PREDICTION AND DETERMINATION OF HOUSEHOLD PERMANENT INCOME*

Ramses H Abul Naga
University of Lausanne

and

Robin Burgess
London School of Economics and Political Science

Discussion Paper
No. DARP/32
November 1997

The Toyota Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020-7955 6678

---

* We are grateful to seminar participants at the University of Geneva, London School of Economics, University of Lausanne, University of Namur, Stanford University, University College London, and to Christian Schlüter, Marcel Fafchamps, Arthur Goldberger, Alberto Holly and John Muellbauer for useful comments and suggestions.

Abstract

This paper is about the determination and prediction of permanent income in household data. Standard static welfare indicators (e.g. per capita expenditure and income) are imperfect in this respect as they typically contain a high transitory component. The framework we employ is consistent with the permanent income hypothesis but is supplemented with a causes equation where unobservable permanent income is explicitly modelled as a function of causal variables which play a key part in its determination. Simultaneous estimation of the model allows us to compare how well different standard static welfare indicators identify permanent income but more importantly enables us to predict permanent income using information contained both in the causal variables and in the standard static welfare indicators. The paper is closed by an application of the methodology to household data from the rural sectors of two Chinese provinces.

Keywords: Permanent income, prediction, determination, living standards, rural China.

JEL Nos.: I30, C31, O12.

Contact address: Dr Robin Burgess, STICERD, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, email: **r.burgess@lse.ac.uk**

1 Introduction

In this paper we develop a framework to look at the determination and prediction of permanent income in household data. The framework we employ is consistent with the permanent income hypothesis (PIH) proposed by Friedman (1957) but is supplemented with a causes equation where unobservable permanent income (Y) is explicitly modelled as a function of observable variables (Z) which play a key part in its determination.[1] The framework we propose is conceptually interesting in that it provides a (reduced form) picture of the living standard determination process and places an explicit focus on a range of determinants of permanent income which do not typically fall within the ambit of redistributive welfare policy. This in turn provides insights into the appropriate design of public policy to support living standards in rural China and elsewhere.

Identification of the permanent income (Y) of households is also of interest as many forms of public policy seek to target households with chronically low standards of living as opposed to those experiencing transient poverty. Standard static welfare indicators (X - e.g. per capita expenditure (PCE) and income (PCI)) are imperfect in this respect as they typically contain a high transitory component. A growing literature in the area seeks to identify the most satisfactory indicator (X) (often defined as the least noisy correlate) of long-term income (Y).[2] The framework we propose enables us to compare how well different standard static welfare indicators (X - e.g. PCI, PCE) identify permanent income (Y). However, we also attempt to move beyond this type of analysis and examine whether it is possible to construct welfare measures which are more informationally efficient and which perform better than observed static income or consumption in terms of proxying permanent income and identifying households with chronically low standards of living. Household permanent income (Y) is modelled as an unobservable and information relevant to its determination (Z, X) is used to obtain predictors of this unobserved variable. Because they make more intensive use of welfare relevant information contained in household data, we can show that these predictors outperform standard static alternatives and represent attractive candidates for practical usage in the identification of the long term poor.[3]

Our analytical framework builds on the model of factor analysis which seeks to account for the correlation between a set of p observed indicators X using a smaller set of c unobserved (latent) variables Y, with $c < p$. In its simplest form, the model has one latent variable ($c = 1$) and is written as:

$$X = \beta Y + u \tag{1}$$

[1] Inclusion of these Z variables allows the model to be identified.

[2] See Glewwe and van der Gaag, 1990; Anand and Harris 1990; Chauduri and Ravallion, 1994.

[3] Their attractiveness is also enhanced by the fact that they can be estimated using cross-sectional household data thus imposing no additional data requirements. This is important as the bulk of household survey information in developing countries is for a single year.

where u is a p-dimensional vector of disturbances. We then augment (1) with an equation explaining Y:

$$Y = \gamma' Z + \epsilon \tag{2}$$

where Z is $k \times 1$ column vector and ϵ is a disturbance term. Equation (2) is commonly referred to as the causes equation, and thus the system consisting of (1) and (2) is known as the multiple indicators-multiple causes ($MIMIC$) model.

The $MIMIC$ model was initially suggested by Zellner (1970). Its estimation has further been refined by Goldberger (1972) and Joreskog and Goldberger (1975). The paper extends the works of these previous authors by deriving alternative predictors for the unobserved variable Y and by comparing them in the light of various statistical criteria. We shall see below the model lends itself for a straightforward empirical specification of the Permanent Income/Life Cycle hypothesis (PIH). Following Deaton (1992) we define permanent income as the annuity value of current human and non-human wealth. The key assumption in our model is that the observed correlation between income and expenditure is induced by their joint dependence on the household's long-run or permanent income.[4] Such a formulation is consistent with the PIH but also covers a broader class of stochastic models where consumption exhibits the martingale property (see Deaton, 1992: 22-29). For the purpose of the present paper, it suffices to see current income and expenditure being modelled as linear functions of a common set of explanatory variables Z, a common error term ϵ, as well as specific disturbances u_i for each of the X variables. Empirical applications on the Chinese data indicate that the model provides an adequate fit for the underlying data.

The analysis gives particular attention to the problem of measurement error. We lay emphasis on the fact that indicators of permanent income are what they are: namely noisy measures of the latter variable, not error free measures as many researchers claim them to be. In this sense our analysis goes in the direction of attending Friedman's complaint about the practice of econometrics in the academic profession:

"Similarly, in academic studies, the common practice is to regress a variable Y on a vector of variables X and to then accept the regression coefficients as supposedly unbiased estimates of the structural parameters, without recognising that all variables are only proxies for the variables of real interest, if only because of measurement error, though generally also because of transitory factors that are peripheral to the subject under consideration" (Friedman, 1992: 2131).

The outline for the remainder of the paper is as follows. Section 2 discusses the specification of the PIH in terms of the $MIMIC$ model and discusses its estimation through the analysis of its covariance matrix. Section 3 deals with the problem of how to predict permanent income. Unlike the Factor Analysis model suggested by Abul

[4]For specifications and tests of the PIH, the reader may consult Attfield, 1976, 1980; Bhalla, 1978, 1979, 1980; Musgrove, 1979; Muellbauer, 1983.

Naga (1994) for the prediction of long-run income status the $MIMIC$ framework offers three potential routes to predicting permanent income. One possibility may consist in using the indicator variables (equation (1)) in order to predict Y. Another possibility is to predict Y using the causes equation (2), while finally it may be possible to use both sets of observed variables (X and Z) in order to derive a predictor of the latent variable. There are statistical and other considerations which guide our choice of predictor and these are also discussed in Section 3.

Section 4 discusses in what ways the $MIMIC$ model may be useful with respect to the problem of identifying the permanent income of households. Decomposition of the variance of static indicators (income, consumption) into permanent and transitory components allows us to identify which represents the least noisy correlate of permanent income thus guiding our choice of welfare indicator. Estimation of the causes equation (2) may also shed some light on which Z variables have the greatest weight in the living standard determination process. This has implications for public policy as it illustrates how influencing variables other than income and consumption can be instrumental in raising long-run living standards. Finally we discuss how constructed predictors of permanent income outperform static indicators (current income and consumption) in terms of identifying the long run living standards of households because they make more intensive use of welfare relevant information in household data.

In Section 5 we illustrate the proposed framework using household data from the rural sectors of two Chinese provinces, Sichuan and Jiangsu. This data represents an interesting testing ground to examine the usefulness of the $MIMIC$ model as information both on income and consumption is collected, and because government policies to support living standards have focussed mainly on providing all rural households with a basic opportunity set. Provision of basic entitlements to cultivable land and primary education are both examples of this approach. This focus on causal (Z) variables is hardly surprising given the logistic and other difficulties associated with operating cash or in kind transfer schemes – which focus on X variables – in a poor country the size of China. The fact that Sichuan and Jiangsu represent two faces of rural China, the former being poorer, inland and predominately agricultural whilst the latter is richer, coastal and has a high degree of rural industrialisation also allows us to identify key differences in the living standard determination process between the two provinces. Implications both for identification of the long-term poor and for the design of policy to support living standards are drawn out from this empirical analysis. Section 6 concludes by reviewing what insights have been obtained from the analysis.

2 Specification of the *PIH* and Living Standards

In their work on income inequality in the United States, Friedman and Kuznets (1945) suggested that observed income at a particular point in time could be decomposed into the sum of an individuals long run income together with a transitory component. In his subsequent work on the consumption function, Friedman (1957) proposed a similar decomposition for current consumption: observed consumption was hypothesised to be a function of the households permanent income and a transitory component.

Letting Y denote the household's unobserved permanent income, X_1 and X_2 denote their current income and consumption expenditure respectively and defining u_1 and u_2 as the corresponding transitory components, we can write the following system:

$$X_1 = Y + u_1 \tag{3}$$

$$X_2 = \beta_2 Y + u_2 \tag{4}$$

Assume X_1 and X_2 are centred around their means and u_1 and u_2 are uncorrelated.[5] It can readily be seen that the system consisting of equations (3) and (4) is generally not identifiable. We require the estimation of β_2, the variance of Y and the variances of the two disturbance terms, four unknowns, using three moment equations ($var(X_1), var(X_2)$ and $cov(X_1, X_2)$). One approach would consist of finding a third correlate of Y. This would render the above model identifiable (see, for example, the discussion in Goldberger, 1972). The alternative approach, suggested by Zellner (1970), is to introduce an equation that explains Y:

$$Y = \gamma' Z + \epsilon \tag{5}$$

Where Z is a $k \times 1$ vector of observed causes and ϵ is a disturbance term. To give a concrete example of the variables contained in Z, we can mention the work of Muellbauer (1983) on Sri Lankan data (see also Glewwe, 1991). Muellbauer includes in Z a set of variables which denote the demographic, asset ownership, housing, educational and occupational status of the household. Permanent income is thus viewed as a function of the human and non-human capital of the household conditioned by its composition which controls for position in the life-cycle (see Deaton, 1992).[6] A possible formulation therefore might be:

$$Y = \gamma_1 D_h + \gamma_2 E_h + \gamma_3 A_h + \gamma_4 C_h + \epsilon \tag{6}$$

Where permanent income (Y) is seen to depend on household composition (D_h) the educational and occupational status (E_h), stock of physical assets (A_h), and community or environmental characteristics such as access to amenities (C_h).[7] These

[5] We are aware that this assumption rules out the possibility that consumption may be sensitive to transitory shocks in income. This assumption is being relaxed in a follow-up research to the present work.

[6] And hence the marginal propensity to consume out of permanent income.

[7] This type of formulation is consistent with household production theory where in a rural setting physical asset stocks might include both monetary (e.g. savings) and non-monetary components (e.g. land, grain stocks, housing, household durables, productive assets - see Singh, Squire and Strauss, 1986).

Z variables thus represent longer term characteristics of the household which have bearing on the determination of permanent income and are likely to be measured with less noise than income or expenditure. The γ coefficients provide us with insights into the determination of long-run living standards. If these explanatory variables can be regarded as exogenous then it should be possible to make causal inferences which can be used to guide welfare support policies.[8]

The causal structure of the $MIMIC$ model (equations (3) to (5)) can be depicted in terms of a path diagram (see Figure 1). In path diagrams observed variables are squared, latent variables are circled, and disturbance terms are unsigned. The path diagram indicates that the error terms, ϵ and u, are all assumed to be uncorrelated and the X's are influenced by the Z's and ϵ (through their dependence on Y). Throughout the paper we maintain these assumptions, in particular:

$$\begin{aligned} E(u_i^2) &= \omega_{ii},\ E(u_i u_j) = 0 \\ cov(Z\varepsilon) &= [0],\ cov(Yu) = [0],\ cov(\epsilon u) = [0] \end{aligned} \tag{7}$$

where [0] denotes a matrix of zeros of appropriate order. The reduced form of the model can be obtained by substituting for Y (equation (5)) into (3) and (4):

$$\begin{aligned} X_1 &= \gamma' Z + \epsilon + u_1 \\ X_2 &= \beta_2 \gamma' Z + \beta_2 \epsilon + u_2 \\ i.e.\ X &= \beta\gamma' Z + \beta\epsilon + u \end{aligned} \tag{8}$$

where $\beta' = [1\ \beta_2]$ and $X' = [X_1 X_2]$.[9] Let ν_j denote the disturbance term for the reduced form equation (8):

$$\nu_j = \beta_j \epsilon + u_j \tag{9}$$

Also let $\Pi = \beta\gamma'$ denote the matrix of reduced form parameters.

2.1 The *SURE* Framework

The $MIMIC$ model can be expressed in the form of a system of Seemingly Unrelated Regression Equations ($SURE$) (Zellner, 1962):

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} \Pi_1 \\ \Pi_2 \end{bmatrix} Z + \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} \tag{10}$$

As pointed out by Goldberger (1972), the two main features of the model are that (i) the vector of reduced form coefficients is of the multiplicative form:

$$\Pi_i = \beta_i \gamma' \tag{11}$$

[8] Though we must keep in mind that they necessarilly reflect unobserved household tastes (e.g. for saving), which in the long run would impact on wealth.

[9] An alternative normalisation device is to set variance of the causes equation equal to unity (i.e. $\sigma_{\epsilon\epsilon} = 1$) and β_1 to be unconstrained (see Joreskog and Goldberger, 1975).

so that the matrix $\Pi = \beta\gamma'$ is of rank 1, and (ii) the covariance matrix of ν has a factor analytic structure:

$$\Sigma_\nu = E\left(\left[\begin{array}{c}\nu_1\\ \nu_2\end{array}\right]\left[\begin{array}{cc}\nu_1 & \nu_2\end{array}\right]\right) = \left[\begin{array}{cc}\sigma_{\epsilon\epsilon}+\omega_{11} & \beta_2\sigma_{\epsilon\epsilon}\\ \beta_2\sigma_{\epsilon\epsilon} & \beta_2^2\sigma_{\epsilon\epsilon}+\omega_{22}\end{array}\right] \tag{12}$$

viz. it is the sum of a matrix of unit rank and a diagonal matrix of full rank:

$$\Sigma_\nu = \sigma_{\epsilon\epsilon}\beta\beta' + \Omega \tag{13}$$

where Ω is the diagonal covariance matrix of the u's whose elements are defined in equation (7).

One important point to note is that in estimating the model as a $SURE$ system one will not obtain efficient parameter estimates unless one introduces the constraint on the reduced form parameters to have the multiplicative form (11). While we do not wish to make the claim that efficiency should be the prime criterion for statistical analysis, we discuss below an alternative method for estimating the $MIMIC$ model which provides a more general way of analysing latent variable models.

2.2 The *LISREL* Framework

Perhaps a more familiar way to study latent variable models is through the analysis of the covariance structure implied by the proposed model. Let δ denote the set of structural parameters and $\Sigma(\delta)$ be the covariance matrix implied by the particular model. The researcher chooses estimates of δ in order to minimize a distance function $F[\Sigma(\delta), S]$ between the model covariance matrix and its sample counterpart S.

Letting W denote the $(p+k)\times 1$ column vector consisting of $[X, Z]'$, we have:

$$\Sigma_W = \left[\begin{array}{cc}\Sigma_X & \Sigma_{XZ}\\ \Sigma_{ZX} & \Sigma_Z\end{array}\right] \tag{14}$$

where Σ_X is a square $p\times p$ matrix, Σ_Z is $k\times k$, and Σ_{XZ} is a $p\times k$ matrix, and Σ_{ZX}, the transpose of Σ_{XZ} is a $k\times p$ matrix. From the reduced form equations (8) we have:

$$\Sigma_X = \beta var(Y)\beta' + \Omega = \beta(\gamma' ZZ'\gamma + \sigma_{\epsilon\epsilon})\beta' + \Omega \tag{15}$$

also:

$$\Sigma_{ZX} = cov[(\beta\gamma' Z + \beta\epsilon + u)Z'] = \beta\gamma' ZZ' \tag{16}$$

When Z is a stochastic matrix we have $\Sigma_{ZX} = \beta\gamma' E(ZZ')$. Here we treat Z as being fixed. However, we note that the extension of the model to the case where Z is stochastic is straightforward and does not pose any new problems.[10]

[10]See Joreskog and Goldberger, 1975 for further details.

Under such circumstances, the covariance structure implied by the $MIMIC$ model can be written as:

$$\Sigma_W = \begin{bmatrix} \beta(\gamma' ZZ'\gamma + \sigma_{\epsilon\epsilon}) + \Omega & \beta\gamma' ZZ' \\ ZZ'\gamma\beta' & ZZ' \end{bmatrix} \tag{17}$$

Let S_W denote the sample correlation matrix of W. Generalised least squares estimation of (17) consists of minimising the sum of squares of the residuals weighted by a positive definite matrix which converges to the inverse of the covariance matrix of the elements of S. In virtue of the symmetry property of covariance matrices, we note that if Σ_W is of order c, the vector of residuals will be of order $\frac{c(c+1)}{2}$. Temporarily suppressing the subscript W, we can pile up the diagonal and below diagonal elements of Σ and S into the vectors σ and s:

$$\begin{aligned} \sigma &= (\sigma_{11}, \sigma_{21}, \sigma_{22}, \sigma_{31} \dots\dots\dots\dots \sigma_{cc}) \\ s &= (s_{11}, s_{21}, s_{22}, s_{31} \dots\dots\dots\dots s_{cc}) \end{aligned}$$

The GLS estimator is the solution $\widehat{\delta}$ which minimizes the distance:

$$F[\Sigma(\delta), S] = (s - \sigma(\delta))'\Theta^{-1}(s - \sigma(\delta)) \tag{18}$$

where Θ contains the fourth-order moments $\sigma_{ij,gh}$ which are consistent estimates of the asymptotic covariance between s_{gh} and s_{ij}. In order to assess the adequacy of a covariance model, we propose the Tanaka and Huba (1985) measure:

$$GFI = 1 - \frac{(s - \sigma(\widehat{\delta}))'\Theta^{-1}(s - \sigma(\widehat{\delta}))}{s'\Theta^{-1}s}$$

which is one minus the ratio of the minimum of the distance function after the model has been fitted to the term $s'\Theta^{-1}s$. The goodness of fit index in its adjusted definition is:

$$AGFI = 1 - \frac{c(c+1)}{2d}(1 - GFI) \tag{19}$$

where d is the number of degrees of freedom. We use this latter definition of the Tanaka and Huba measure.

3 Prediction

Since permanent income is assumed unobservable in our framework, one of our essential tasks is to draw inferences about this latent variable using the data on living standards available to us.[11] As the $MIMIC$ model contains two sets of variables, X,

[11] With expansion of household surveys, information about different aspects of the living standards of households has become much more widely available. Rather than relying on a single indicator, our framework allows us to pool information from different correlates and determinants of the long-run living standards of households.

the correlates, and Z, the causes, we can essentially follow three alternative routes to prediction (i) The $X-$ *predictor* (Y_X^*); a predictor of Y using the correlates X, (ii) The $Z-$ *predictor* (Y_Z^*) a predictor of Y based on the causes equation (2), (iii) The $W-$ *predictor* (Y_W^*); a predictor of Y which combines the available information in both X and Z variables.

We first note that the Best Linear Predictor (BLP) of a variable U constructed from a set of variables W has the general form:

$$U^* = E(UW)'(E(WW'))^{-1}W \tag{20}$$

where U^* is the predictor of U, and W is written as a column vector. Also, note that under the assumption of multivariate normality, the posterior mean $E(U|W)$ is equal to the best linear predictor U^*. The best linear unbiased predictor on the other hand obtains as a solution for the constrained optimization problem:

$$\min E[U - b'W]^2 \ s.t. \ E[b'W] = E(U) \tag{21}$$

That is, the best linear unbiased predictor is the minimum mean-square error linear predictor chosen subject to an unbiasedness restriction (Goldberger, 1962).[12] Since this predictor is chosen from a restricted set of linear predictors, its mean square error (MSE) cannot be lower than that of the BLP. When the unbiasedness condition is binding, one will generally have to sacrifice some prediction precision in the MSE sense, in order to obtain unbiasedness.[13]

Noting that $E(YZ') = \gamma'ZZ$, applying the formula (20), the BLP of Y using Z can be derived as:

$$Y_Z^* = \gamma'ZZ(ZZ')^{-1} = \gamma'Z \tag{22}$$

The $Z-$ *predictor* also readily obtains from the "causes" equation (2):

$$Y_Z^* = E(Y) = \gamma'Z$$

From this it follows that the $Z-$ *predictor* can be obtained from the regression function of Y on the causal variables (Z). In the context of cross-section data where error terms are assumed uncorrelated across family units, the regression function constitutes the minimum square error predictor of Y using Z.

While the BLP of Y using Z is unbiased, best linear predictors of Y that use X variables will generally be biased. For the sake of deriving minimum MSE unbiased predictors of Y that use X, or that combine X and Z information simultaneously, we thus need to solve the constrained optimisation problem of the type (21) above.

[12] See Abul Naga (1996) for a definition of unbiasedness in the context of factor analysis.

[13] Minimising with respect to the unbiasedness constraint in the form shown in (21) implies that all predictors share the same mean and can be compared on the basis of their respective $MSEs$ (see Figures 2 and 3).

The derivation of the X and $W - predictors$ are detailed in the appendix of the paper (Section 7). The best linear predictor of Y using X, obtained subject to an unbiasedness constraint (see Abul Naga, 1996), can be expressed as:

$$Y_X^* = (\beta'\Sigma_X^{-1}\beta)^{-1}\beta'\Sigma_X^{-1}X \tag{23}$$

where Σ_X is the covariance matrix of X given in expression (15). For the $W - predictor$ (the predictor combining X and Z information simultaneously) we obtain:

$$Y_W^* = \sigma_{\epsilon\epsilon}\beta'\Sigma_\nu^{-1}X + (1 - \sigma_{\epsilon\epsilon}\beta'\Sigma_\nu^{-1}\beta)\gamma Z \tag{24}$$

By substituting (8) into (23), we have that:

$$Y_X^* = (\beta'\Sigma_X^{-1}\beta)^{-1}\beta'\Sigma_X^{-1}(\beta\gamma'Z + \beta\epsilon + u)$$

i.e.

$$Y_X^* = \gamma'Z + \eta$$

where η is white noise, viz $E(\eta) = 0$. Likewise for the $W - predictor$:

$$Y_W^* = \sigma_{\epsilon\epsilon}\beta'\Sigma_\nu^{-1}(\beta\gamma'Z + \beta\epsilon + u) + (1 - \sigma_{\epsilon\epsilon}\beta'\Sigma_\nu^{-1}\beta)\gamma Z$$

which can be expressed as:

$$Y_W^* = \gamma'Z + \phi$$

where also $E(\phi) = 0$. For both the X and $W - predictors$ we can then write the decomposition:

$$Y_i^* = Y_Z^* + white\ noise,\ i = X, W \tag{25}$$

It is in this sense that the predictors are unbiased. It also follows that the X and $W - predictors$ will have higher variances than the $Z - predictor$.[14]

Going back to equation (24) we can express Y_W^* as a weighted average of Y_X^* and Y_Z^*:

$$Y_W^* = \theta Y_X^* + (1-\theta)Y_Z^*\ where\ \theta = \sigma_{\epsilon\epsilon}\beta'\Sigma_\nu^{-1}\beta \tag{26}$$

It is shown in the Appendix that Y_W^* dominates either of Y_X^* and Y_Z^* in terms of MSE, and that $0 < \theta < 1$. Because θ is increasing in $\sigma_{\epsilon\epsilon}$ (equation (A7)), we suggest the following interpretation for Y_W^* : in order to minimise prediction mean-square error it is necessary to pool the information about Y contained in Y_X^* and Y_Z^* . As $\sigma_{\epsilon\epsilon}$ goes to zero, (that is, as Z accounts for more of the variation in Y) θ goes to zero, so that Y_W^* relies on Z variables. Conversely, as $\sigma_{\epsilon\epsilon}$ rises, θ increases and more weight is being placed on X variables in Y_W^*. Note finally that while the X and $Z - predictors$

[14] Also, from the results on the measurement of inequality, both Y_X^* and Y_W^* can be viewed as mean preserving spreads of Y_Z^* (Atkinson, 1970). Thus inequality of permanent incomes will always be higher when measured using Y_X^* and Y_W^* than when using Y_Z^*. Likewise, expected poverty (Ravallion, 1988) will appear to be higher when using the X and $W-predictors$ than when measuring permanent income using the $Z - predictor$.

can be compared in variance terms (see (25)), they cannot be ranked in terms of prediction MSE (see equations (A9) and (A10) of the Appendix.

Given that the variance of the $Z - predictor$ underestimates the true variance of permanent income, this raises the possibility that the $Z - predictor$ is worse at predicting the permanent incomes of individuals at the extremes, so that it may be of limited use in identifying the households with chronically low standards of living. Since the variance of permanent income lies somewhere between the variance of current income and that of the $Z - predictor$[15] there may be grounds for believing that on the whole the $Z - predictor$ may be more precise at predicting middle ranges of the distribution of Y, but that magnitudes of prediction errors may be larger at the tails. It is perhaps there that using the X variables adds most information about the permanent incomes of households. In this sense focussing our attention on unbiasedness and efficiency criteria exclusively may not be appropriate.

There is yet another way in which the superiority of the $W - predictor$ over the Z and $X - predictors$ may be advocated: in the case of living standard analysis, where typically long run income is not observed, we believe that a predictor of permanent income should ideally be chosen in a way as to exhaust all the sample information on the underlying latent variable. In technical terms, a statistic which contains as much information about permanent income as the data itself, is said to be *sufficient* (Bartholomew, 1984). While sufficient statistics do not exist for all inferential problems, the concept is useful for identifying approaches to the prediction of permanent income which do not meet this particular criterion. Thus, for instance, a researcher who possesses data on a households consumption expenditure and annual earnings, and chooses to predict permanent income using consumption expenditure only (on the grounds of lower noise) will not be using a sufficient statistic for the unobserved variable. Unless, earnings contain no other information about permanent income than the information already available in consumption, the approach of predicting permanent income using its least noisy indicator cannot be recommended on sufficiency grounds. Furthermore, for LDC data, where measurement errors may be high, the need to exploit all the available information on long run income is all the more necessary.

We would therefore argue strongly that unbiasedness and efficiency should not be the sole criteria for the selection of a predictor. Since both X and Z contain information on Y, a statistic which does not use both sets of variables cannot exhaust all the available sample information on Y. Writing the joint density of the random variables as follows:

$$f(X, Z, Y) = f(Y|X, Z) \times f(X, Z)$$

we can use the results of Bartholomew (1984) to conclude that a sufficient statistic for Y must be based on the predictive distribution $f(Y|X, Z)$. Predictors based on X without Z, or on Z excluding X, do not exhaust all the sample information

[15] This follows from reading the reduced form equations (8).

on Y, and thus will fail to meet the sufficiency requirement. On the basis of this sufficiency result, it is clearly preferable to employ Y_W^* as a predictor of permanent income as opposed to Y_X^* or Y_Z^*.[16] In essence, the $Z-predictor$ throws away the welfare relevant information contained in the correlates of permanent income (i.e. current income and consumption) which on *a priori* grounds would be thought to be highly relevant to the prediction of permanent income. These considerations lead us to select the $W-predictor$ as our preferred indicator of permanent income. This predictor effectively balances the contribution of X_1 (current income), X_2 (current consumption) and Z information to prediction according to the degree to which they are correlated with unobserved permanent income (Y).

This does not imply the $X-predictor$ is useless for practical work. Just like the $W-predictor$ its use can be recommended on other grounds than prediction MSE (i.e. efficiency). First of all, the researcher or policy maker may not have the necessary data on the Z variables to make a prediction about the family's long run income. Observing current income and expenditure, on the other hand, may be at times less costly. In our applications sections we show that the weights on income and consumption in the $X-predictor$ tend to fall within the $0.45-0.55$. As a first hand approximation one could therefore compute a predictor of permanent income as an average of current income and consumption.

There is also the fact that the $X-predictor$ is based solely on correlates of permanent income. It is known from previous studies on living standards that income and consumption do not identify the same households as being in poverty (e.g. Chaudhuri and Ravallion, 1994). On such basis the $X-predictor$ can offer a sensible compromise between relying on income data on the one hand, versus consumption exclusively, in identifying households with low standards of living. As shown in our applications section, by predicting permanent income using Z variables we do not necessarily identify a larger set of poor individuals than that which is common to income and expenditure definitions. This is not an undesirable feature of Y_Z^* (and Y_W^*), but rather a statement of the fact that poverty of productive resources or opportunities (Z space) need not be equivalent to poverty measured as a function of outcomes (i.e. X variables).

4 Analysis of Living Standards

Researchers, policy makers and administrators routinely face the problem of selecting an observable indicator of welfare from cross-sectional data. These indicators are expected to convey information about the welfare of households well beyond the

[16] Also, because they make more intensive use of the living standard relevant information in the data set, any of the X, Z or $W-predictors$ would be preferable to using per capita expenditure *or* income which are the standard measures employed in empirical analysis.

survey period.[17] Static welfare measures (X) are seen as (noisy) indicators of the permanent income of the household which are the objects of interest if we are trying to alleviate chronic as opposed to transient poverty. To be consistent with economic theory what are needed are measures that approximate money metric utility.[18] The two leading practical candidates in this respect are per capita (or equivalised) income and consumption.[19] The underlying problem is that these static welfare measures (X) are imperfect measures of unobservable long-term welfare (Y). We are therefore faced with problems of choosing between static indicators or of combining information from different indicators in the identification of longer term circumstances of households.

In some cases the choice between income and expenditure is dictated by the fact that information is only collected on one variable.[20] In the bulk of other cases the choice is usually made by resorting to *a priori* arguments. In the developing country setting, consumption based measures are typically preferred for a variety of reasons. Based on the permanent income hypothesis it is argued that these represent smoother indicators of permanent income than current income in particular when data is collected over short periods. For households which are unrestricted in their opportunities to buffer their income variability, their short-term consumption levels will reveal their permanent income at that date. Incentives to understate income may be greater whereas expenditure is typically calculated as the aggregate of a number of items, reporting of which are less sensitive to downward bias. It is also argued that consumption represents a more natural framework to impute the value of home production which is central to the welfare of rural households.

The presumed superiority of consumption measures, however, rest mainly on practical data issues which dwarf the theoretical considerations. Given this state of affairs it is largely an *empirical* issue as to which observable measure of money metric utility should be preferred in a given country. This dilemma has spawned a growing literature on the choice of static indicators of permanent income from cross-sectional data.[21]

Given the specification of the causes equation (2) the $MIMIC$ model allows us to choose the least noisy indicator of permanent income without having to rely on *a priori* arguments which are not based on any analysis of the data at hand. Researchers and policy makers can then employ the preferred indicator of permanent income in poverty alleviation and other policies which rely on the identification of long-run living standards. What is more, the $MIMIC$ framework allows us to identify the variances of the permanent and transitory components of observed income and expenditure.

[17]This correspondence is made necessary by the fact that surveys cannot be carried out continuously and because there are costs associated with reallocation of benefits (see Chaudhuri and Ravallion, 1994)

[18]See Deaton and Muellbauer, 1980: Chapter 7.

[19]See Deaton, 1994.

[20]For example the National Sample Surveys in India only collect information on consumption.

[21]See Glewwe, 1990, Glewwe and van der Gaag, 1990; Anand and Harris, 1990; Chaudhuri and Ravallion, 1994 and Deaton, 1994.

Placing (2) into (1) and calculating variances we have that:

$$\frac{\beta_i^2 var(Y) + var(u_i)}{var(X_i)} = 1 \tag{27}$$

We can identify the variances of the permanent and transitory components of a given indicator as:

$$\begin{aligned} \frac{var(u_i)}{var(X_i)} &= \textit{variance share of transitory component} \\ 1 - \frac{var(u_i)}{var(X_i)} &= \textit{variance share of permanent component} \end{aligned} \tag{28}$$

Given the specification of the model, these shares provide us with a precise and observable measure of indicator noisiness which can be used to guide the choice of welfare measure in practical applications.[22]

Though the approach is useful for choosing with a greater degree of confidence between welfare indicators it is also useful for choosing between different definitions of the same welfare indicator. One can examine, for example, how correcting for noisy elements of expenditure (e.g. durables, housing) affect the performance of the consumption indicator (PCE) as assessed through (28). These critical decisions are often made on the basis of somewhat arbitrary assumptions typically without reference to the data set being analysed. The $MIMIC$ methodology offers some scope to improve on this practice. As we show in the empirical section (Section 5) these choices of definition can have a critical bearing on the relative performance of a given welfare indicator. Finally, it is also possible within the $MIMIC$ approach to compare the relative correlations of Xs and Zs with permanent income thus allowing us to examine the hypothesis that observed Zs may be more appropriate (i.e. less noisy) indicators of long-run living standards than observed Xs. Given that some Zs may be easier to observe than Xs, this raises the possibility that Zs may be better measures on which to base the targeting of resources and other redistributive schemes.

Through the causes equation (2), the $MIMIC$ approach also provides us with insights into the determination of living standards. Permanent income is seen to depend on household composition, educational and occupational status, stocks of physical and monetary assets (e.g. housing, land, consumer durables, productive assets, saving deposits) and community characteristics such as access to amenities (see Muellbauer, 1983; Glewwe, 1991).[23] The causal equation (2) should be viewed as a reduced form estimate of various structural relationships (e.g. asset returns, agricultural production function, earnings function).[24] If they can be treated as exogenous, coefficients on the Z variables tell us which factors may be important in determining

[22] This type of decomposition analysis parallels that for earnings mobility (see Lillard and Willis, 1978).

[23] See (6) for a formulation of the causes equation.

[24] Policies to support living standards through Z interventions would have to be based on a more

adjustment for the noisy elements in total expenditure, namely durables and housing. In the causal equation (2), permanent income is modelled as a function of eleven Z variables which serve as proxies of the human and non-human capital stocks which determine the permanent income of the household. Following Muellbauer (1983) and Glewwe (1991), permanent income (Y) is seen to depend on household composition, educational and occupational status, stock of physical assets, and community characteristics such as access to amenities (see (6)).[30]

5.2 Estimation

Estimation results of the income/consumption $MIMIC$ system of equations are given in Table 2. The system was estimated using the $LISREL$ framework presented in Section 2.2 using the normalisation device that $\beta_1 = 1$. Absolute t values are presented in parenthesis.

Table 2 suggests that the choice of permanent income indicator between observed current income and consumption is sensitive to which definition of per capita consumption is used (see Table 1 for alternative definitions of consumption). These results are most clearly brought by decomposing the variance of these static welfare indicators into transitory and permanent components using equation (28).

Uncorrected measures like $LNPCE1$ which include all current spending on current goods and services (including noisy items such as housing and durables) are outperformed by $LNPCI$ in both provinces. For intermediate measures like $LNPCE2$ where only the flow of value from durables is imputed, income and consumption perform almost equally well as each other, consumption slightly outperforming income in Sichuan and *vice versa* in Jiangsu. Corrected measures such as $LNPCE3$, where the flows of value from both housing and durable stocks are imputed in place of (noisy) current expenditures, clearly outperform $LNPCI$. The fact that Jiangsu constitutes a more monetised economy may help to explain why income appears to perform relatively better as an indicator of permanent income in this province *vis a vis* Sichuan.

The overall impression from Table 3 is that replacement of noisy elements of current expenditure with imputed value flows leads to a lowering of the transitory share in overall variance and a corresponding rise in its performance *vis a vis* income. Thus, assuming that the causal equation (2) is well specified, corrected consumption will be the preferred observed welfare measure to identify households with low permanent income. Taken together the results, however, cast doubt on the widespread belief that current consumption is always a better indicator of chronic poverty than current income. Our results are thus in line with Chaudhuri and Ravallion (1994) who do not find that the preference for consumption-based measures is supported in Indian

[30]The model, where possible, is specified in log terms to allow for the presence of non-linearities in the reduced form estimates.

longitudinal data.[31] Arguments based either on consumption smoothing or on the superiority of consumption data do not seem to be well supported in either study. What is emphasized in our study is the need to carefully correct for noisy elements in current consumption in order to improve its performance in identifying the chronic poor.

It is also possible to compare the performance of Xs versus Zs as indicators of permanent income. Our interest here is not in the living standard determination process *per se* but rather the issue of whether some Z variables might outperform X variables as indicators of Y. To make this comparison in a consistent way we calculate correlation coefficients between X and Z variables and unobservable Y.[32] The appropriate formula for the calculation of correlation coefficients (ζ) are:

$$\begin{aligned}\zeta_{XY} &= \frac{\beta(\rho^2+\sigma^2_{\epsilon\epsilon})}{\sqrt{var(X_i)(\rho^2+\sigma^2_{\epsilon\epsilon})}} \textit{where } \rho^2 = \gamma' Z Z' \gamma \qquad (29)\\ \zeta_{ZY} &= \frac{(ZZ'\gamma)_i}{\sqrt{var(Z_i)(\rho^2+\sigma^2_{\epsilon\epsilon})}}\end{aligned}$$

where $(ZZ'\gamma)_i$ is the ith entry of the vector $ZZ'\gamma$. These results follow from equations A1-A5 in the Appendix (Section 7). Coefficients for the uncorrected definition of ($LNPCE1$) are shown in Table 4.

This comparison confirms that either income or consumption, on the whole, perform better than the various Z variables in identifying permanent income. Thus if a single welfare indicator needs to be chosen these results indicate that it would be better to choose a money metric utility proxy (X measure) as opposed to a Z measure. This does not, however, imply that a X based approach to poverty identification is preferable to a Z approach. Such a choice would have be informed by consideration of a number of incentive, institutional and administrative considerations.[33]

Overall the estimation and correlation results suggest that it is unwise to not include both X and Z information in the prediction of permanent income given that both sets of variables contain information relevant to the identification of permanent income. These results can therefore be taken as empirical grounds for preferring the $W-predictor$ over the $Z-predictor$ (or $X-predictor$).

We now turn to a discussion of the coefficients on the Z variables in the causal equation (6) the form of which is fleshed out in Table 1. As these variables are mea-

[31] Comparing current income and consumption measures to six year means in the ICRISAT panel, Chaudhuri and Ravallion (1994) find that current income is a better indicator of chronic poverty than current consumption when poverty is defined with respect to mean income.

[32] This type of reasoning has been used as an argument to base targeting on landholding, housing characteristics, holding of physical assets etc (see van de Walle and Nead, 1995 for a review).

[33] For example, we are assuming that X can be correctly observed for all households. This typically is *not* the case. Problems with identifying and monitoring the X metric for different households might lead us to base public policy more firmly on more readily observable Z measures (e.g. landholding, demographics).

sured in different units the discussion will focus not on the coefficients (γ) themselves but on t statistics which give an idea of the weight that different Z variables have in the determination of Y. Several clear patterns emerge.[34] As would be expected, both household size (LNN) and the dependency ratio ($CHILDP$) have negative effects on permanent income (Y). The latter child effect appears to be stronger in Sichuan which may reflect lower off-farm employment opportunities in this province.[35] The educational level of the household head ($EDUHD$) has a small but significant effect on Y which is more pronounced in Jiangsu which may be due to off-farm employment opportunities raising the returns to education. The proportion of the household labour force in primary activities ($PRIMP$) is negatively and significantly related to Y in both provinces. This suggests that rural diversification into off-farm activities has a large positive impact on permanent income. This is an important finding given the large structural changes that have been taking place in the rural economy associated in particular with the growth of rural industry in the 1980s. Housing characteristics, both per capita floor area ($AREAPC$) or the electrification dummy ($ELEC$) are both positively associated with Y. Housing, in particular, appears to represent an important stock of wealth in Jiangsu. Stronger coefficients on the $ELEC$ variable (which can also be viewed as a community characteristic) in Jiangsu might also reflect complementarity with off-farm production activities.[36] As might be expected cultivable landholding per capita ($CULTPC$) has a significant and positive influence on permanent income confirming its central role in determining the permanent income of households.[37] $BWTV$ and $WATCH$ which proxy for the durable stocks[38] held by the household are both significantly and positively related to permanent income.[39] This suggests that durables represent an important stock of wealth in rural China. Per capita savings ($DEPOSIT$) also have a significant role in the determination of permanent income.[40] Interestingly this effect is stronger in poorer Sichuan which might be reflective of farmers in this province placing their surpluses

[34] Table 2 shows that as we move from the least to most corrected definitions of consumption the size of these coefficients (in absolute terms) tends to increase slightly. The overall pattern of results, however, is robust to changing the definition of PCE.

[35] Increasing population on a fixed land resource may lead to a more negative impact on welfare than where excess labour can be absorbed into other activities.

[36] This is an important finding as it suggests that expanding provision of rural infrastructure may be an important means of encouraging diversification and raising long run living standards.

[37] The centrality of access to land to the determination of household welfare is thus confirmed. Though households do not formally own land, the majority are allocated a plot of land by the village council for them to farm under contract (see Burgess, 1997).

[38] The value of the total household durable stock which is a more complete durable variable cannot be entered directly given that the imputed flow of value from this stock enters into the correction of the consumption terms $LNPCE2$ and $LNPCE3$.

[39] Interestingly the basic durable ownership indicator ($WATCH$) has a more signficant effect in the poorer province whilst the reverse is true for the luxury durable ownership indicator ($BWTV$).

[40] In 1990 nearly all household savings would be held with government banks (i.e. there was a monopoly), therefore this measure should be relatively good measure of the level of monetary savings kept in formal financial institutions. It can also be an inverse measure of liquidity constraints.

in financial institutions whereas a larger number of households in Jiangsu may be investing their surpluses directly in off-farm activities which are more risky but have higher mean expected returns.[41] A greater role for savings in rural Sichuan may also reflect that they play a greater role as a consumption smoothing device in this province as households cannot rely as much on off-farm earning streams. Per capita holdings of productive assets ($PASSETV$) also has a positive and significant impact on household welfare. Given that these assets are almost entirely agricultural in nature it is not surprising that this effect is stronger in Sichuan.

Not too much should be read into these causal equations as they represent the reduced form estimates of a host of structural equations which are not directly observed. They do, however, offer broad directions in terms of identifying which factors are important for the determination of rural living standards. Not all these factors can be thought of as policy dependent but a number of policy suggestions do follow from the analysis. The promotion of diversification, for example, through deregulation or the provision of basic infrastructure (e.g. electricity, telecoms, roads) can help to raise long-run living standards in both provinces. Expansion of education programmes will have a more significant impact in Jiangsu than Sichuan, though this effect is likely to be contingent on the level of diversification. Providing close to universal access to land is shown to be an important element in the support of rural living standards in both provinces. This confirms the central role that access to land plays in the determination of permanent income in rural China. Given small farm sizes, making access to land less egalitarian, for example, by altering the allocation mechanism or by introducing land markets, may have negative consequences for the long term welfare of rural households.[42] The liberalisation of financial markets which allows rural households to benefit from rapid private sector growth by increasing the returns on savings might also help to increase welfare. Finally, the results suggest that the subsidisation of agricultural productive assets will have greater impact in Sichuan than Jiangsu. It is thus possible through this crude analysis to pick up some common factors of importance in the two provinces as well as some key differences in terms of what might be advisable in policy terms.

5.3 Prediction

Prediction results using the formulae from Section 3 are shown in Table 5. We employ the normalisation $\beta = 1$ so that the predicted levels of permanent income are measured in the units of per capita income (*yuan*). Using this specification, different determinants (Z) and correlates (X) can thus be collapsed into a single

[41] Taken as a whole these findings are consistent with the analysis of demand patterns in rural China which shows that particularly in Jiangsu, richer rural households invest their surpluses primarily in building up housing and durable stocks or in pooling resources with other households to invest in off-farm rural enterprises (as opposed to placing these surpluses in low or negative interest bank accounts).

[42] In particular, where off-farm labour markets remain relatively undeveloped.

money metric predictor (Y) which is comparable in magnitude to observed income and consumption.[43]

In Table 5, as a means of avoiding clutter, we restrict our attention to the model using uncorrected consumption expenditure ($LNPCE1$).[44] In order to produce unbiased estimates of Y which are comparable in levels to PCE and PCI, both the Z and W prediction equations contain constants. The regression constant γ_0 is obtained under the condition that the regression line intersects the sample mean. In the case of the $Z-predictor$ this constant (γ_0) is given by:

$$\gamma_{0Z} = \overline{X_1} - \gamma'\overline{Z} \tag{30}$$

where bars denote sample means. In the case of the $W-predictor$ the size of this constant is scaled down to take into account the fact that the Xs are now playing a role in prediction:

$$\gamma_{0W} = \frac{\gamma_{0Z}}{(1-\sigma_{\epsilon\epsilon}\beta'\Sigma_{\nu}^{-1}\beta)} \tag{31}$$

Figures 2 and 3 which plot the non-parametric densities of the different predictors along with that of the observed income provide confirmation of the key results derived in Section 3. First, all three predictors are shown to be unbiased and are centred around the mean of $LNPCI$. Second, it is clear that in both provinces all three predictors have a lower variance than either observed consumption or income. Third, the $Z-predictor$ is shown to have the lowest variance. Though there is an increase in variance in moving from the Z to the $W-predictor$ there is a compensating gain in terms of the information set being covered which results in a reduction in MSE. Because the $W-predictor$ utilises X information, it is capable of explaining some of the residual variance that is left unaccounted for by sole use of Z information and this may have value in terms of identifying the long-run income status of households.

This gain in information (sufficiency), can be demonstrated by first ranking households according to the different predictors and then creating two way quintile frequency tables which give some idea of the degree of agreement between the different rankings (Table 6). The weight of households on the diagonal expressed as fraction of total sample size provides us with a measure of the degree of agreement between two predictors in terms of the identification of the welfare status of households. The degree of disagreement in welfare rankings between the X and $Z-predictors$ is large in both provinces . Only $37-38\%$ of household are classified to belong to the same quintile according to these criteria. Table 6 demonstrates that there is greater agreement on welfare rankings between the $W-predictor$ and either the X or the $Z-predictor$ than there is between the X and the $Z-predictors$. The $W-predictor$, by spanning

[43] The numbers presented in Table 5 reflect the weights on each of the variables which are measured in different units.

[44] Models using different definitions of consumption, where noisy housing and durable expenditures have been corrected, produce very similar results except that the relative weighting of income versus consumption changes.

both outcome (X) and opportunity (Z) space, thus forces agreement on the relative welfare rankings and should be preferred if we believe that *both* these spaces contain information which is relevant to the prediction of household permanent income (Y).

A similar analysis can be carried out to compare static welfare indicators (PCI, PCE) and the various predictors (Table 7). As can be seen from Table 7 the degree of agreement between observed income and consumption is fairly low; less than 50% of households are identified as being in the same quintile. The use of the $X-predictor$ substantially increases the degree of agreement on the welfare status of households. In cases where there is substantial disagreement according to income and consumption criteria, the $X-predictor$ thus has real value in steering a middle route between using only income information or only consumption information. Interestingly, in Table 7 we see that using the $Z-predictor$ leads to a *reduction* in agreement as regards welfare identification. This result is consistent with Table 6 and suggests that the ranking of households is substantially different when they are ranked in outcome (X) space as opposed to opportunity (Z) space. The subset identified as "poor" according to these two sets of criteria is likely to be substantially different irrespective of where the poverty line is drawn. Implementation of the $W-predictor$ leads to an improvement in agreement on the welfare ranking of households. Where there is uncertainty regarding the *true* welfare ranking the $W-predictor$ effectively balances the contributions of these types of information in prediction according to their correlation with the unobserved variable of interest (Y). In this way the $W-predictor$ makes fullest use of welfare relevant information in a given data set.

6 Conclusions

The central conclusion of the paper must be that it is possible to construct welfare measures which are more informationally efficient and which perform better than observed static income or consumption in terms of proxying for permanent income and identifying the chronically poor. Thus, in principle, for poverty and inequality analysis, we should view the choice between static welfare measures as being of secondary importance if new informationally more efficient welfare measures can be constructed.

The $MIMIC$ approach also has the potential to provide insights into both the measurement and determination of living standards. On the measurement front, the framework allows us to model long-run household status as an unobservable and thus compare the performance of various commonly used observed welfare indicators in identifying this status. The results of this analysis in the case of rural China are not straightforward and suggest that consumption cannot be *assumed* to be a better welfare indicator in a poor developing country setting. Preference for consumption based welfare measures rests on careful corrections being made for the noisy durable and housing elements in expenditure.

Where the costs of continuous household surveys are prohibitive, by making more intensive use of welfare relevant data in a given household survey, predicted welfare measures of the types suggested here can play a useful role in providing a better approximation of permanent income than is obtainable using observed income or consumption. This in turn can result in savings in welfare programmes if use of this new welfare indicator (e.g. $W - predictor$) results in the needy being more correctly identified. If cross-sectional data is going to be used in this fashion then careful thought needs to be put into survey design, in particular as regards ensuring that a complete and standardised set of Xs and Zs is gathered.

The $MIMIC$ framework has also proven to be useful for looking at the determination of living standards. In the case of rural China what this analysis has shown is that various Z factors have had a great deal of purchase in terms of improving the long-run status of households. Factors such as access to cultivable land and agricultural productive assets, access to basic infrastructure (e.g. electricity), economic diversification and financial savings are all shown to play an important role in the determination of living standards. Interesting differences emerge in the causal equation between the two provinces. The $MIMIC$ approach thus allows us to examine a richer set of influences on living standards than is standard in welfare policy. This is relevant as institutional features which characterize developing countries limit the scope for redistribution through monetary transfers due to administrative and logistic constraints and problems associated with monitoring and identifying needy individuals (Burgess and Stern, 1991, 1993).

References

[1] Abul Naga, R.H. (1994): "Identifying the Poor: A Multiple Indicator Approach", Distributional Analysis Research Programme Discussion Paper No. 9, LSE.

[2] Abul Naga, R.H. (1996): "Prediction and Sufficiency in the Model of Factor Analysis", discussion paper, DEEP, University of Lausanne.

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

[4] Atkinson, A. B. (1970): "On the Measurement of Inequality" Journal of Economic Theory 23: 244-263.

[5] Attfield, C.L.F. (1976): "Estimation of the Structural Parameters in a Permanent Income Model", Economica 43: 247-254.

[6] Attfield, C.L.F. (1980): "Testing the Assumptions of the Permanent Income Model", Journal of the American Statictical Association 75: 32-36.

[7] Bartholomew, D. (1984): "The Foundations of Factor Analysis" Biometrika.

[8] Bhalla, S. (1978): "The Role of Sources of Income and Investment Opportunities in Rural Savings" Journal of Development Economics 5: 259-307.

[9] Bhalla, S. (1979): "Measurement Errors and the Permanent Income Hypothesis: Evidence from Rural India" American Economic Review 69: 295-307.

[10] Bhalla, S. (1980): "Measurement of Permanent Income and it Application to Savings Behaviour" Journal of Political Economy 88: 722-744.

[11] Burgess, R.S.L. (1997): "Land, Welfare and Efficiency in Rural China" mimeo London School of Economics.

[12] Burgess, R.S.L. and Stern, N.H. (1991): "Social Security in Developing Countries: What, Why, Who and How?" in Ahmad, E., Dreze, J., Hills, J. and Sen A.K. Social Security in Developing Countries (Oxford: Oxford University Press).

[13] Burgess, R.S.L. and Stern, N. H. (1993): "Taxation and Development" Journal of Economic Literature 31: 762-830.

[14] Chaudhuri, S. and Ravallion, M. (1994): "How Well do Static Indicators Identify the Chronically Poor" Jouirnal of Public Economics 53: 367-394.

[15] Deaton, A. (1992): Understanding Consumption (Oxford: Clarendon Press).

[16] Deaton, A. (1994): The Analysis of Household Surveys: Microeconometric Analysis for Development Policy (mimeo Princeton University).

[17] Deaton, A. and Muellbauer, J. (1980): Economics and Consumer Behaviuor (Cambridge: Cambridge University Press).

[18] Friedman, M. (1957): A Theory of the Consumption Function (Princeton: Princeton Univ Press).

[19] Friedman, M. (1992): "Do Old Fallacies Ever Die" Journal of Economic Literature Vol XXX: 2129-2132.

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

[21] Glewwe, P. (1990): "Efficient Allocation of Transfers to the Poor" Living Standards Measurement Working Paper No. 70, World Bank, Washington D.C.

[22] Glewwe, P. (1991): "Investigating the Determinants of Household Welfare in Cote d'Ivoire" Journal of Development Economics 35: 307-337.

[23] Glewwe, P. and van der Gaag, J. (1990): "Identifying the Poor in Developing Countries: Do Different Definitions Matter?" World Development 18: 803-814.

[24] Goldberger, A. (1962): "Best Linear Unbiased Prediction in the Generalised Linear Model", Journal of the American Statistical Association 57: 369-375.

[25] Goldberger, A. (1972): "Maximum Likelihood Estimation of Regressions Containing Unobservable Independent Variables", International Economic Review 13: 1-15.

[26] Joreskog, 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.

[27] Lillard, L. and Willis, R. (1978): "Dynamic Aspects of Earning Mobility" Econometrica, 46(5).

[28] Muellbauer, J. (1983): "The Measurement of Long-Run Living Standards with a Permanent Income Model" mimeo Nuffield College, Oxford.

[29] Musgrove, P. (1979): "Permanent Income and Consumption in Urban South America" American Economic Review 69: 598-609.

[30] Rao, C.R. and Toutenberg (1995): Linear Models, Least Squares and Alternatives (New York: Springer)

[31] Ravallion, M. (1988): "Expected Poverty under Risk-induced Welfare Variability" Economic Journal, 98.

[32] Singh, I., Squire, L. and Strauss, J. (1986): Agricultural Household Models: Extensions, Applications and Policy, (Baltimore: Johns Hopkins University Press).

[33] van de Walle, D. and Nead, K. (1995): Public Spending and the Poor: Theory and Evidence (Oxford, Oxford University Press).

[34] Zellner A. (1962): "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias", Journal of the American Statictical Association 70: 631-639.

[35] Zellner, A. (1970): "Estimation of Regression Relationships containing Unobservable Variables", International Economic Review 11: 441-454.

7 Appendix

In this appendix we derive the best linear predictors Y_X^* and Y_W^* subject to unbiasedness constraints. We also show that Y_W^* has the lowest MSE. Throughout the section we make use of the following results.

$$\Sigma_X = \beta\beta'(\rho^2 + \sigma_{\epsilon\epsilon}) + \Omega where\ \rho^2 = \gamma' ZZ'\gamma \tag{A1}$$

$$E(XZ') = \beta\gamma' ZZ' \tag{A2}$$

$$E(ZY') = ZZ'\gamma \tag{A3}$$

$$E(XY) = \beta(\rho^2 + \sigma_{\epsilon\epsilon}) \tag{A4}$$

$$var(Y) = \rho^2 + \sigma_{\epsilon\epsilon} \tag{A5}$$

7.1 The *X* - *predictor*

The best linear predictor of Y using X variables obtained subject to a sample unbiasedness constraint is a function $a'X$ with coefficients vector a chosen to minimise the Lagrangean:[45]

$$L = E[Y - a'X]^2 - \lambda E[a'X - Y]$$

The constraint can be written as:

$$E(a'\beta Y - Y) = 0$$

$$i.e.\ a'\beta - 1 = 0$$

Minimisation of L is thus equivalent to minimisation of L':

$$L' = E[YY' - 2a'XY + a'XX'a] - 2\lambda(a'\beta - 1)$$

Taking first derivatives with respect to a, we get:

$$E[-XY + XX'a] - \lambda\beta = 0$$

$$i.e.\ a = \Sigma_X^{-1}[\lambda\beta + E(XY)]$$

Using (A4) we obtain:

$$a = (\lambda + \rho^2 + \sigma_{\epsilon\epsilon})\Sigma_X^{-1}\beta \tag{*}$$

Pre-multiplying (*) by β' noting that $\beta' a = 1$, we have:

$$1 = (\lambda + \rho^2 + \sigma_{\epsilon\epsilon})\beta'\Sigma_X^{-1}\beta$$

$$i.e.\ \lambda = (\beta'\Sigma_X^{-1}\beta) - \rho^2 - \sigma_{\epsilon\epsilon})$$

and substituting for λ in (*) we get :

$$a = (\beta'\Sigma_X^{-1}\beta)^{-1}\Sigma_X^{-1}\beta$$

$$i.e.\ Y_X^* = (\beta'\Sigma_X^{-1}\beta)^{-1}\beta'\Sigma_X^{-1}X$$

[45] See Goldberger (1962) and Abul Naga (1996).

7.2 The W - *predictor*

We wish to choose a vector b such that:

$$E[b'W - Y]^2$$

is minimum subject to the unbiasedness condition for random variables:

$$E[b'W - Y] = 0 \tag{UC}$$

Write the $W - predictor$ as:

$$Y_W^* = b_1'X + b_2'Z$$

We have:

$$Y_W^* = b_1'(\beta\gamma'Z + \beta\epsilon + u) + b_2'Z$$

thus the unbiasedness condition is equivalent to:

$$E[(b_1'\beta\gamma'Z + b_1'\beta\epsilon + b_1'u) + (b_2'Z) - (\gamma'Z + \epsilon)] = 0 \tag{UC'}$$

where we have made use of the causal equation $Y = \gamma'Z + \varepsilon$. (UC') is thus equivalent to:

$$b_1'\beta\gamma' + b_2' - \gamma' = 0 \tag{UC''}$$

On the basis of (UC") we have replaced b_2' by $\gamma' - b_1'\beta\gamma'$. Our problem can now be stated as one of choosing a vector b_1 to minimise the MSE criterion:

$$E[b_1'X + (\gamma' - b_1'\beta\gamma')Z - Y]^2$$

Taking first derivatives, first order conditions imply:

$$E[XX'b_1 + XZ'\gamma - 2\beta\gamma'ZX'b_1 - XY - \beta\gamma'ZZ'\gamma + \beta\beta'b_1\gamma'ZZ'\gamma + YZ'\gamma\beta] = 0$$

Using (A1)-(A4) above, we obtain:

$$b_1 = \Sigma_\nu^{-1}\beta\sigma_{\epsilon\epsilon}$$

Substituting for b_1 in (UC") we obtain the following expression for b_2:

$$b_2 = (1 - \beta'\Sigma_\nu^{-1}\beta\sigma_{\epsilon\epsilon})\gamma$$

and therefore:

$$Y_W^* = (\sigma_{\epsilon\epsilon}\beta'\Sigma_\nu^{-1})X + (1 - \sigma_{\epsilon\epsilon}\beta'\Sigma_\nu^{-1}\beta)\gamma'Z$$

7.3 Comparison of Mean Square Errors of the Three Predictors

Let θ be a scalar such that

$$\sigma_{\epsilon\epsilon}\beta'\Sigma_{\nu}^{-1} = \theta(\beta'\Sigma_X^{-1}\beta)^{-1}\beta'\Sigma_X^{-1} \tag{A6}$$

i.e. θ is the scalar which maps the coefficients of Y_X^* on those of X variables in Y_W^*. Post multiplying (A6) by β, we obtain

$$\sigma_{\epsilon\epsilon}\beta'\Sigma_{\nu}^{-1}\beta = \theta$$

It follows that Y_W^* can be written as in (26).

Using theorem A.18 of Rao and Toutenburg (1995: 291) we can invert Σ_{ν} to obtain:

$$\Sigma_{\nu}^{-1} = \Omega^{-1} - \frac{\sigma_{\epsilon\epsilon}\Omega^{-1}\beta\beta'\Omega^{-1}}{1+\sigma_{\epsilon\epsilon}\beta'\Omega^{-1}\beta}$$

from which it follows that:

$$\theta = \sigma_{\epsilon\epsilon}\beta'\Sigma_{\nu}^{-1}\beta = \frac{\sigma_{\epsilon\epsilon}\beta'\Omega^{-1}\beta}{1+\sigma_{\epsilon\epsilon}\beta'\Omega^{-1}\beta} \tag{A7}$$

i.e. $0 < \theta < 1$. The $W - predictor$ is the best combination of X and Z variables which minimises MSE subject to the unbiasedness condition (UC"). In particular, it will dominate any other convex combination of Y_X^* and Y_Z^* in the MSE sense.

It is instructive to consider the two limiting cases of Y_W^*, that is when $\theta \to 0$ and when $\theta \to 1$. From (A7) it can be seen that as $\sigma_{\epsilon\epsilon} \to \infty$, $\theta \to 1$, and that as $\sigma_{\epsilon\epsilon} \to 0$, $\theta \to 0$. That is, the more Z accounts for the variation in Y, the higher will be the weight on Z variables in Y_W^*.

Computations of $MSEs$ give us the following expressions:

$$MSE(Y_W^*) = (1-\theta)^2[\sigma_{\epsilon\epsilon}(\sigma_{\epsilon\epsilon}\beta'\Omega^{-1}\beta + 1)] \tag{A8}$$

$$MSE(Y_X^*) = (\beta'\Omega^{-1}\beta)^{-1} \tag{A9}$$

$$MSE(Y_Z^*) = \sigma_{\epsilon\epsilon} \tag{A10}$$

These relations are used to establish that $MSE(Y_X^*) > MSE(Y_W^*)$, and that $MSE(Y_Z^*) > MSE(Y_W^*)$.

Table 1: MIMIC Model Specification

Variable	Variable name	Description
Y		Unobservable logged per capita permanent income
X1	LNPCI	Logged per capita disposable income
X2	LNPCE1	Logged per capita expenditure including current durable and housing expenditures
	LNPCE2	Logged per capita expenditure including imputed rent from durable stock and current housing expenditures
	LNPCE3	Logged per capita expenditure including imputed rents from durable and housing stock
Z1	LNN	Logged household size
Z2	CHILDP	Proportion of children in household size
Z3	EDUHD	Educational status of household head
Z4	PRIMP	Proportion of household labour force in primary occupation (agriculture, fisheries, forestry)
Z5	AREAPC	Per capita housing floor area
Z6	ELEC	Dummy for whether house electrified
Z7	CULTPC	Cultivable land per capita
Z8	BWTV	Per capita black and white TVS
Z9	WATCH	Per capita watches
Z10	DEPOSIT	Per capita value of savings
Z11	PASSETV	Per capita value of productive assets

Notes: In the corrected measures of consumption housing is imputed at 6% of house value, whilst durables are imputed at 12% of the current value of the household durable stock.

Table 2: Estimation Results for the Income/Consumption MIMIC Model

		SICHUAN 1990			JIANGSU 1990		
Var	Coef	LNPCE1	LNPCE2	LNPCE3	LNPCE1	LNPCE2	LNPCE3
LNPCI	β_1	1.000	1.000	1.000	1.000	1.000	1.000
LNPCE	β_2	0.966	0.985	1.001	0.956	0.979	1.040
LNN	γ_1	-0.100 (7.486)	-0.105 (8.046)	-0.133 (11.425)	-0.109 (5.351)	-0.110 (5.582)	-0.119 (7.579)
CHILDP	γ_2	-0.240 (8.834)	-0.236 (8.896)	-0.279 (11.825)	-0.224 (5.884)	-0.207 (5.626)	-0.196 (6.687)
EDUHD	γ_3	0.002 (2.770)	0.002 (2.853)	0.002 (3.108)	0.008 (6.500)	0.008 (6.618)	0.008 (8.127)
PRIMP	γ_4	-0.231 (9.897)	-0.226 (9.925)	-0.241 (11.906)	-0.247 (8.663)	-0.244 (8.848)	-0.181 (8.227)
AREAPC	γ_5	0.006 (13.390)	0.006 (13.830)	0.006 (16.160)	0.009 (17.992)	0.009 (18.585)	0.011 (25.276)
ELEC	γ_6	0.070 (5.089)	0.070 (5.260)	0.052 (4.395)	0.157 (7.163)	0.157 (7.405)	0.147 (8.710)
CULTPC	γ_7	0.086 (11.420)	0.081 (11.108)	0.071 (10.858)	0.129 (12.576)	0.118 (11.866)	0.076 (9.636)
BWTV	γ_8	0.640 (17.863)	0.618 (17.640)	0.605 (19.242)	0.363 (7.065)	0.368 (7.409)	0.373 (9.382)
WATCH	γ_9	0.294 (17.447)	0.304 (18.430)	0.306 (20.644)	0.436 (15.071)	0.449 (15.997)	0.398 (17.415)
DEPOSIT	γ_{10}	0.040 (14.877)	0.040 (15.176)	0.029 (12.529)	0.012 (7.539)	0.011 (7.496)	0.009 (7.288)
PASSETV	γ_{11}	0.017 (9.192)	0.017 (9.303)	0.019 (11.574)	0.007 (3.829)	0.006 (3.782)	0.008 (5.840)
	ω_{11}	0.039	0.043	0.051	0.075	0.084	0.107
	ω_{22}	0.039	0.034	0.011	0.113	0.093	0.027
AGFI		0.977	0.977	0.957	0.979	0.973	0.949

igure 1: Path Diagram for the MIMIC Model

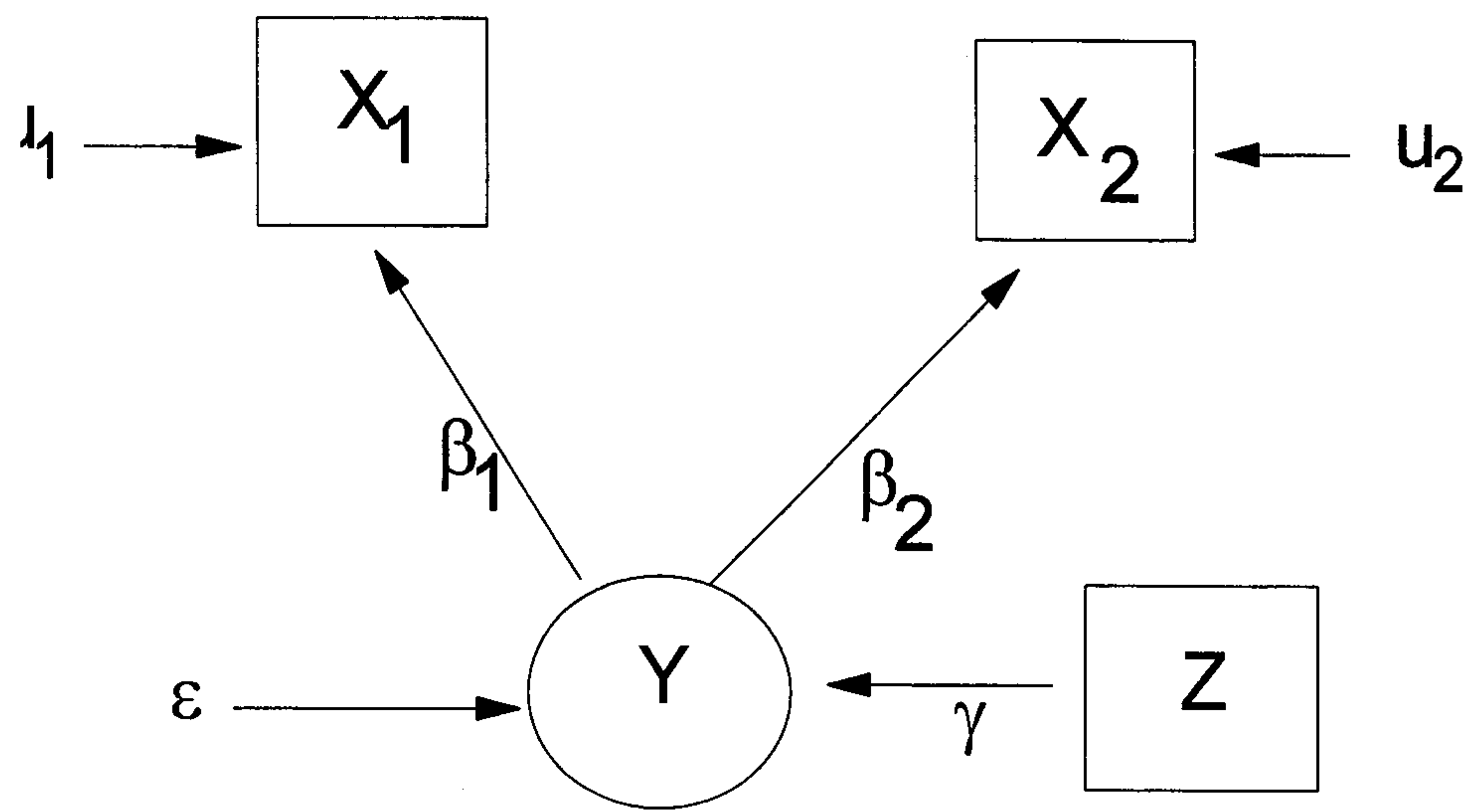

Figure 2: Density Plots for Sichuan

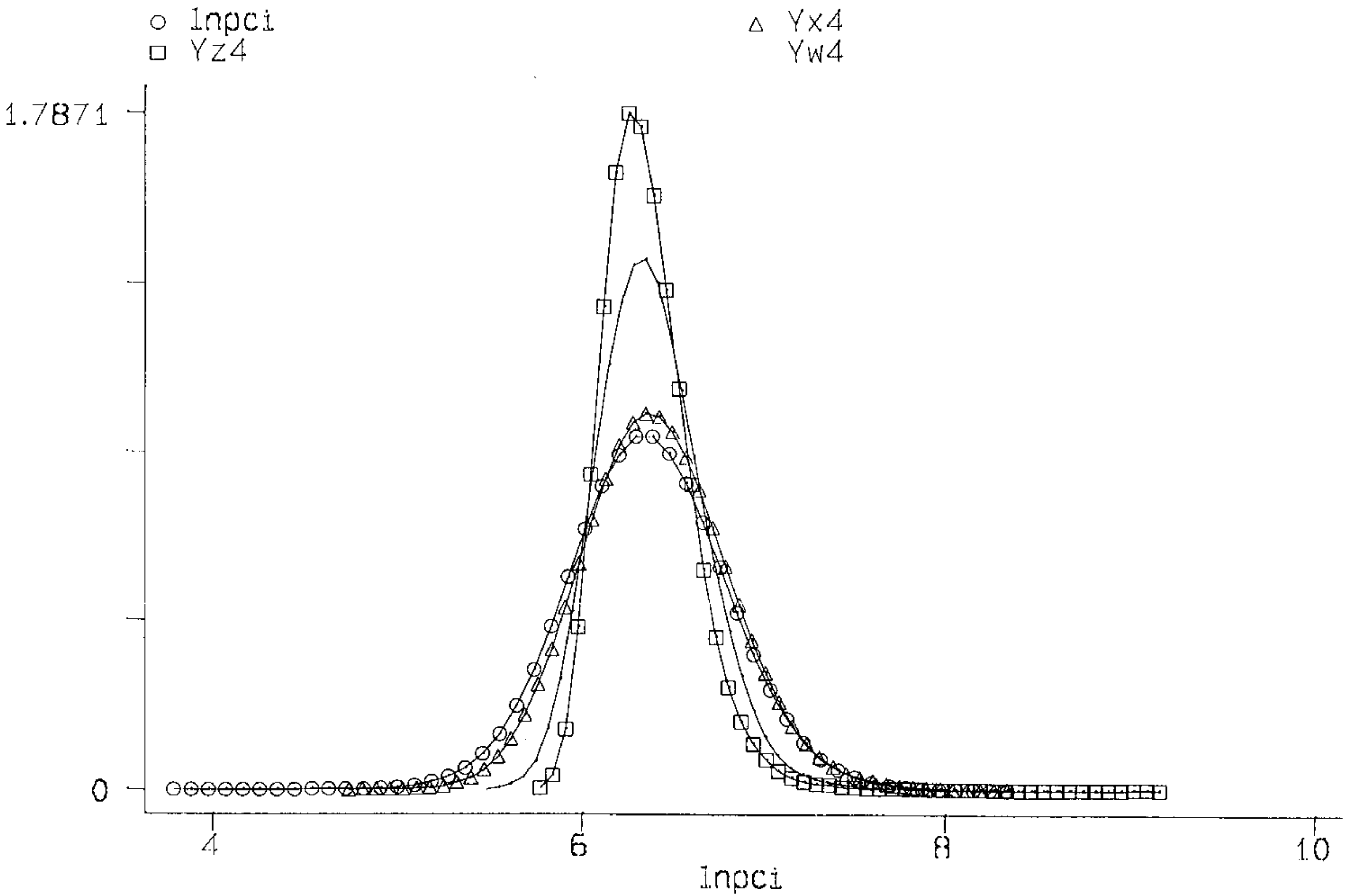

Figure 3: Density Plots for Jiangsu