the review of income and wealth

Review of Income and Wealth
Series 57, Number 1, March 2011

ESTIMATING SPATIAL CONSUMER PRICE INDICES THROUGH ENGEL CURVE ANALYSIS

BY DIPANKOR COONDOO, AMITA MAJUMDER AND SOMNATH CHATTOPADHYAY*

Indian Statistical Institute

In this paper we propose a method of estimating spatial multilateral price index numbers from cross-section consumer expenditure data on different items using Engel curve analysis. The novelty of the procedure is that it overcomes the problem of data inadequacy, a problem that is shared by most of the developing countries. The procedure does not require item-specific price/unit value data and price index numbers can be calculated from consumer expenditure data grouped by per capita income/total consumer expenditure class in a situation where unit level data are not available. To illustrate the method, we use published state-specific data of the 50th round (1993–94) and 55th round (1999–2000) consumer expenditure surveys of India's National Sample Survey Organization (NSSO) and calculate the spatial consumer price index numbers for 15 major states of India, with All-India taken as base, separately for the rural and the urban sector for each round.

INTRODUCTION

Appropriate consumer price index numbers are essential for comparison of real income levels or consumption patterns over time, across regions or across population groups. When more than two (regions/countries/population) groups are involved in a comparison of price or real income levels, the price index number problem is resolved in one of two major ways. The simpler and straightforward approach is to use a set of binary price index numbers and make pair-wise comparisons. Examples of this approach are Sen (1976), Bhattacharya, Joshi, and Roychowdhury (1980), Bhattacharya, Chatterjee, and Pal (1988), Coondoo and Saha (1990), Deaton (2003), and Deaton and Tarozzi (2005). Use of the binary comparisons approach, however, does not guarantee transitivity of price level comparisons except under unrealistic assumptions.

A second approach is to have a multilateral price level comparison, whereby a set of internally consistent price index numbers, popularly known as Purchasing Power Parities (PPPs), are constructed on the basis of a set of group-specific price and quantity data for a common set of commodities (see Geary, 1958; Khamis, 1972; Kravis *et al.*, 1978; Balk, 1996; Hill, 1997; Prasada Rao, 1997; Diewert, 1999; Neary, 2004). As in the case of binary price index numbers, computation of a set of multilateral price index numbers requires price and quantity data of uniform quality, which is often rather difficult to obtain. To resolve the data problems arising from quality variation of items across groups and from gaps in the avail-

Note: We thank Ms Saswati Das for technical assistance. We also thank two anonymous referees for their helpful comments and for drawing our attention to several recent references. The usual disclaimer applies.

*Correspondence to: Somnath Chattopadhyay, Economic Research Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata 700108, India (pikluchatterjee@gmail.com).

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main St, Malden, MA, 02148, USA.

able price data, the Country Product Dummy (CPD) methodology was proposed (Summers, 1973). The CPD procedure, which is essentially a hedonic approach, offers a regression analysis-based methodology for constructing multilateral price index numbers (see Kokoski *et al.*, 1999; Prasada Rao, 2001). A large part of the literature on multilateral price index numbers today is concerned with the construction of PPPs from commodity-specific price and quantity/expenditure share data using the CPD methodology. Application of CPD and similar methods to household level data has been proposed in recent works of Aten and Menezes (2002) and Coondoo *et al.* (2004).

In this paper we propose a procedure for estimating regional consumer price index numbers based on the estimation of item-specific region-wise Engel curves. Given the problem of data inadequacy in developing countries, the basic question we try to answer here is: "Is it possible to find a method of estimation of a set of spatial consumer price index numbers using Engel curves that does not require data (i) at the household level, or (ii) on prices/unit values of goods and services consumed?"

Costa (2001) and Hamilton (2001) pioneered the use of Engel curves in the context of Consumer Price Indices (CPI). The basic idea underlying these studies is that if a given CPI is an accurate measure of cost of living, the CPI-deflated Engel curves (log-linear/ log-quadratic food share equations) estimated at different time points should coincide and temporal drift in CPI-deflated Engel curves will reflect systematic bias in measurement of CPI (Barrett and Brzozowski, 2008). There have been various extensions of this approach in terms of specification (introduction of household demographics: Logan (2008); flexible semi-parametric Engel curves: Beatty and Larsen (2005), Larsen (2007); flexible Almost Ideal Demand System (AIDS): Barrett and Brzozowski (2008)) as well as application (in the context of regional price index: Papalia (2006); PPP: Almas (2008)). All these studies, however, are based on pooled time series of household-level cross-section data. Also, a basic data requirement for these kinds of exercises is availability of estimates of relative price changes over time /region.

The data requirement for the procedure proposed here is minimal. It does not require region-specific data on prices of individual items. Formally, given a system of demand functions derived from an underlying cost function, it may be possible to derive estimates of the parameters appearing in the cost function from the estimated demand functions. One should then, in principle, be able to estimate the True Cost of Living Index (TCLI) number corresponding to a specified utility level. When a consumer expenditure dataset covers regions facing different price situations, the region-specific Engel curves for individual items estimated from such a dataset contain information about regional price level differentials, which if retrieved, can be used to construct regional TCLIs. This kind of procedure has already been suggested by Fry and Pashardes (1989). They investigate the conditions under which the Törnqvist price index number can be a reasonable approximation to the TCLI underlying a Price Independent Generalized Log linear (PIGLOG) demand system. Using the decomposition of the TCLI under PIGLOG as the sum of a basic index (the cost of living index at some minimum level of consumer expenditure) and a marginal index (Deaton and Muellbauer, 1980a), they apply the Törnqvist method to estimate the TCLI in a systems framework. It makes explicit reference to expenditure levels, commodity prices, and household characteristics in the context of the AIDS of Deaton and Muellbauer (1980b) and the Translog model of demand, both members of the PIGLOG class.

Here we propose a more general approach to estimating TCLI based on a two-component decomposition (a basic index and a marginal index) of the TCLI underlying a quadratic PIGLOG system. The proposed procedure has several useful features:

(1) It overcomes the already mentioned problem of data inadequacy: it does not require item-specific price or unit-value data, and more importantly, allows inclusion of items of expenditure for which separate data on price and quantity are usually not recorded (e.g., meals away from home, expenditure on recreation, educational and health services, etc.).

(2) The method is essentially based on (single equation) Engel curve analysis and hence it is computationally simpler and no explicit algebraic form for the coefficients of the Engel curves (which are functions of prices) is required. An underlying assumption here is that the Engel curve is quadratic logarithmic (in budget share form)[1] and the form is the same for all the regions being compared.

(3) It is not necessary that all items must be consumed in all regions.[2]

(4) The procedure does not require household level expenditure data and can be applied to consumer expenditure data grouped by per capita income/total consumer expenditure class.

Estimation of TCLI using this method involves three steps. In the first step, a set of item-specific Engel curves, relating item-specific budget shares to the logarithm of per capita income/total consumer expenditure, are estimated for each region. The first component of the TCLI (the basic index) is estimated in the second step based on a pooled regression over items and regions. In the third step, the marginal index and the TCLI are estimated. In an illustrative exercise reported here, the procedure is applied to grouped expenditure compiled from the published reports of the Indian National Sample Survey Organization's (NSSO) 50th and 55th rounds of consumer expenditure surveys conducted during the years 1993–94 and 1999–2000. The data cover 15 major states of India, and state-specific TCLIs (with All-India taken as base) are estimated separately for the rural and urban sectors for each of the two rounds.

The Proposed Procedure

The cost function underlying Quadratic Logarithmic (QL) systems (e.g., the Quadratic Almost Ideal Demand System (QUAIDS) of Banks *et al.* (1997) and the Generalized Almost Ideal Demand System (GAIDS) of Lancaster and Ray (1998)) is of the form

$$C(u, p) = a(p).\exp\left(\frac{b(p)}{(1/\ln u) - \lambda(p)}\right), \tag{1}$$

[1]This is the most popular and commonly used form of budget share equation in the literature.

[2]A requirement, however, is that the set of items of the base region must be the union of all items consumed in different regions, as will be seen in the estimation procedure.

where p is the price vector, $a(p)$ is a homogeneous function of degree one in prices, $b(p)$ and $\lambda(p)$ are homogeneous functions of degree zero in prices, and u denotes the level of utility. By Shephard's lemma, the budget share functions corresponding to the cost function (1) are of the form

$$(2) \quad w_i = \frac{\partial \ln a(p)}{\partial \ln p_i} + \frac{\partial \ln b(p)}{\partial \ln p_i} \ln \frac{y}{a(p)} + \frac{1}{b(p)} \frac{\partial \lambda(p)}{\partial \ln p_i} \left(\ln \frac{y}{a(p)} \right)^2, i = 1, 2, \ldots, n,$$

$$\text{or, } w_i = a_i(p) + b_i(p) \ln \frac{y}{a(p)} + \frac{\lambda_i(p)}{b(p)} \left(\ln \frac{y}{a(p)} \right)^2, i = 1, 2, \ldots, n,$$

where y denotes nominal per capita income and i denotes item of expenditure.

The corresponding TCLI in logarithmic form comparing price situation p^1 with price situation p^0 is given by

$$(3) \quad \ln P(p^1, p^0, u^*) = [\ln a(p^1) - \ln a(p^0)] + \left[\frac{b(p^1)}{1/\ln u^* - \lambda(p^1)} - \frac{b(p^0)}{1/\ln u^* - \lambda(p^0)} \right],$$

where u^* is the reference utility level. The first term of the R.H.S. of (3) is the logarithm of the basic index (measuring the cost of living index at some minimum benchmark utility level) and the second term is the logarithm of the marginal index. Note that for $p^1 = \theta p^0$, $\theta > 0$, $a(p^1) = \theta a(p^0)$, so that the basic index takes a value θ and hence, may be interpreted as that component of TCLI that captures the effect of uniform or average inflation on the cost of living. On the other hand, for $p^1 = \theta p^0$, $\theta > 0$, $b(p^1) = b(p^0)$ and $\lambda(p^1) = \lambda(p^0)$, the marginal index takes a value of unity. Hence, the marginal index may be interpreted as the other component of TCLI that captures the effect of a change in the relative price structure. If prices are normalized such that $b(p^0) = 1$ and $\lambda(p^0) = 1$, the TCLI for a reference utility level u^* becomes

$$(4) \quad P(p^1, p^0, u^*) = \frac{a(p^1)}{a(p^0)} \exp\left[\frac{b(p^1)}{1/\ln u^* - \lambda(p^1)} - \frac{1}{1/\ln u^* - 1} \right].$$

To build up our procedure, look at the region specific Engel curves of the form (2). Let p^r denote the price vector of region r, $r = 0,1,2 \ldots R$. Then, from (2), the budget share equations for region r can be written as

$$w_{ir} = a_i(p^r) + b_i(p^r) \ln \frac{y_r}{a(p^r)} + \frac{\lambda_i(p^r)}{b(p^r)} \left(\ln \frac{y_r}{a(p^r)} \right)^2$$

or

$$(5) \quad w_{ir} = \alpha_{ir} + \beta_{ir} \ln\left(\frac{y_r}{P_r} \right) + \gamma_{ir} \left(\ln\left(\frac{y_r}{P_r} \right) \right)^2, \quad i = 1, 2, \ldots, n.$$

In (5), P_r denotes the price level for region r, homogeneous of degree one in prices of the region.[3] The parameters α_{ir}, β_{ir}, γ_{ir}, P_r are functions of the price vector and are parameters for a given cross-sectional data situation where prices are fixed.

The budget share curves in (5), which are quadratic in logarithm of income, correspond to those of QUAIDS and GAIDS having underlying *cost functions* of the forms

$$C(u, p^r) = a(p^r)\exp\left(\frac{b(p^r)}{(1/\ln u) - \lambda(p^r)}\right), \tag{6}$$

where $b(p^r) = \prod_{i=1}^{n} p_{ir}^{\beta_i}$, $\lambda(p^r) = \sum_{i=1}^{n} \lambda_i \ln p_{ir}$ for both QUAIDS and GAIDS, and

$\ln a(p^r) = \alpha_0^* + \sum_{i=1}^{n} \alpha_i^* \ln p_{ir} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij} \ln p_{ir} \ln p_{jr}$ for QUAIDS, and

$\ln a(p^r) = \alpha_0^* + \frac{1}{1-\sigma}\ln\left(\sum_{i=1}^{n} \alpha_0^* \ln p_{ir}^{(1-\sigma)}\right) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij} \ln p_{ir} \ln p_{jr}$ for GAIDS.

The parameters of the budget share functions derived from (6) using the above expressions relate to the parameters in (5) as follows (see Banks *et al.*, 1997; Lancaster and Ray, 1998):

$$\alpha_{ir} = \alpha_i^* + \sum_{k=1}^{n} \delta_{ik} \ln p_{kr} \text{ for QUAIDS,}$$

$$\alpha_{ir} = \frac{\alpha_i^* p_{ir}^{1-\sigma}}{\sum_{k=1}^{n} \alpha_k^* p_{kr}^{1-\sigma}} + \sum_{k=1}^{n} \delta_{ik} \ln p_{kr} \text{ for GAIDS and}$$

$$\beta_{ir} = \beta_i,\ \gamma_{ir} = \frac{\lambda_i}{b(p^r)},\ P_r = a(p^r) \text{ for both QUAIDS and GAIDS.}$$

If the QUAIDS or the GAIDS system is estimated using an appropriate panel dataset that contains adequate price variation for individual regions, then reliable estimates of all the parameters of the system will be obtained. Using these and the formula given in (4), TCLIs measuring cross-sectional regional price level differentials may then be easily calculated.

Our objective here, however, is to explore whether or not such regional price level differences can be estimated using data from a single cross-section, without item level information on prices and without specifying the algebraic forms of $a(p^r)$, $b(p^r)$, and $\lambda(p^r)$ explicitly. Thus, the method is essentially based on (single equation) Engel curve analysis as opposed to the systems approach.

To estimate $a(p^r)$, i.e., P_r, rewrite the budget share functions (5) as

$$w_{ir} = (\alpha_{ir} - \beta_i\pi_r - \gamma_{ir}\pi_r^2) + (\beta_i - 2\gamma_{ir}\pi_r)y_r^* + \gamma_{ir}y_r^{*2}, \tag{7}$$

[3]Ratios of P_r's will measure the basic price index number of a region with the other region taken as base.

where $y_r^* = \ln(y_r)$, $\pi_r = \ln(P_r)$.[4] As mentioned earlier, for a single cross-sectional dataset corresponding to a given price situation, α_{ir}, γ_{ir}, β_i and π_r are parameters to be estimated from the given data. Now, (7) is a set of $(R + 1)$ complete systems of *n-commodity* Engel curves, which are non-linear in parameters. With an appropriate stochastic specification, (7) will be a large system of non-linear SUR equations, which can, in principle, be estimated from a given set of data. However, such a simultaneous estimation of all the equations of (7) under parametric restrictions would be extremely difficult, if not impossible. We have, therefore, used the following alternative indirect estimation route.

As mentioned earlier, our suggested procedure for estimating TCLIs in (4) involves three stages. In the first stage, a set of item-specific Engel curves relating budget shares to the logarithm of income are estimated (using equation (8) defined later) for each region. In the second stage $a(p^r)$, $r = 0,1,2 \ldots, R$ is estimated. In the third stage $b(p^r)$ and $\lambda(p^r)$, $r = 1,2 \ldots, R$ are estimated using the normalization $b(p^0) = \lambda(p^0) = 1$ (where p^0 denotes the price vector of the base region). Using these, the TCLIs are estimated for a given reference level of utility of the base region. It may be emphasized that $a(p^r)$, $b(p^r)$, and $\lambda(p^r)$ are estimated as composite variables and no explicit algebraic forms for these functions are assumed. The three stages are described below in detail.

Stage 1

Estimate the following log-quadratic budget share function, which is in the form of a linear regression equation:

$$w_{irj} = a_{ir} + b_{ir} y_{rj}^* + c_{ir} y_{rj}^{*2} + \varepsilon_{irj}, \tag{8}$$

where the subscript j $(= 1,2, \ldots, H_r)$ denotes the per capita income/total consumer expenditure (PCE) class of a region, ε_{irj} is a random disturbance term, and a_{ir}, b_{ir}, c_{ir} are the parameters.[5]

Stage 2

Let $\hat{a}_{ir}$, $\hat{b}_{ir}$, and $\hat{c}_{ir}$ be the estimates of a_{ir}, b_{ir}, c_{ir}.[6] Given the estimates $\hat{a}_{ir}$, $\hat{b}_{ir}$, $\hat{c}_{ir}$, from (7) and (8), we have say

[4]Note that the β_{ir}'s have been replaced by β_i's. That is, they do not have any *region effect*, or to put it differently, they are independent of prices. This is in line with the specifications in QUAIDS and GAIDS. Also, a justification for this form of budget share can be found in Banks *et al.* (1997).

[5]The subscript j would denote the j-th sample household of region r, when household level expenditure data are used to estimate these Engel curves. Note that as compared to household level data, for grouped data H_r will be small, leading to smaller degrees of freedom.

[6]We have estimated the item-specific Engel curves for regions from grouped expenditure data by single-equation weighted least squares, using the estimated population proportion of individual PCE classes as weights. This should take care of the heteroscedasticity arising out of grouping of data. The heteroscedasticity problem due to dependence of the error variance on y^*, if any, should be largely taken care of by the use of Engel curve formulation in budget share form in our case due to the grouped nature of expenditure data used. However, for household level data, the issue of heteroscedasticity needs to be addressed appropriately (see Deaton, 1997).

(9a) $$\hat{c}_{ir} = \gamma_{ir} + e_{ir}^{c},$$

(9b) $$\hat{b}_{ir} = (\beta_i - 2\gamma_{ir}\pi_r) + e_{ir}^{b},$$

(9c) $$\hat{a}_{ir} = \left(\alpha_{ir} - \beta_i\pi_r - \gamma_{ir}\pi_r^2\right) + e_{ir}^{a},$$

where e_{ir}^{b}, e_{ir}^{c}, e_{i0}^{b} are the errors in estimation of the parameters using equations (9a) and 9(b). The π_r's are then estimated as follows:

(10) $$\hat{b}_{ir} - \hat{b}_{i0} = \pi_0(2\hat{c}_{i0}) - \pi_r(2\hat{c}_{ir}) + e_{ir}, \quad i = 1, 2, \ldots, n; \quad r = 1, 2, \ldots, R$$

where e_{ir} is a composite error term, which is a linear combination of the individual errors e_{ir}^{b}, e_{ir}^{c}, e_{i0}^{b} and e_{i0}^{c}. Thus, the regression error is assumed to be present only because of estimation errors in the first stage, and since the first stage parameters are consistently estimated, asymptotically equation (10) would hold exactly. However, typically, the error e_{ir} and the explanatory variables will be correlated. As an approximation if we treat this as a multivariate errors-in-variables set-up, consistent estimates of the π_r's can be obtained from[7]

(11) $$\hat{\pi} = \frac{1}{2}\left(\hat{C}'\Sigma^{-1}\hat{C} - NE_C'\Sigma^{-1}E_C\right)^{-1}\left(\hat{C}'\Sigma^{-1}\hat{B} - NE_C'\Sigma^{-1}E_B\right),^{7}$$

where $\hat{C}$ is the matrix of explanatory variables in (10), $\hat{B}$ is the vector of dependent variables in (10), E_C is the matrix of estimation errors in C's, E_B is the vector of estimation errors in B, N $(= nR)$ is the sample size, and Σ is the variance–covariance matrix of the error terms in (10). From the consistency property of OLS estimates it can be shown that asymptotically $E_C'\Sigma^{-1}E_C$ will converge to a null matrix and $E_C'\Sigma^{-1}E_B$ will converge to a null vector, thus yielding the usual GLS estimates.[8]

Three points are noteworthy so far as this estimation procedure is concerned. First, the procedure does not necessarily require that the number of items of expenditure and the composition of the set of items of expenditure be same for all regions. Second, whereas in the literature π_0 is not estimated and is fixed exogenously (Deaton and Muellbauer, 1980a; Banks *et al.*, 1997) here an estimate of π_0 is obtained from the estimation process itself. Finally, as already pointed out, the estimates of π_r's obtained by the above procedure are conditional upon the fact that the β_i's in equation (7) do not have any region effect.

Stage 3

Once the estimates of π_r's (and hence of ln $a(p^r)$'s) are obtained, the next step involves the estimation of $b(p^r)$ and $\lambda(p^r)$ for every r. However, this set of estimated parameters would not suffice for calculating TCLIs reflecting regional price level

[7]See Deaton (1997).

[8]Alternatively, an estimate of the variance–covariance structure can be obtained using the bootstrap method. The necessary adjustment for the correction of bias can thus be made. We have not attempted this here.

differentials by equation (4). This is because without detailed price information $b(p^r)$ and $\lambda(p^r)$ cannot be calculated, even though estimates of $\gamma_{ir}\left(=\frac{\lambda_i}{b(p^r)}\right)$ are available.

To resolve this problem, we proceed as follows. Treat region $r = 0$ as the base region and take the utility levels of the base region as reference utility levels. Using equation (6) and the normalization $b(p^0) = \lambda(p^0) = 1$, the money metric utility u_0^h of a household of the base region that has nominal per capita income $y_{0h}(= C(u_0^h, p^0))$ is given by

$$\frac{1}{\ln u_0^h} = \frac{1}{\ln \frac{y_{0h}}{a(p^0)}} + 1. \tag{12}$$

Now, combine equations (6) and (12) to obtain for region r

$$\frac{1}{\ln\left(\frac{y_{rh}}{a(p^r)}\right)} = \frac{1}{b(p^r)}\left(\frac{1}{\ln \frac{y_{0h}}{a(p^0)}} + 1\right) - \frac{\lambda(p^r)}{b(p^r)}, \tag{13}$$

where y_{rh} denotes the per capita nominal income required by a household of region r to have u_0^h utility level.[9]

Using the relationship (13) we propose estimation of $b(p^r)$ and $\lambda(p^r)$ from the following regression equation[10]

$$n\left(\frac{1}{\ln \bar{y}_{rq} - \hat{\pi}_r}\right) = \frac{1}{b(p^r)}\left(\frac{1}{\ln \bar{y}_{oq} - \hat{\pi}_0} + 1\right) - \frac{\lambda(p^r)}{b(p^r)} + error, \tag{14}$$

where $\hat{\pi}_r$ is the estimate of $ln\ a(p^r)$, which has already been obtained in the second stage of estimation and $\bar{y}_{rq}$ $(q = 1, 2, .., Q)$ is the q-th quantile of PCE. Since the number of expenditure groups may vary across regions, we estimate the regression equations using region-specific data on PCE by quantiles, viz. $(\bar{y}_{rq}; q = 1, 2, \ldots, Q)$.[11] It is assumed that all the q-th quantile households of a given region have comparable utility levels.

Here again, it may be noted that both the regressor and the regressand contain estimated values of π's and hence are measured with error. However, under some mild conditions, the use of OLS can be justified.[12] Once estimates of π_r, $b(p^r)$, and $\lambda(p^r)$ are obtained this way, consistent estimates of the TCLIs for regions corresponding to given quantile levels, may be calculated using equation (4).

[9]See Appendix 1 for derivation of equation (13).

[10]See Appendix 2 for a justification of the regression set-up.

[11]This may produce noisy estimates if Q is not large. Thus, while for household level data there may not be any problem, for grouped data this may typically be the case. In our empirical exercise, decile values of PCE have been used.

[12]See Appendix 2 for a justification of the use of OLS.

Data and Results

As an illustrative application, the procedure presented above has been applied to a set of consumer expenditure data, aggregated by monthly per capita total consumer expenditure class (grouped data), compiled from the published reports of the Indian National Sample Survey Organisation's (NSSO) 50th and 55th rounds of consumer expenditure surveys conducted during the years 1993–94 and 1999–2000.[13] Data for 15 major Indian states are included in this dataset and these states are treated as regions in the empirical exercise. The states included are Andhra Pradesh (AP), Assam (AS), Bihar (BI), Gujarat (GU), Haryana (HA), Karnataka (KA), Kerala (KE), Madhya Pradesh (MP), Maharashtra (MA), Orissa (OR), Punjab (PU), Rajasthan (RA), Tamil Nadu (TN), Uttar Pradesh (UP) and West Bengal (WB). The base region is "All-India," which relates to the data for all states combined.[14] The dataset covers 17 and 19 item expenditure categories for the 50th and 55th rounds. These categories, a number of which contain non-food and service items, jointly comprise total consumer expenditure.[15] The exercise has been done separately for rural and urban sectors. For each sector, state-specific TCLIs corresponding to the level of living of households having the All-India median level PCE are estimated taking All-India as the base region.

Estimated state-specific TCLIs (with All-India as base) for NSSO 50th and 55th rounds are presented in Table 1 separately for the rural and urban sectors.[16] The reference utility level has been taken to be the utility value (obtained from equation (6)) at the median level of expenditure for the reference region (All-India) for the respective rounds and sectors.[17] For the purpose of comparison and to judge the plausibility of our results, corresponding Törnqvist price indices estimated by Deaton (2003) using household level data for the same rounds are reproduced in Table 1. It may be seen that, by and large, our estimated TCLIs and the corresponding Törnqvist price indices of Deaton yield comparable results in terms of order of magnitude. However, our estimates are somewhat smaller for the rural and urban sectors of Bihar and Orissa and for the rural sectors of Assam, Madhya Pradesh, and West Bengal (55th round). On the other hand, for the rural sectors of Haryana, Kerala, and Punjab, our estimates are larger. In the remaining cases, the two sets of estimates do not appear to be widely different. To examine the extent of agreement or otherwise of the ordering of the states in terms of the

[13]This illustration shows that the method is applicable even when household level data are not available.

[14]It may be mentioned that the All-India estimates cover all the states of India, including those not considered here.

[15]See Appendix 3 for the list of items and description of the PCE classes.

[16]For consideration of space, the estimated Engel curves and values of other parameters estimated at intermediate stages are not presented here. These will be made available to interested readers on request. Estimated values of π_r, and coefficients of equation (14), however, are presented in Tables A1 and A2 in Appendix 4. It may be mentioned that in estimating π_r's [equation (10)], for the 50th round the number of observations is 255 and that for the 55th round is 285, with equation R^2 very close to 1 in all cases. For equation (14), as mentioned in footnote 11, the number of observations is 10 in each case and the R^2 value ranges from 0.965 to 0.999. It may also be pointed out here that the Engel curve specification does not involve household size, which, if omitted, can seriously contaminate the results when household level data are used. Items that are affected by economies of scale will have their impact on TCLI. However, in our case, because we are using grouped data, this factor is unlikely to influence our results.

[17]Change in the reference utility level may induce marginal changes in the estimates of TCLI.

TABLE 1

Price Indices for States Relative to All-India

	Price Indices from the Proposed Model				Price Indices (Törnqvist) of Deaton (2003)			
	Rural Sector		Urban Sector		Rural Sector		Urban Sector	
States	50th Round	55th Round	50th Round	55th Round	50th Round	55th Round	50th Round	55th Round
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AP	1.044	0.863	0.875	0.929	0.979	1.020	0.940	0.977
AS	0.997	0.810	1.008	1.043	1.093	1.120	1.059	1.079
BI	0.833	0.772	0.751	0.689	0.981	0.978	0.957	0.914
GU	1.142	1.098	1.016	1.206	1.165	1.111	1.052	1.074
HA	1.317	1.357	1.071	1.272	1.033	1.024	1.009	1.005
KA	0.942	0.936	0.930	1.148	1.035	1.063	0.994	1.035
KE	1.354	1.307	1.092	1.197	1.127	1.232	1.005	1.085
MP	0.865	0.781	0.908	0.892	0.942	0.952	0.948	0.927
MH	0.925	0.961	1.161	1.153	1.057	1.054	1.106	1.096
OR	0.810	0.698	0.892	0.698	0.928	0.990	0.906	0.886
PU	1.572	1.407	1.191	1.110	1.050	1.043	1.017	0.973
RA	1.159	1.049	0.977	1.072	1.055	1.067	0.997	0.991
TN	1.018	0.927	0.945	1.171	1.070	1.109	1.004	1.046
UP	0.966	0.913	0.864	0.833	0.918	0.924	0.941	0.943
WB	1.028	0.739	1.074	1.018	0.966	1.011	1.000	0.984
All-India	1.000	1.000	1.000	1.000	1	1	1	1

two price indices, Spearman's rank correlation coefficient between the ordering of the states in terms of the values of the two price indices are calculated. For the rural sector, the rank correlation coefficients work out to be 0.482 and 0.509 for the 50th and 55th rounds, respectively, the former being significant at the 6 percent level and the latter at the 5 percent level. The corresponding values for the urban sector are estimated to be 0.836 and 0.803 for the 50th and 55th rounds, both of which are significant at the 1 percent level.[18]

Conclusion

In this paper we have proposed a procedure for estimating a set of regional consumer price index numbers that are the TCLIs of a quadratic PIGLOG demand system. The procedure is based on estimating region-specific Engel curves for a set of expenditure categories. The most important contribution of the procedure is that it addresses the issue of data inadequacy, a major problem in the context of developing countries. In other words, the method works even in a situation where unit level data are not available. This procedure does not require

[18]Strictly speaking, our estimated TCLIs and Deaton's Törnqvist price indices are non-comparable for two reasons. First, while Deaton's price indices cover only those items of expenditure (food, beverage, tobacco and other intoxicants, and fuel) for which both quantity and expenditure data are available, the dataset used in our exercise includes expenditure on all the categories of consumption, including those of non-food and service items. So, the observed differences between the two sets of price indices may partly be due to the regional price differentials of *non-food* and *service item* prices. Second, Deaton's Törnqvist price indices are the estimates of TCLIs underlying a PIGLOG demand system, and therefore these indices may not correspond well with our estimates which are TCLIs underlying a more general quadratic PIGLOG demand system.

the expenditure categories to exhaust the consumer's budget. However, if the set of expenditure categories considered is exhaustive, the estimated consumer prices index numbers will be more accurate measures of the underlying true price level differentials.

The other notable features are as follows. First, as it is intimately related to the quadratic PIGLOG demand system, it has a well defined theoretical underpinning. Second, the data requirement is minimal in the sense that it can be implemented even on a set of grouped consumer expenditure data covering several regions. More importantly, region-specific separate data on quantity and price of individual consumer goods are not required for this procedure. Therefore, items of expenditure like "services consumed," and "medical expenses," for which only expenditure data are available and separate quantity and price are often not well defined, can also be included. Third, no explicit algebraic form for the coefficients of the Engel curves (which are functions of prices) is required. Finally, as the results presented here would suggest, the empirical performance of the proposed procedure is satisfactory (this is also evident from Table A1 in the Appendix, as each component of the TCLI turns out to be highly significant).

A few issues, however, need to be addressed for future application of the methodology. Some outstanding questions, which remain unresolved in the present exercise, regarding the statistical properties of the estimates at different stages, need to be explored, possibly using Monte Carlo studies.

Appendix 1

For the reference utility level u_0^h, equation (6) can be written as

$$C(u_0^h, p^r) = a(p^r)\exp\left(\frac{b-(p^r)}{(1/\ln u_0^h) - \lambda(p^r)}\right)$$

or,

$$y_{rh} = a(p^r)\exp\left(\frac{b(p^r)}{(1/\ln u_0^h) - \lambda(p^r)}\right),$$

where y_{rh} denotes the per capita nominal income required by a household of region r to have u_0^h utility level.

We thus have

$$\ln\frac{y_{rh}}{a(p^r)} = \frac{b(p^r)}{(1/\ln u_0^h) - \lambda(p^r)}$$

or,

$$\frac{1}{\ln\left(\dfrac{y_{rh}}{a(p^r)}\right)} = \frac{1}{b(p^r)}\left(\frac{1}{\ln u_0^h}\right) - \frac{\lambda(p^r)}{b(p^r)}.$$

Now, substituting for $(1/\ln u_0^h)$ from equation (12) in the above equation we get equation (13).

Appendix 2

Let $\hat{\pi}_r = \pi_r + \delta_r$, say, for $r = 0,1, \ldots, R$, where δ_r is the error of estimation. Then,

$$
\begin{aligned}
\frac{1}{\ln \bar{y}_{rq} - \hat{\pi}_r} &= \frac{1}{\ln \bar{y}_{rq} - \pi_r - \delta_r} \\
&= \frac{1}{\ln \bar{y}_{rq} - \pi_r} \frac{\ln \bar{y}_{rq} - \pi_r}{\ln \bar{y}_{rq} - \pi_r - \delta_r} = \frac{1}{\ln \bar{y}_{rq} - \pi_r} + \frac{\delta_r}{(\ln \bar{y}_{rq} - \pi_r)(\ln \bar{y}_{rq} - \pi_r - \delta_r)} \\
&= \frac{1}{\ln \bar{y}_{rq} - \pi_r} + \delta_r^*, \text{ say.}
\end{aligned}
$$

Therefore, equation (12) becomes

$$
\frac{1}{(\ln \bar{y}_{rq} - \hat{\pi}_r)} + \delta_r^* = \frac{1}{b(p^r)} \left(\frac{1}{(\ln \bar{y}_{0q} - \hat{\pi}_0)} + \delta_0^* + 1 \right) - \frac{\lambda(p^r)}{b(p^r)}
$$

or,

$$
\frac{1}{\ln \bar{y}_{rq} - \hat{\pi}_r} = \frac{1}{b(p^r)} \left(\frac{1}{\ln \bar{y}_{0q} - \hat{\pi}_0} + 1 \right) - \frac{\lambda(p^r)}{b(p^r)} + \left(\frac{\delta_0^*}{b(p^r)} - \delta_r^* \right),
$$

which can be written in the form of equation (13) as

$$
\frac{1}{\ln \bar{y}_{rq} - \hat{\pi}_r} \frac{1}{b(p^r)} \left(\frac{1}{\ln \bar{y}_{0q} - \hat{\pi}_0} + 1 \right) - \frac{\lambda(p^r)}{b(p^r)} + error.
$$

This again gives rise to the issue that regression error is present only because of estimation errors in the second stage, where the *error* term involves δ_0^* and δ_r^*. However, in the absence of a linear association between the error term and the regressor, we have used OLS to estimate $\frac{1}{b(p^r)}$ and $\frac{\lambda(p^r)}{b(p^r)}$. This issue needs to be further explored, possibly using a Monte Carlo study.

Appendix 3

List of Items (50th Round)

1. Cereals and cereal substitutes	7. Fruits (fresh and dry)	13. Fuel and light
2. Pulses and products	8. Sugar	14. Clothing
3. Milk and milk products	9. Salt	15. Footwear
4. Edible oils	10. Spices	16. Miscellaneous goods and services*
5. Meat, eggs, and fish	11. Beverages etc.	
6. Vegetables	12. Betel leaf, tobacco, intoxicants	17. Durable goods

*This item group has been split into "Education," "Medical," and "Miscellaneous goods and services" in the 55th round, thus making the number of items to be 19.

Monthly Per Capita Expenditure Class (Rs.) (at current prices for each round)

1. Less than 120	4. 165–190	7. 235–265	10. 355–455
2. 120–140	5. 190–210	8. 265–300	11. 455–560
3. 140–165	6. 210–235	9. 300–355	12. 560 and above

Appendix 4

TABLE A1

Estimated Values of π_r and the Coefficients of Equation (14)

States	Rural Sector 50th Round $\pi_r = \ln(a(p^r))$	Rural 50th $\frac{1}{b(p^r)}$	Rural 50th $\frac{\lambda(p^r)}{b(p^r)}$	Rural Sector 55th Round $\pi_r = \ln(a(p^r))$	Rural 55th $\frac{1}{b(p^r)}$	Rural 55th $\frac{\lambda(p^r)}{b(p^r)}$	Urban Sector 50th Round $\pi_r = \ln(a(p^r))$	Urban 50th $\frac{1}{b(p^r)}$	Urban 50th $\frac{\lambda(p^r)}{b(p^r)}$	Urban Sector 55th Round $\pi_r = \ln(a(p^r))$	Urban 55th $\frac{1}{b(p^r)}$	Urban 55th $\frac{\lambda(p^r)}{b(p^r)}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
AP	5.490	0.917	0.911	6.347	0.852	0.830	5.990	0.910	0.896	6.691	0.951	0.941
	(180.8)	(21.37)	(17.95)	(283.3)	(54.66)	(45.75)	(267.6)	(64.65)	(54.41)	(179.5)	(71.70)	(61.48)
AS	5.729	0.641	0.575	6.279	0.768	0.731	6.663	0.987	0.971	6.592	0.897	0.883
	(537.1)	(49.96)	(37.93)	(566.8)	(34.42)	(28.18)	(151.2)	(34.94)	(29.36)	(289.8)	(41.20)	(35.17)
BI	5.513	0.790	0.754	6.216	0.741	0.701	6.305	1.109	1.114	6.823	1.057	1.052
	(303.4)	(68.21)	(54.99)	(142.8)	(32.87)	(26.72)	(151.7)	(30.99)	(26.61)	(219.4)	(35.76)	(30.84)
GU	5.824	0.845	0.818	6.378	0.818	0.796	6.012	0.794	0.764	6.594	0.782	0.754
	(253.9)	(54.85)	(44.91)	(167.6)	(44.66)	(37.33)	(297.0)	(78.41)	(64.53)	(119.6)	(103.97)	(86.90)
HA	5.871	1.021	1.030	6.740	0.886	0.871	5.999	0.780	0.749	6.795	0.858	0.838
	(347.9)	(25.36)	(21.62)	(125.5)	(18.12)	(15.31)	(344.9)	(68.56)	(56.33)	(208.3)	(30.71)	(26.01)
KA	5.735	0.979	0.974	6.558	0.902	0.885	6.041	0.950	0.943	6.646	0.920	0.911
	(332.7)	(87.02)	(73.16)	(230.2)	(75.06)	(63.25)	(387.6)	(42.08)	(35.70)	(371.2)	(70.69)	(60.68)

Table A1 continued on next page

TABLE A1 (*continued*)

	Rural Sector						Urban Sector					
	50th Round			55th Round			50th Round			55th Round		
States	$\pi_r = \ln(a(p^r))$	$\frac{1}{b(p^r)}$	$\frac{\lambda(p^r)}{b(p^r)}$	$\pi_r = \ln(a(p^r))$	$\frac{1}{b(p^r)}$	$\frac{\lambda(p^r)}{b(p^r)}$	$\pi_r = \ln(a(p^r))$	$\frac{1}{b(p^r)}$	$\frac{\lambda(p^r)}{b(p^r)}$	$\pi_r = \ln(a(p^r))$	$\frac{1}{b(p^r)}$	$\frac{\lambda(p^r)}{b(p^r)}$
KE	5.745	0.943	0.942	6.404	0.885	0.877	6.071	0.905	0.894	6.582	0.897	0.886
	(284.8)	(27.80)	(23.48)	(158.5)	(18.61)	(15.85)	(123.1)	(42.44)	(35.83)	(107.8)	(77.23)	(66.18)
MP	5.435	0.923	0.915	6.305	0.887	0.869	6.029	0.909	0.895	6.420	0.880	0.864
	(293.9)	(135.58)	(113.55)	(313.0)	(134.34)	(113.10)	(271.5)	(50.03)	(42.10)	(390.5)	(78.71)	(67.01)
MH	5.811	1.123	1.141	6.251	0.884	0.873	6.193	1.060	1.073	6.493	0.963	0.964
	(220.8)	(110.59)	(94.97)	(219.8)	(149.99)	(127.19)	(326.6)	(36.13)	(31.25)	(201.6)	(21.50)	(18.66)
OR	5.291	0.799	0.769	5.863	0.796	0.771	6.589	1.113	1.115	6.840	1.019	1.007
	(417.9)	(125.22)	(102.01)	(148.6)	(30.02)	(25.00)	(420.2)	(38.52)	(33.00)	(143.9)	(80.43)	(68.94)
PU	5.871	0.840	0.821	6.403	0.775	0.750	5.922	0.719	0.683	6.263	0.728	0.697
	(293.3)	(15.23)	(12.59)	(198.9)	(14.91)	(12.41)	(234.8)	(50.22)	(40.79)	(168.5)	(54.11)	(44.90)
RA	5.879	0.931	0.919	6.258	0.711	0.673	6.070	0.840	0.815	6.579	0.782	0.751
	(470.1)	(63.22)	(52.75)	(821.7)	(52.32)	(42.53)	(591.2)	(55.26)	(45.85)	(243.7)	(69.34)	(57.77)
TN	5.674	1.031	1.040	6.412	0.989	0.990	6.064	0.976	0.972	6.500	0.958	0.958
	(271.4)	(78.08)	(66.56)	(229.3)	(117.54)	(101.07)	(220.1)	(68.58)	(58.43)	(228.5)	(18.69)	(16.21)
UP	5.715	0.990	0.988	6.417	0.882	0.864	6.135	0.942	0.929	6.590	0.972	0.964
	(493.7)	(91.35)	(77.07)	(395.0)	(82.00)	(69.05)	(265.1)	(25.75)	(21.71)	(446.2)	(59.39)	(51.09)
WB	5.581	0.804	0.775	6.476	0.900	0.878	6.231	0.988	0.986	6.806	1.005	1.002
	(350.0)	(40.25)	(32.77)	(99.03)	(70.36)	(58.97)	(269.5)	(47.90)	(40.85)	(299.3)	(47.78)	(41.31)
All-India	5.743	1.000	1.000	6.574	1.000	1.000	6.147	1.000	1.000	6.644	1.000	1.000
	(551.8)	–	–	(266.2)	–	–	(541.0)	–	–	(406.0)	–	–

Notes: Figures in parentheses are the asymptotic t-ratios.

A possible reason why the t-ratios for the estimated π_r's are so large is that asymptotically equation (10) holds exactly. A similar argument holds for t-ratios corresponding to $\frac{1}{b(p^r)}$ and $\frac{\lambda(p^r)}{b(p^r)}$ in equation (14).

TABLE A2

Estimated Values of the Basic Index $\left(=\frac{a(p^r)}{a(p^0)}\right)$, $b(p^r)$, and $\lambda(p^r)$

States	Rural Sector						Urban Sector					
	50th Round			55th Round			50th Round			55th Round		
	$\frac{a(p^r)}{a(p^0)}$	$b(p^r)$	$\lambda(p^r)$	$\frac{a(p^r)}{a(p^0)}$	$b(p^r)$	$\lambda(p^r)$	$\frac{a(p^r)}{a(p^0)}$	$b(p^r)$	$\lambda(p^r)$	$\frac{a(p^r)}{a(p^0)}$	$b(p^r)$	$\lambda(p^r)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
AP	0.776	1.091	0.994	0.797	1.173	0.974	0.855	1.099	0.984	1.049	1.051	0.989
AS	0.986	1.561	0.898	0.745	1.303	0.953	1.676	1.013	0.983	0.949	1.115	0.985
BI	0.794	1.265	0.953	0.699	1.349	0.946	1.172	0.902	1.005	1.196	0.946	0.995
GU	1.084	1.184	0.968	0.822	1.222	0.973	0.874	1.260	0.963	0.951	1.278	0.964
HA	1.137	0.979	1.008	1.180	1.129	0.983	0.862	1.283	0.961	1.164	1.165	0.977
KA	0.992	1.021	0.994	0.984	1.108	0.981	0.900	1.052	0.992	1.002	1.087	0.990
KE	1.002	1.060	0.999	0.844	1.130	0.991	0.927	1.105	0.988	0.940	1.115	0.988
MP	0.735	1.083	0.991	0.764	1.127	0.980	0.889	1.100	0.984	0.800	1.136	0.982
MH	1.070	0.890	1.016	0.724	1.131	0.987	1.047	0.943	1.012	0.860	1.038	1.001
OR	0.636	1.252	0.964	0.491	1.256	0.969	1.556	0.898	1.002	1.217	0.982	0.989
PU	1.136	1.191	0.977	0.843	1.291	0.969	0.799	1.391	0.950	0.683	1.374	0.957
RA	1.145	1.074	0.987	0.729	1.406	0.946	0.926	1.191	0.971	0.937	1.279	0.961
TN	0.933	0.970	1.008	0.850	1.011	1.001	0.920	1.025	0.997	0.866	1.044	1.000
UP	0.972	1.010	0.998	0.855	1.134	0.980	0.988	1.061	0.986	0.948	1.029	0.992
WB	0.850	1.243	0.963	0.907	1.111	0.975	1.088	1.012	0.998	1.176	0.995	0.997
All-India	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

References

Almas, Ingvild, "International Income Inequality: Measuring PPP Bias by Estimating Engel Curves for Food," LIS Working Paper No. 473, 2008.

Aten, B. and T. Menezes, "Poverty Price Levels: An Application to Brazilian Metropolitan Areas," World Bank ICP Conference, Washington D.C., March 11–15, 2002.

Balk, B. M., "A Comparison of Ten Methods for Multilateral International Price and Volume Comparison," *Journal of Official Statistics*, 12, 199–222, 1996.

Banks, J., R. Blundell, and A. Lewbel, "Quadratic Engel Curves and Consumer Demand," *Review of Economics and Statistics*, 79, 527–39, 1997.

Barrett, Garry F. and M. Brzozowski, "Using Engel Curves to Estimate the Bias in the Australian CPI," Working Paper, October, 2008.

Beatty, Timothy and Erling R. Larsen, "Using Engel Curves to Estimate Bias in the Canadian CPI as a Cost of Living Index," *Canadian Journal of Economics*, 38, 482–99, 2005.

Bhattacharya, N., G. S. Chatterjee, and P. Pal, "Variations in Level of Living Across Regions and Social Groups in India, 1963–64 and 1973–74," in T. N. Srinivasan and P. K. Bardhan (eds), *Rural Poverty in South Asia*, Oxford University Press, Oxford, 1988.

Bhattacharya, S. S., P. D. Joshi, and A. B. Roychowdhury, "Regional Price Indices Based on NSS 25th Round Consumer Expenditure Data," *Sarvekshana*, 3(4), 107–21, 1980.

Coondoo, D. and S. Saha, "Between-State Differentials in Rural Consumer Prices in India: An Analysis of Intertemporal Variations," *Sankhya*, Series B, 52(3), 347–60, 1990.

Coondoo, D., A. Majumder, and R. Ray, "A Method of Calculating Regional Consumer Price Differentials with Illustrative Evidence from India," *Review of Income and Wealth*, 50, 51–68, 2004.

Costa, Dora L., "Estimating Real Income in the United States from 1888 to 1994: Correcting CPI Bias Using Engel Curves," *Journal of Political Economy*, 109, 1288–310, 2001.

Deaton, A. S., *The Analysis of Household Surveys—A Microeconometric Approach to Development Policy,* Johns Hopkins University Press, Baltimore and London, 1997.

———, "Prices and Poverty in India, 1987–2000," *Economic and Political Weekly*, January 25, 362–8, 2003.

Deaton, A. S. and J. Muellbauer, *Economics and Consumer Behaviour*, Cambridge University Press, Cambridge, 1980a.

———, "An Almost Ideal Demand System," *American Economic Review*, 70, 312–26, 1980b.

Deaton, A. S. and A. Tarozzi, "Prices and Poverty in India," in A. S. Deaton and V. Kozel (eds), *Data and Dogma: The Great Indian Poverty Debate*, Macmillan, New Delhi, 2005.

Diewert, W. E., "Axiomatic and Economic Approaches to International Comparisons," in A. Heston and R. E. Lipsey (eds), *International and Interarea Comparisons of Income, Output and Prices*, University of Chicago Press, Chicago, 13–87,1999.

Fry, V. and P. Pashardes, "Constructing the True Cost of Living Index from the Engel Curves of the PIGLOG Model," *Journal of Applied Econometrics*, 4, 41–56, 1989.

Geary, R. C., "A Note on Comparison of Exchange Rate and Purchasing Power Parities Between Countries," *Journal of the Royal Statistical Society*, 121 (Part 1), 97–9, 1958.

Hamilton, B., "Using Engel's Law to Estimate CPI Bias," *American Economic Review*, 91, 619–30, 2001.

Hill, R. J., "A Taxonomy of Multilateral Methods for Making International Comparisons of Prices and Quantities," *Review of Income and Wealth*, 43, 49–69,1997.

Khamis, S. H., "Properties and Conditions for the Existence of a New Type of Index Numbers," *Sankhya*, Series B, 32 (Parts 1 & 2), 81–98, 1972.

Kokoski, M. F., B. R. Moulton, and K. D. Zeischang, "Interarea Price Comparisons for Heterogeneous Goods and Several Levels of Commodity Aggregation," in R. E. Lipsey and A. Heston (eds), *International and Interarea Comparisons of Prices, Income and Output*, National Bureau of Economic Research, Chicago University Press, Chicago, 327–64, 1999.

Kravis, I. B., A. Heston, and R. Summers, *International Comparison of Real Product and Purchasing Power*, John Hopkins University Press, Baltimore, 1978.

Lancaster, J. and R. Ray, "Comparison of Alternative Models of Household Equivalence Scales: The Australian Evidence on Unit Record Data," *Economic Record*, 74, 1–14, 1998.

Larsen, Erling R., "Does the CPI Mirror Costs-of-Living? Engel's Law Suggests Not in Norway," *Scandinavian Journal of Economics*, 109, 177–95, 2007.

Logan, Trevon D., "Are Engel Curve Estimates of CPI Bias Biased?" NBER Working Paper Series, Working Paper 13870, March, 2008.

Neary, J. P., "Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income," *American Economic Review*, 94, 1411–28, 2004.

Papalia, Rosa Bernardini, "Estimating Real Income at Regional Level Using a CPI Bias Correction," Unpublished Manuscript, University of Bologna, 2006.

Prasada Rao, D. S., "Aggregation Methods for International Comparison of Purchasing Power Parities and Real Income: Analytical Issues and Some Recent Developments," *Proceedings of the International Statistical Institute*, 51st Session, 197–200, 1997.

———, "Weighted EKS and Generalised CPD Methods for Aggregation at Basic Heading Level and Above Basic Heading Level," Joint World Bank–OECD Seminar on Purchasing Power Parities—Recent Advances in Methods and Applications, Washington D.C., 2001.

Sen, A., "Real National Income," *Review of Economic Studies*, 43, 19–39. 1976.

Summers, R., "International Price Comparisons Based Upon Incomplete Data," *Review of Income and Wealth*, 19, 1–16, 1973.