_{W}$, we …nd that the de…cit curve of the self-employed initially lies above that of the other group. Our results would therefore suggest that, at the bottom of the distribution there exists a range of poverty lines such that the self-employed are poorer than employees, but that the ranking of these two groups changes as we consider higher levels of the poverty line. Note …nally that in …gure 5 the lower range of the horizontal axis contains some negative values for $´^{¤}_{W}$. As discussed in our methodology section, this pattern re‡ects the translation of income and consumption expenditure arising from the adjustment for demographic variation.

6. Conclusion

It is a well documented fact that the set of households which are identi…ed as being poor is not invariant to the choice of welfare indicator. If income and consumption expenditure do not convey the same information on an unobserved variable on the basis of which families plan their consumption over time, there is a rationale for combining these two (and other) welfare indicators in order to identify the poor. The multiple indicators and MIMIC approaches are statistical: starting from a set of welfare indicators/determinants of long-run status, we have constructed our various predictors of permanent income as linear combinations of the data, chosen to minimize criteria based on mean-square (prediction) error.

The methods proposed here may also be applied for the study of inequality of permanent incomes. In a companion paper, Abul Naga and Bolzani (2000), we show how the multiple indicators and multiple causes approach may be used in order to provide upper and lower bounds for the Lorenz curve pertaining to the unobserved household permanent incomes. Similar results are also available for the related generalized Lorenz curve used to rank distributions in terms of welfare.

We have used Swiss household data in order to illustrate the potential insights the data analyst may gain from the proposed methodology. When income and consumption expenditure produce conflicting evidence regarding the poverty status of groups of individuals, multiple indicators and MIMIC approaches offer more scope for agreement with these two welfare standards. Nonetheless, there are also problems yet to be resolved. Working for instance with an absolute rather than relative poverty threshold may be problematic if the underlying variable is a permanent income concept. The fact that a portion of the distribution of predicted permanent incomes lies in the negative range of the real line may be dealt with by scaling rather than translating welfare indicators when controlling for household demographic structure. Yet, we believe that the problem of defining an absolute poverty line for a weighted sum of income and expenditure (as in the case of the Y-predictor) is not such a straightforward task. At this stage, then, it is perhaps more reasonable to think of these approaches as being related to relative concepts of poverty, where it is agreed at the outset that the poverty line is a fraction of the median resource level, or some quantile of the distribution as in the case of our empirical applications.

As shown by means of our intertemporal resource allocation problem of section 2, the linear relation between consumption expenditure and permanent income only obtains in the specific case of quadratic utility. We could decide to do away with uncertainty about future income, in which case a relation between consumption and permanent income may be obtained for a wider class of problems. But how reasonable would it be to assume that households, especially the poorer ones, do not save for precautionary purposes? Consider another problem distinct from that of uncertainty - borrowing constraints on the credit market. If there exist limitations to borrowing, then the consumption expenditure of rationed families will equal their current income. In this respect, an econometric framework where current income and expenditure only correlate via their common dependence on permanent income, such as the one we have considered, is in some way restrictive. More ‡exible frameworks for cross-section data, allowing current income and expenditure to correlate more freely, would add realism to the methods proposed here, especially if one is interested in drawing inferences about the welfare of the poorer groups, who may have limited means of re-allocating their consumption over time and states of nature.

7. References

Abul Naga, R. (1994): "Identifying the Poor: A Multiple Indicator Approach", Distributional Analysis Research Program Discussion paper No. 9, LSE.

Abul Naga, R. and Bolzani, E. (2000): "Lorenz Curve Analysis for Household Permanent Incomes", work in progress, University of Lausanne.

Abul Naga, R. and Burgess, R. (1997): "Prediction and Determination of Household Permanent Income", Discussion Paper, DEEP, University of Lausanne.

Anand, S. and Harris, C. (1990): "Food and Standard of Living: An Analysis based on Sri Lankan Data", in Drèze, J. and Sen, A. (eds.): The Political Economy of Hunger, Vol.1, Oxford University Press, Oxford.

Atkinson, A. (1989): Poverty and Social Security, Hemel Hempstead, Harvester Wheatsheaf.

Bartholomew, D. (1987): Latent Variable Models and Factor Analysis, New York, Oxford University Press.

Bartlett, M.S. (1937): "The Statistical Conception of Mental Factors", British Journal of Psychology, 28, 97-104.

Blundell, R. and Preston, I. (1998): "Consumption Inequality and Income Uncertainty", Quarterly Journal of Economics, 113, 603-640.

Buhmann, B., Rainwater, L., Schmaus, G. and T.M. Smeeding (1988): "Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study (LIS) Database", Review of Income and Wealth, 34, 115-142.

Chamberlain, G. (1977): "An Instrumental Variables Interpretation of Identi…cation in Variance-Components and MIMIC models", in P. Taubman ed.: Kinometrics, Amsterdam, North Holland.

Chauduri, S. and Ravaillon, M. (1994): "How Well Do the Static Indicators Identify the Chronically Poor?", Journal of Public Economics, 53, pp.367-394.

Davidson, R. and MacKinnon, J. (1992): Estimation and Inference in Econometrics, New York, Oxford University Press.

Deaton, A. and Muellbauer, J. (1980): Economics and Consumer Behavior, Cambridge, Cambridge University Press.

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

Friedman, M. and Kuznets, S. (1945): Income from Independent Professional Practice, New York, National Bureau of Economic Research.

Glewwe, P. and Van der Gaag, J. (1990): Identifying the Poor in Developing Countries: Do Di¤erent De…nitions Matter?", Word Development, 18, 803-814.

Godfrey, L. (1988): Misspeci…cation Tests in Econometrics, Cambridge, Cambridge University Press.

Gorman, W. (1959): "Professor Friedman's Consumption Function and the Theory of Choice", reprinted in Blackorby, C. and Shorrocks, A. eds (1995): Separability and Aggregation: The Collected Works of W.M. Gorman, Vol I, Oxford, Clarendon Press.

Hall, R. and Mishkin, F. (1982): "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households", Econometrica, 50, 461-481.

Jöreskog, K. and Goldberger, A. (1975): "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable", Journal of the American Statistical Association, 70, 631.639.

Lawley, D. and Maxwell, A. (1971): Factor Analysis as a Statistical Method, London, Butterworth.

Leu, R., Buhmann, B. and Frey, R. (1986): "Die personelle Einkommens- und Vermögensverteilung der Schweiz", Revue Suisse d'Economie Politique et de Statistique, 122/2, 111-142.

O¢ce Fédéral de la Statistique (1992): "Enquête sur la Consommation 1990: les Dépenses et les Revenus des Ménages Privés", Bern, 1992.

Mercader-Prats, M. (1997): "Identifying Low Standards of Living: Evidence from Spain", Research on Income Inequality, 8, 155-173.

Muellbauer, J. (1983): "The Measurement of Long-Run Living Standards with a Permanent Income Model", mimeo Nu¢eld College, Oxford.

Pollak, R. and Wales, T. (1992): Demand System Speci…cation and Estimation, Oxford, Oxford University Press.

Slesnick, D. (1998): "Empirical Approaches to the Measurement of Welfare", Journal of Economic Literature, 36, 2108-2165.

Tobin, J. (1970): "On Limiting the Domain of Inequality", Journal of Law and Economics, 13, 263-277.

Zellner, A. (1962): "An E¢cient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias", Journal of the American Statistical Association, 70, 661-639.

Zellner, A. (1970): "Estimation of Regression Relationships Containing Unobserved Variables", International Economic Review, 11, 441-454.

Tables

Table 1 – Variables Description

Variable	Description	Mean	Std. Dev.
Cs	Consumption expenditure of the household	56263.870	25685.701
Inc	Total income of the household	59037.770	31034.075
Empl	Dummy for employed worker	0.815	0.388
Ind	Dummy for independent worker	0.087	0.282
Nchild	Number of children under 10	0.517	0.898
Hsize	Household size	2.705	1.332
Ger	Dummy for residence in Germanic area	0.784	0.412
Age	Age of family head	41.348	11.698
Male	Dummy for male householder	0.839	0.367
Swiss	Dummy for Swiss household head	0.895	0.307
Edu	Education	2.711	1.559

Notes: 1. Data are weighted for sample representativeness.
2. For the education variable the following scale is in use: Edu = 1 if person has no more education than compulsory schooling; Edu = 2 when individual has completed an apprenticeship or vocational training; Edu = 3 if the family head has completed secondary education; Edu = 4 if the householder has an apprenticeship and further educational training or specialization; Edu = 5 when one has higher education other than university; Edu = 6 for a university degree (or equivalent).

Table 2 – Estimation Results for Some Regressions [Dependent Variable: ln(y)]

Variables	**I**	**II**	**III**	**IV**	**V**
C_eduhd	0.101 (10.167)	0.099 (9.956)	0.099 (9.840)	0.103 (12.584)	0.092 (11.311)
Age	0.043 (4.933)	0.034 (3.760)	0.030 (3.238)	0.036 (4.829)	0.024 (3.140)
Age^2/100	-0.050 (-5.193)	-0.041 (-4.159)	-0.038 (-3.276)	-0.038 (-4.531)	-0.028 (-3.246)
Ind				-0.658 (-14.560)	-0.658 (-14.831)
Agr				-1.142 (-20.65)	-1.338 (-23.034)
Nworkers				0.160 (10.823)	0.113 (7.072)
Ft				0.286 (6.445)	0.226 (5.011)
Nchild			-0.083 (-3.438)		-0.133 (-6.454)
Prop			0.016 (0.440)		0.117 (3.974)
Hsize			0.044 (2.185)		0.094 (5.695)
Ger			0.019 (0.504)		0.084 (2.771)
Male		0.085 (1.716)	0.074 (1.486)		0.075 (1.818)
Mar		0.096 (2.361)	0.082 (1.685)		0.027 (0.666)
Swiss		-0.152 (-3.037)	-0.156 (-3.072)		-0.024 (-.586)
Adjusted R^2	0.080	0.093	0.098	0.384	0.414

Notes: 1. t values in parentheses
2. n = 1673
3. Agr is a dummy taking a unit value if the householder is a farmer
4. Ft is a dummy for full-time worker
5. Prop is a dummy for houseowner-occupancy
6. Mar is a dummy for married household heads

Table 3 – Parameter Estimates of the Factor Analysis Model

	I			II			III		
Variable	β_i	ω_{ii}	**Predict**	β_i	ω_{ii}	**Predict**	β_i	ω_{ii}	**Predict**
Rinc	1.000	0.155 (14.574)	**0.307**	1.000	0.154 (14.579)	**0.319**	1.000	0.154 (14.453)	**0.322**
Rcs	0.782 (12.393)	0.048 (8.261)	**0.775**	0.771 (12.672)	0.050 (8.721)	**0.758**	0.767 (12.633)	0.050 (8.821)	**0.760**
Redu	1.856 (13.063)	1.913 (26.074)	**0.046**	1.869 (13.146)	1.888 (25.863)	**0.049**	1.864 (13.112)	1.888 (25.844)	**0.049**
Rswiss	0.043 (1.700)	0.092 (28.854)	**0.022**	0.042 (1.679)	0.092 (28.851)	**0.023**			
Rmale				0.085 (3.382)	0.088 (28.657)	**0.048**	0.085 (3.394)	0.088 (28.658)	**0.048**
$\sigma_{\eta\eta}$	0.120 (9.926)			0.122 (10.095)			0.122 (10.091)		
χ^2 **test** **[P-value]**	1.81 [0.404]			10.05 [0.074]			8.05 [0.018]		

Notes: 1. t values appear inside parentheses
2. All variables are constructed as residuals from prior regressions on a string of household demographics (described in the text).

Table 4a – Classification Matrix for Consumption and Income

	Q1	Q2	Q3	Q4	Q5
Q1	186	78	34	28	9
Q2	78	119	67	47	23
Q3	40	77	107	64	47
Q4	19	45	100	101	70
Q5	12	16	26	95	185

Notes: 1. Data are ranked from the lowest (Q1) to the highest (Q5) quintile.
2. n_{ij} is the number of observations that fall in the i^{th} quintile of consumption and j^{th} quintile of income.
3. All variables are constructed as residuals from prior regressions on a string of household demographics (described in the text).

Table 4b – Classification Matrix for Income and the Multiple Indicator Index

	Q1	Q2	Q3	Q4	Q5
Q1	236	71	20	5	3
Q2	75	145	76	32	7
Q3	18	75	136	95	10
Q4	5	38	81	127	84
Q5	0	6	22	76	230

Notes: 1. Data are ranked from the lowest (Q1) to the highest (Q5) quintile.
2. n_{ij} is the number of observations that fall in the i^{th} quintile of income and j^{th} quintile according to the multiple indicator index.
3. All variables are constructed as residuals from prior regressions on a string of household demographics (described in the text).

Table 4c – Classification Matrix for Consumption and the Multiple Indicator Index

	Q1	Q2	Q3	Q4	Q5
Q1	269	58	7	1	0
Q2	50	210	64	10	0
Q3	13	60	176	80	6
Q4	2	5	86	184	58
Q5	0	2	2	60	270

Notes: 1. Data are ranked from the lowest (Q1) to the highest (Q5) quintile.
2. n_{ij} is the number of observations that fall in the i^{th} quintile of consumption and j^{th} quintile according to the multiple indicator index.
3. All variables are constructed as residuals from prior regressions on a string of household demographics (described in the text).

Table 5 – Estimation Results for the Income/Consumption MIMIC Model

Variable	Coefficient	I	II	III
Rinc	β_1	1.000	1.000	1.000
Rcs	β_2	0.783 (12.360)	0.774 (12.782)	0.746 (13.117)
Redu	γ_1	0.095 (11.831)	0.096 (12.057)	0.096 (12.179)
Rswiss	γ_2	0.038 (1.204)	0.024 (0.772)	0.012 (0.382)
Rger	γ_3		0.082 (3.444)	0.083 (3.410)
Rmale	γ_4		0.050 (1.553)	0.053 (1.615)
Rprop	γ_5			0.068 (2.892)
	ω_{11}	0.156	0.155	0.150
	ω_{22}	0.048	0.049	0.052
	$\sigma_{\varepsilon\varepsilon}$	0.099	0.098	0.101
χ^2 test **P value**	χ^2	1.51 [0.219]	2.40 [0.493]	9.35 [0.053]

Notes: 1. t values appear inside parentheses
2. All variables are constructed as residuals from prior regressions on a string of household demographics (described in the text).

Table 6 –Prediction Results for the Income/Consumption MIMIC Model

	I			II			III		
Variable	**Y-pred**	**Z-pred**	**W-pred**	**Y-pred**	**Z-pred**	**W-pred**	**Y-pred**	**Z-pred**	**W-pred**
Rinc	0.334		0.219	0.345		0.223	0.384		0.245
Rcs	0.850		0.557	0.846		0.547	0.826		0.526
Redu		0.095	0.033		0.096	0.034		0.096	0.035
Rswiss		0.038	0.013		0.024	0.009		0.012	0.004
Rger					0.082	0.029		0.083	0.030
Rmale					0.050	0.018		0.053	0.019
Rprop								0.068	0.025

Note: 1. All variables are constructed as residuals from prior regressions on a string of household demographics (described in the text).

Figures

Figure 1 – Identifying the Poor Using Income, Consumption Expenditure and the Multiple Indicator Index: Venn Diagram

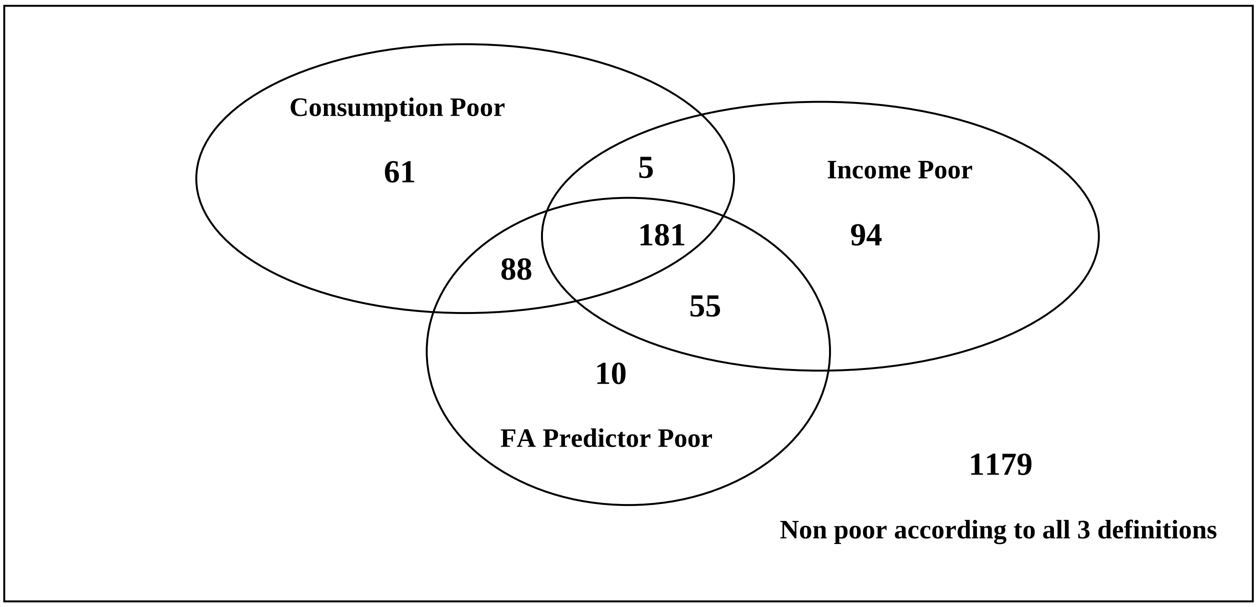
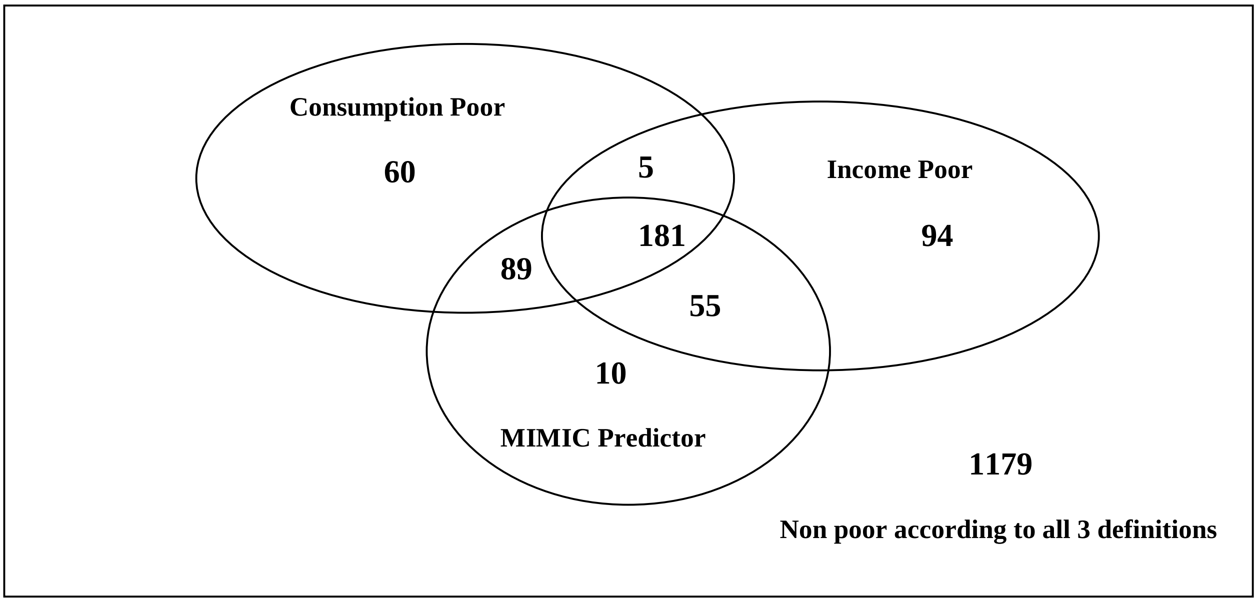


Figure 2 – Identifying the Poor Using Income, Consumption Expenditure and the MIMIC Predictor: Venn Diagram



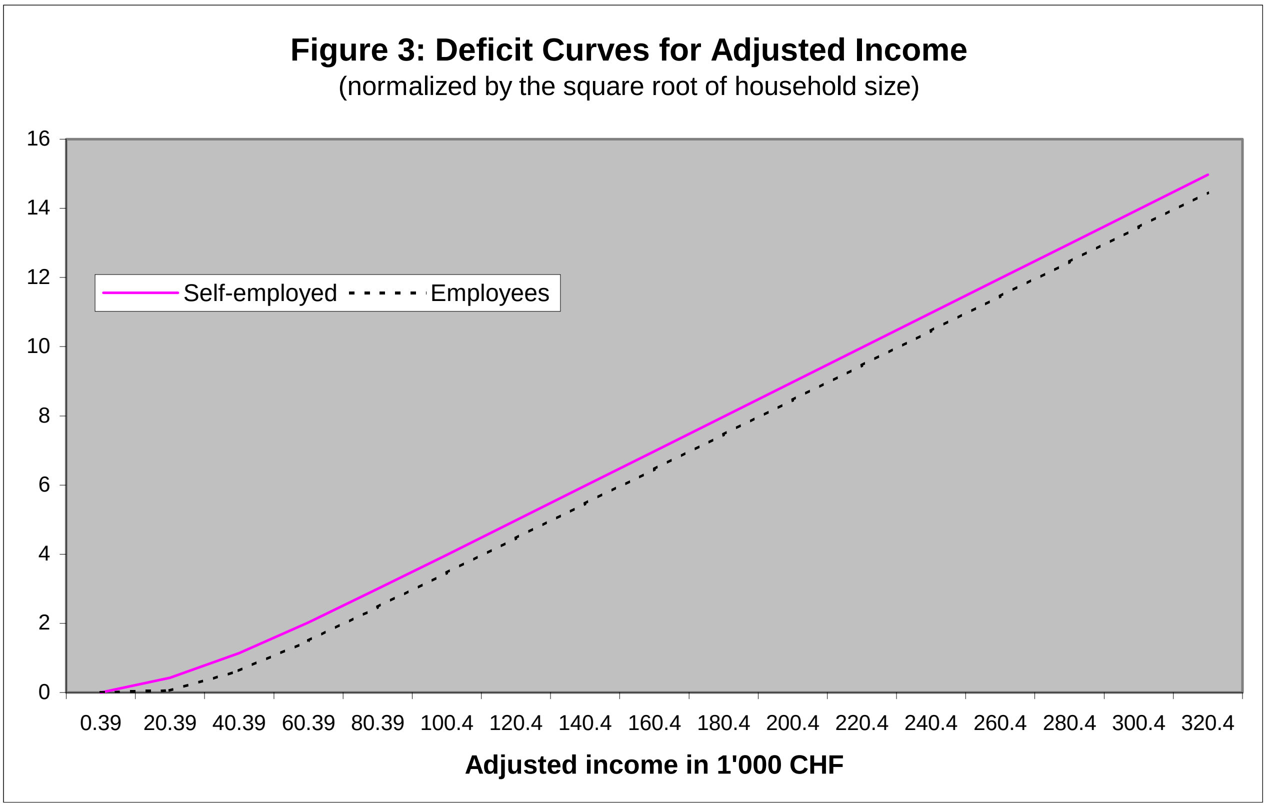


Figure 3: Deficit Curves for Adjusted Income
(normalized by the square root of household size)
Self-employed
Employees
0
2
4
6
8
10
12
14
16
0.39
20.39
40.39
60.39
80.39
100.4
120.4
140.4
160.4
180.4
200.4
220.4
240.4
260.4
280.4
300.4
320.4
Adjusted income in 1'000 CHF

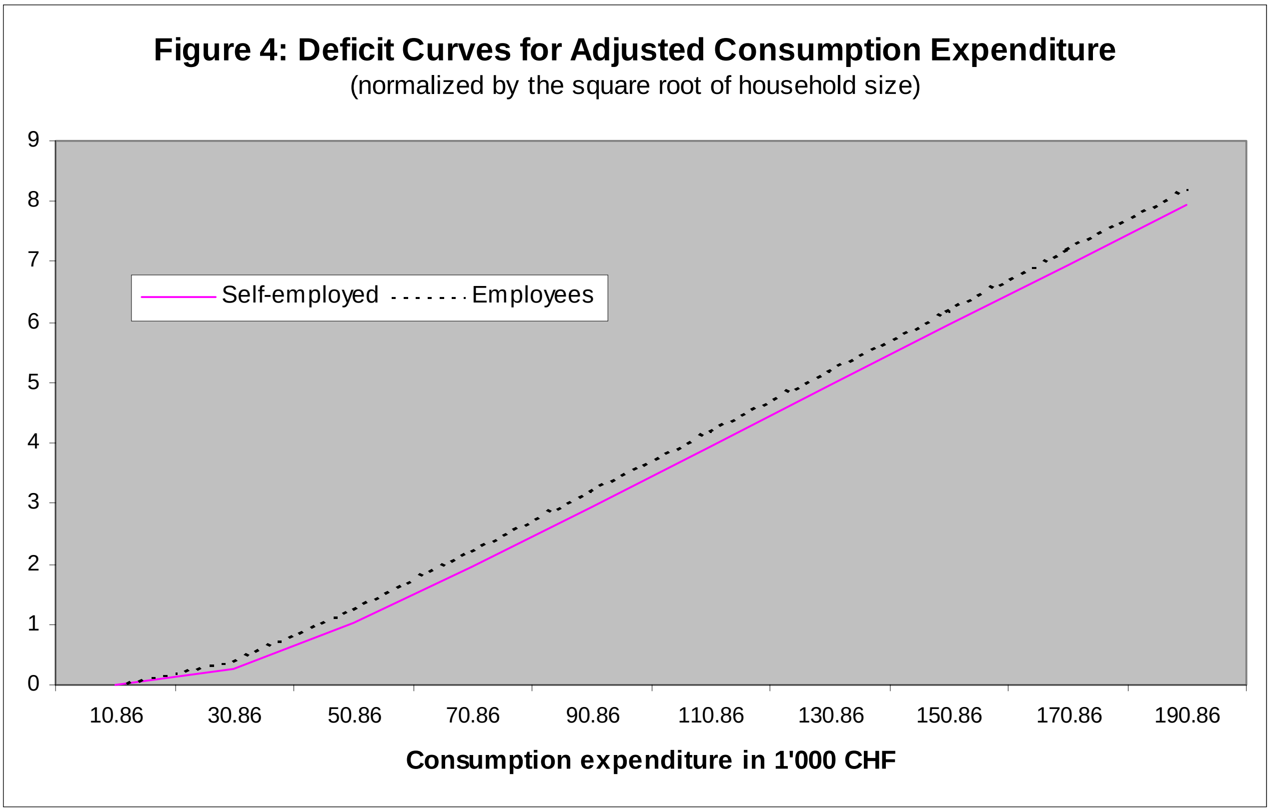


Figure 4: Deficit Curves for Adjusted Consumption Expenditure
(normalized by the square root of household size)
Self-employed
Employees
0
1
2
3
4
5
6
7
8
9
10.86
30.86
50.86
70.86
90.86
110.86
130.86
150.86
170.86
190.86
Consumption expenditure in 1'000 CHF

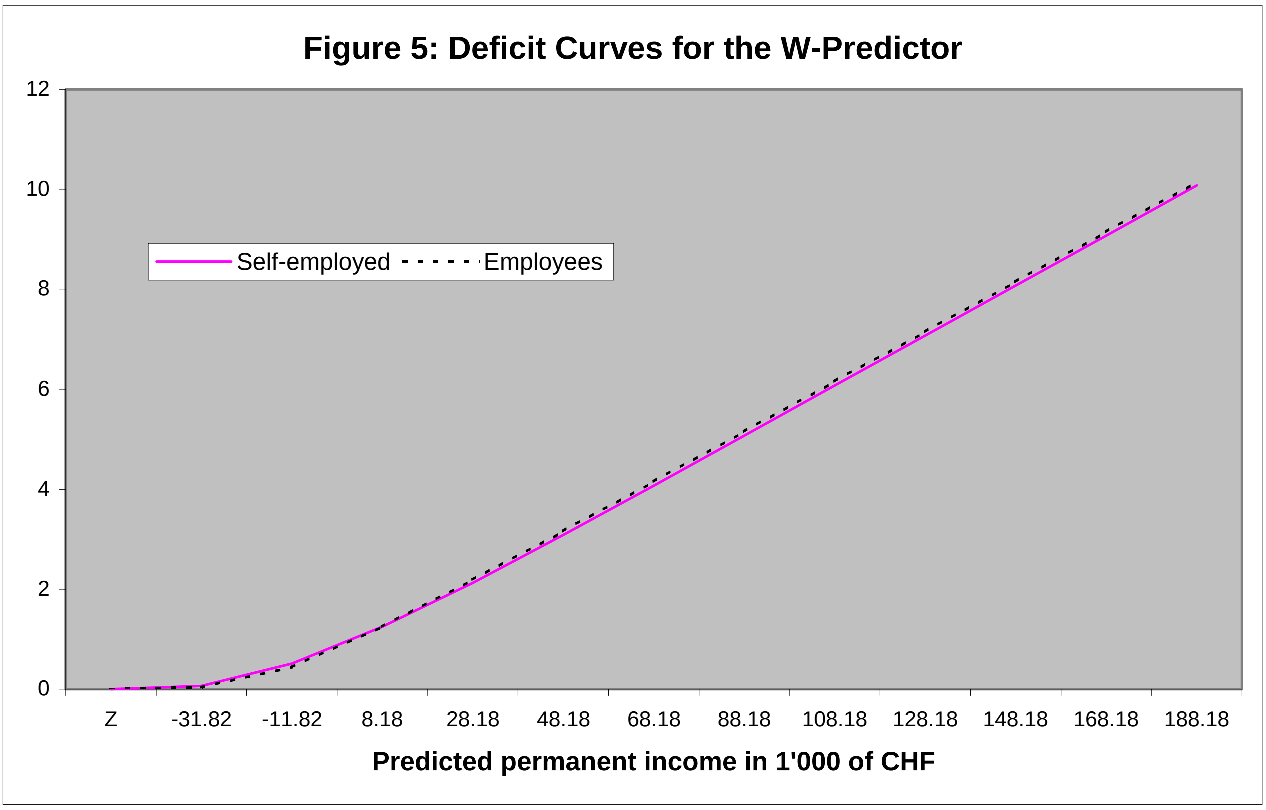


Figure 5: Deficit Curves for the W-Predictor
12
10
8
6
4
2
0
Self-employed
Employees
Z
-31.82
-11.82
8.18
28.18
48.18
68.18
88.18
108.18
128.18
148.18
168.18
188.18
Predicted permanent income in 1'000 of CHF