Volume Title: International and Interarea Comparisons of Income, Output, and Prices

Volume Author/Editor: Alan Heston and Robert E. Lipsey, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-33110-5

Volume URL: http://www.nber.org/books/hest99-1

Publication Date: January 1999

Chapter Title: Interarea Price Comparisons for Heterogeneous Goods and Several Levels of Commodity Aggregation

Chapter Author: Mary F. Kokoski, Brent R. Moulton, Kimberly D. Zieschang

Chapter URL: http://www.nber.org/chapters/c8387

Chapter pages in book: (p. 123 - 169)

3 Interarea Price Comparisons for Heterogeneous Goods and Several Levels of Commodity Aggregation

Mary F. Kokoski, Brent R. Moulton,
and Kimberly D. Zieschang

There has been steady interest for many years among academic, business, and policy users of economic data for "purchasing power parities" permitting conversions of the gross domestic product of two or more countries, which information is compiled in national currency, into common units so that the real size of one economy can be measured relative to another. In the available literature on international comparisons, the microeconomic foundations for such conversion measures have been most definitively considered by Diewert (1986). In this paper, we focus on a part of the consumption component of GDP and consider comparing the price of consumption among areas within a given country rather than between countries. Judging from periodic contacts from data users with Bureau of Labor Statistics (BLS) staff, such place-to-price comparisons for areas within the United States are very much in demand. In our approach to this problem, we contribute to the resolution of several technical issues in compiling both interarea and international price indexes, including imposing transitivity (the so-called aggregation problem of international comparisons) and the problem of adjusting for interarea differences in the composition and quality of consumption, among others. Although not specifically designed for producing interarea comparisons, the database from

Mary F. Kokoski is an economist in the division of price and index number research at the Bureau of Labor Statistics, U.S. Department of Labor. Brent R. Moulton is associate director for national income, expenditure, and wealth accounts at the Bureau of Economic Analysis, U.S. Department of Commerce. This paper was completed while he was chief of the division of price and index number research at the Bureau of Labor Statistics, U.S. Department of Labor. Kimberly D. Zieschang is a senior economist with the International Monetary Fund. This paper was completed while he was associate commissioner for compensation and working conditions at the Bureau of Labor Statistics, U.S. Department of Labor.

The opinions expressed in this paper are those of the authors and do not represent the policies of the Bureau of Labor Statistics (BLS) or the views of other BLS staff members.

which the U.S. consumer price index (CPI) is calculated is nevertheless a rich source of geographic price information, one that we exploit in this study.

We adopt a Törnqvist measurement framework and derive a very general form for a transitive, multilateral system of parities within that framework. The Caves, Christensen, and Diewert (CCD) (1982b) implementation of the Eltetö, Köves, and Szulc (EKS) methodology (described in English in Dreschler [1973]) for multilateral price measurement uses a special case of this general multilateral form, one involving the estimation of $2N - 1$ parameters when there are N items in the index aggregate. The unknown parameters represent a reference set of value shares and prices against which the shares and prices of all areas are compared. We show that, if Törnqvist aggregates are to be formed from lower-level aggregates on a single, that is, commodity, aggregation tree, the adjustment can be applied successively from the lowest to the highest levels of aggregation to produce a set of reference prices that, while fixed across areas, have components corresponding to each level of item aggregation.

We adapt earlier index number results from Zieschang (1985, 1988) and Fixler and Zieschang (1992) to incorporate information on the characteristics of the products that are systematically randomly sampled for the CPI using country-product-dummy regression at "entry-level item" product detail. In this index number framework, coefficients from these regressions are used in constructing quality-adjustment factors for the place-to-place price comparisons.

We show how the parameters of the transitive Törnqvist system can be estimated with a particular regression model to impose transitivity with minimal adjustment of the data. We believe that the model represents, if not a completely new approach, a substantive refinement of the regression-based multilateral adjustment that underlies the versions of EKS expounded by, for example, Cuthbert and Cuthbert (1988) and Selvanathan and Rao (1992).

Kokoski, Cardiff, and Moulton (1994) estimate country-product-dummy regressions on micro data from the U.S. consumer price index for the thirteen-month period June 1988–July 1989. The models are fitted at the index item level to produce quality-adjusted price indexes for the lowest, item level of aggregation. The regression methodology developed in this paper for imposing transitivity is demonstrated on data for a small, three-aggregation-level example problem for fruits and vegetables from an extract of 1993 data based on the Kokoski, Cardiff, and Moulton (1994) study, presaging the computation of price indexes for successively higher commodity aggregates for the forty-four major urban centers and region-city size groups covered by the CPI.

We close by summarizing the methodological and empirical results, by describing an application of the methodology enforcing consistency between the comparisons across areas within a given time period and comparisons across time within a given area. Finally, we point out a notable advantage of this framework for compiling international parities over the narrow specification approach now used in the International Comparisons Project (ICP). Provided that a standard list of item characteristics and item groups is promulgated, item strata can be broadened, increasing the likelihood of finding a useful specification from country to country. Further, this advantage is not bought at the cost of operational feasibility since calculation can be decentralized, obviating the need for central, transnational access to closely guarded national micro data.

3.1 Economic Index Number Concepts Incorporating Information on the Characteristics of Heterogeneous Goods

Let p_i^a be the price in area a, of which there are A areas in total, of commodity i. Let q_i^a be the corresponding quantity purchased, and let x_i^a be the vector of characteristics of the ith item specification transacted in area a. Let e_h^a represent the total expenditures of consumer unit h in area a. We will use interchangeably the terms *economic household* and *consumer unit* for the economic unit of analysis, following BLS terminology. A consumer unit is a group of individuals whose consumption decisions for significant components of expenditure are joint or shared. Let q_h^a denote the vector of goods consumed by household h in area a with vector of characteristics x_h^a and prices p_h^a.

We suppose that each consumer unit in area a minimizes the cost of achieving a given level of welfare at expenditure level e_h^a so that the consumer unit cost of consumption of a given quality of goods as determined by the vector x_h^a would be

$$e_h^a = E_h^a(u_h^a, x_h^a, p_h^a) = \min_{q_h^a} \{p_h^{a\prime} q_h^a : F_h^a(x_h^a, q_h^a) \geq u_h^a\},$$

where $F_h^a(x_h^a, q_h^a)$ is the utility function of consumer unit h, and u_h^a is the unit's welfare index.

We suppose further that the consumer unit in area a faces a hedonic locus of market equilibrium prices across the quality spectrum given by $p_h^a = H^a(x_h^a)$ and that the unit minimizes the cost of achieving welfare level u_h^a over the characteristics of goods, with the result that

$$(1) \qquad \nabla_{x_h^a} E_h^a(u_h^a, x_h^a, p_h^a) + \nabla_{x_h^a} p_h^{a\prime} \nabla_{p_h^a} E_h^a(u_h^a, x_h^a, p_h^a) = 0.$$

Since $\nabla_{x_h^a} p_h^a = \nabla_{x_h^a} H^a$ and $\nabla_{p_h^a} E_h^a(u_h^a, x_h^a, p_h^a) = q_h^a$, the latter by the Shephard/Hotelling lemma, we have

$$(2) \qquad \nabla_{x_h^a} E_h^a(u_h^a, x_h^a, p_h^a) = -\nabla_{x_h^a} H^{a\prime} q_h^a.$$

If H^a is semilog, as generally assumed in hedonic studies, so that

$$(3) \qquad \ln H_i^a = \alpha_i^a + \beta_i^{a\prime} x_i^a,$$

then the characteristics gradient expression can be rewritten

$$(4) \qquad \nabla_{x_h^a} E_h^a(u_h^a,\ x_h^a,\ p_h^a) = -\beta^{a\prime} w_h^a e_h^a,$$

where

$$w_{i,h}^a = \frac{p_{i,h}^a q_{i,h}^a}{\sum_i p_{i,h}^a q_{i,h}^a},$$

$$w_h^a = \begin{bmatrix} w_{1,h}^a \\ \vdots \\ w_{N_q,h}^a \end{bmatrix}, \quad N_q = \text{number of commodities},$$

$$\beta^{a\prime} = \begin{bmatrix} \beta_1^{a\prime} \\ \vdots \\ \beta_{N_x}^{a\prime} \end{bmatrix}, \quad N_x = \text{number of product characteristics}.$$

Turning now to aggregate expenditure over consumer units in an area, Diewert (1987) has considered this problem as a weighted average of individual household index numbers comparing the prices in two areas in the "democratic-weighting" case. In this paper, we follow his characterization of the "plutocratic" expenditure-weighted case with some modifications for the heterogeneity of goods within and between areas. The area aggregate expenditure function is

$$E^a(\vec{u}^a,\ \vec{x}^a,\ \vec{p}^a) = \sum_h E_h^a(u_h^a,\ x_h^a,\ p_h^a),$$

where the arrow over an argument indicates the concatenation of vectors across households. We then consider the expenditure function in terms of log transformed price arguments as

$$Q^a(\vec{u}^a,\ \vec{x}^a,\ \ln \vec{p}^a) = E^a(\vec{u}^a,\ \vec{x}^a,\ \vec{p}^a) = \sum_h E_h^a(u_h^a,\ x_h^a,\ p_h^a)$$
$$= \sum_h Q_h^a(u_h^a,\ x_h^a,\ \ln p_h^a).$$

We "plutocratically" aggregate across households in area a such that the expenditure-weighted averages for characteristics and log prices represent the indicators determining area demand behavior, where area item demand is the sum of the economic household item demands for the area.[1] We do not require strong aggregation conditions but effectively hold the distribution of product characteristics and prices fixed across economic households within area a as in

$$(5) \qquad \tilde{Q}^a(\vec{u}^a,\ \bar{x}^a,\ \overline{\ln p}^a) = Q^a(\vec{u}^a, \iota \otimes \bar{x}^a + v_x^a,\ \iota \otimes \overline{\ln p}^a + v_{\ln p}^a),$$

1. Actually, the results to follow do not depend on the particular form of area aggregation for product characteristics and prices. Although this discussion is couched in terms of an arithmetic area mean for characteristics and a geometric mean for prices, others will work as well.

where $v_x^a = \vec{x}^a - \iota \otimes \bar{x}^a$, $v_{\ln p}^a = \ln\vec{p}^a - \iota \otimes \overline{\ln p}^a$, ι = a vector of ones of dimension equal to the number of households, and $\otimes$ = Kronecker product, all giving the deviations of the area means from the individual household values for commodity characteristics and prices paid.

Using the derivatives of the expenditure function with respect to log prices expressed in terms of observable expenditure shares, Diewert (1976) and Caves, Christensen, and Diewert (1982a) have shown that the Törnqvist index number is exact for the translog flexible functional form. The translog aggregator function differentially approximates any price aggregator function (i.e., cost of utility, input cost, revenue function) to the second order at a point, and it is exact for the Törnqvist index number even when some of the parameters (those on the first-order terms) of the underlying aggregator function are different in the two periods or localities compared. We take the derivative of the area expenditure function with respect to interarea average household characteristics and price arguments to obtain

$$(6)\quad \begin{aligned} \frac{\partial}{\partial \bar{x}_{iz}^a} \ln \tilde{E}^a(\vec{u}^a, \bar{x}^a, \exp(\overline{\ln p^a})) &= \frac{\partial}{\partial \bar{x}_{iz}^a} \ln \tilde{Q}^a(\vec{u}^a, \bar{x}^a, \overline{\ln p^a}) \\ &= \sum_h \frac{\partial}{\partial x_{izh}^a} Q_h^a(u_h^a, x_h^a, \ln p_h^a)/\tilde{Q}^a = \beta_{iz}^a \sum_h w_{ih}^a s_h^a = -\beta_{iz}^a \bar{w}_i^a, \end{aligned}$$

$$(7)\quad \begin{aligned} \frac{\partial}{\partial \overline{\ln p_i^a}} \ln \tilde{E}(\vec{u}^a, \bar{x}^a, \exp(\overline{\ln p^a})) &= \frac{\partial}{\partial \overline{\ln p_i^a}} \ln \tilde{Q}(\vec{u}^a, \bar{x}^a, \overline{\ln p^a}) \\ &= \sum_h \frac{\partial}{\partial \ln p_{ih}^a} Q_h^a(u_h^a, x_h^a, \ln p_h^a)/\tilde{Q}^a = \sum_h w_{ih}^a s_h^q = \bar{w}_i^a, \end{aligned}$$

where

$$w_{ih}^a = \frac{p_{ih}^a q_{ih}^a}{\sum_i p_{ih}^a q_{ih}^a} = \frac{p_{ih}^a q_{ih}^a}{e_h^a},$$

$$s_h^a = \frac{e_h^a}{\sum_h e_h^a},$$

are, respectively, the within-household expenditure shares of commodities and the between-household total expenditure shares of consumer unit h in area a.

Finally, we assume that the area aggregate expenditure function $\ln\tilde{Q}^a(e^a, \bar{x}^a, \overline{\ln p^a})$ has a quadratic, "semi-translog" functional form in its arguments with coefficients of second-order terms independent of location but with possibly location-specific coefficients on linear terms. Following Caves, Christensen, and Diewert (1982a), then, we can derive the following (logarithmic) index number result:

$$\ln I^{ab} = \frac{1}{2}\Big[\ln \tilde{Q}^a(\vec{u}^a, \bar{x}^b, \overline{\ln p^b}) - \ln \tilde{Q}^a(\vec{u}^a, \bar{x}^a, \overline{\ln p^a})$$

$$+ \ln \tilde{Q}^b(\vec{u}^b, \bar{x}^b, \overline{\ln p^b}) - \ln \tilde{Q}^b(\vec{u}^b, \bar{x}^a, \overline{\ln p^a})\Big]$$

(8)

$$= \frac{1}{2}\Big[\nabla_{\ln p} \ln \tilde{Q}^a(\vec{u}^a, \bar{x}^a, \overline{\ln p^a}) + \nabla_{\ln p} \ln \tilde{Q}^b(\vec{u}^b, \bar{x}^b, \overline{\ln p^b})\Big](\overline{\ln p^b} - \overline{\ln p^a})$$

$$+ \frac{1}{2}\Big[\nabla_x \ln \tilde{Q}^a(\vec{u}^a, \bar{x}^a, \overline{\ln p^a}) + \nabla_x \ln \tilde{Q}^b(\vec{u}^b, \bar{x}^b, \overline{\ln p^b})\Big](\bar{x}^b - \bar{x}^a).$$

Substituting (6) and (7) into (8), following Caves, Christensen, and Diewert (1982a) again, and with reference to Fixler and Zieschang (1992), we have

$$\ln I^{ab} = \ln T^{ab}$$

(9)

$$\equiv \frac{1}{2}\sum_i \Big[(\bar{w}_i^a + \bar{w}_i^b)(\overline{\ln p_i^b} - \overline{\ln p_i^a}) - \sum_z (\beta_{iz}^a \bar{w}_i^a + \beta_{iz}^b \bar{w}_i^b)(\bar{x}_{iz}^b - \bar{x}_{iz}^a)\Big].$$

This formula for the bilateral index between areas is an extremely flexible result that permits all parameters of the semilog "hedonic" price equations to differ by area, fully reflecting household optimization over measured product quantities and characteristics.

3.2 Törnqvist Multilateral (Transitive) Systems of Bilateral Index Numbers

In another paper, Caves, Christensen, and Diewert (1982b) noted that the system of bilateral Törnqvist interarea indexes is not transitive but developed a simply calculated multilateral variant satisfying the transitivity property. Returning to lowercase notation for the index arguments for areas, we derive the following general implication of transitivity for this class of index number:

PROPOSITION 1: For the bilateral Törnqvist item index to be transitive, it is necessary and sufficient that, for all a, b, there exist constant vectors w^0 and $\ln p^0$ such that

(10)

$$\sum_i w_i^a \ln p_i^b - \sum_i w^b \ln p_i^a = \sum_i w_i^0(\ln p_i^b - \ln p_i^a)$$

$$- \sum_i \ln p_i^0(w_i^b - w_i^a),$$

where w_i^0 = a reference share for index item i for the entire region, with $\sum_i w_i^0 = 1$, and p_i^0 = a reference price for index item i across the entire region. Furthermore, if this condition holds, the multilateral Törnqvist index has the form

$$\ln T^{ab} \equiv \sum_i \left[\frac{1}{2}(w_i^0 + w_i^b)(\ln p_i^b - \ln p_i^0) - \frac{1}{2}(w_i^0 + w_i^a)(\ln p_i^a - \ln p_i^0)\right]. \tag{11}$$

The proof is given in appendix A. Caves, Christensen, and Diewert (1982b) showed that application of the EKS principle to a system of bilateral Törnqvist indexes yields the formula given above with the reference shares and log prices set at their simple arithmetic averages across areas. Clearly, these simple averages could also be replaced with total expenditure-weighted averages. We consider still another way of estimating the reference shares and prices in the next section.

The overall system can be adjusted to be transitive in both prices and item characteristics by applying the principle underlying proposition 1. This is stated in proposition 2:

PROPOSITION 2: If the area-specific country-product-dummy (CPD) coefficients are known, for the bilateral quality-adjusted Törnqvist item index to be transitive it is necessary and sufficient that, for all a, b,

$$\begin{aligned} &\sum_n w_i^a \left(\ln p_i^b - \left(\sum_z \beta_{iz}^a x_{iz}^b\right)\right) - \left[\sum_i w_i^a \left(\ln p_i^a - \left(\sum_z \beta_{iz}^a x_{iz}^a\right)\right)\right] \\ &= \sum_i \sum_{z^b} [-\beta_{iz}^0 w_i^0](x_{iz}^b - x_{iz}^a) + \sum_i \sum_z x_{iz}^0 (\beta_{iz}^b w_i^b - \beta_{iz}^a w_i^a) \\ &\quad + \sum_i w_i^0 (\ln p_i^b - \ln p_i^a) + \sum_i [-\ln p_i^0](w_i^b - w_i^a), \end{aligned} \tag{12}$$

where x_{iz}^0 = a reference characteristic z for index item i across the entire region, β_{iz}^0 = a reference coefficient for the characteristic z of item i in a semilog hedonic equation explaining specification price across the entire region, p_i^0 = a reference price for item i across the entire region, and w_i^0 = a reference share for item i for the entire region. Furthermore, if this condition holds, the bilateral Törnqvist index for item group i has the form

$$\begin{aligned} \ln T^{ab} = &-\sum_i \sum_z \frac{1}{2}(\beta_{iz}^0 w_i^0 + \beta_{iz}^b w_i^b)(x_{iz}^b - x_{iz}^0) + \sum_i \frac{1}{2}(w_i^0 + w_i^b)(\ln p_i^b - \ln p_i^0) \\ &- \left[-\sum_i \sum_z \frac{1}{2}(\beta_{iz}^0 w_i^0 + \beta_{iz}^a w_i^a)(x_{iz}^a - x_{iz}^0) + \sum_i \frac{1}{2}(w_i^0 + w_i^a)(\ln p_i^a - \ln p_i^0)\right]. \end{aligned} \tag{13}$$

3.3 Multilateral Price Measurement with Subaggregates of Items

Let p_{ijklmn}^a be the price in area a, of which there are A areas in total, of specification n in item group m in stratum class l in basic heading class k in group j

in major group or division i. Let q^a_{ijklmn} be the corresponding quantity purchased. The bilateral Törnqvist index comparing the prices in areas a and b for item aggregate *ijklm* is

$$\ln T^{ab}_{ijklm} \equiv \sum_n \frac{1}{2}(w^a_{ijklmn} + w^a_{ijklmn})(\ln p^b_{ijklmn} - \ln p^a_{ijklmn}),$$

where w^a_{ijklmn} = the value share in area a of specification *ijklmn* within the next-higher group *ijklm*, with $w^a_{ijklmn} \equiv (p^a_{ijklmn}q^a_{ijklmn})/(\sum_n p^a_{ijklmn}q^a_{ijklmn})$ and with q^a_{ijklmn} the quantity of the specification transacted, so that $\sum_n w^a_{ijklmn} = 1$.

3.3.1 Analysis of the Contribution of Subaggregates to Levels of Place-to-Place Indexes

In practice, index numbers are produced for hierarchical classification trees of products, industries, occupations, etc. Because Törnqvist indexes are linear in the log differences of detailed specification prices, the contribution of each subaggregate, say, women's apparel, to the all-items-level ratio between two areas can be readily calculated by exponentiating the appropriate weighted sums of log price differences. These sums would be calculated from the transitive expression for the index given in equation (11), where it is expressed in terms of locality weights averaged with reference weights and price differentials from reference prices. In this case, *all-items* bilateral indexes are constructed as the direct aggregation of the specification prices as

$$\ln T^{ab} \equiv \sum_i \sum_j \sum_k \sum_l \sum_m \sum_n \left[\frac{1}{2}(w^0_{ijklmn} + w^b_{ijklmn})(\ln p^b_{ijklmn} - \ln p^0_{ijklmn}) - \frac{1}{2}(w^0_{ijklmn} + w^a_{ijklmn})(\ln p^a_{ijklmn} - \ln p^0_{ijklmn})\right].$$

The contribution to the level of ln T^{ab} of major commodity group i would simply be the subordinate sum

$$\ln C^{ab}_i \equiv \sum_j \sum_k \sum_l \sum_m \sum_n \left[\frac{1}{2}(w^a_{ijklmn} + w^0_{ijklmn})(\ln p^b_{ijklmn} - \ln p^0_{ijklmn}) - \frac{1}{2}(w^a_{ijklmn} + w^0_{ijklmn})(\ln p^a_{ijklmn} - \ln p^0_{ijklmn})\right].$$

The simplicity of this approach to analysis of the place-to-place price differentials of subaggregates, and its focus on subaggregate change within the larger all-items context, has a great deal of appeal. The extension of this discussion to quality-adjusted price indexes, including characteristics, is straightforward and left to the reader.

3.3.2 Transitivity Simultaneously across Several Aggregation Levels

Nevertheless, when place-to-place subaggregate indexes are to be published in addition to the all-items index, it may not be seen as sufficient to adjust only the all-items index to be transitive. The subaggregates would then be required not only to satisfy transitivity but also to aggregate according to an index number rule to successively higher levels. This is a property distinct from consistency in aggregation, whereby an index formula for an aggregate calculated directly is the same as that calculated with the same formula successively applied to intermediate subaggregates. Rather, assuming that the same index formula is repeatedly applied at each level, as in the latter case, we would like all levels of aggregation to satisfy transitivity while preserving the aggregation rule so that users might also combine low-level aggregates following the same formula and weighting and be assured of obtaining the higher-level aggregates. We show that it is possible to construct such aggregation-consistent place-to-place indexes under a multilevel Törnqvist aggregation rule.

Having dealt with the first level of aggregation in section 3.1 above, we now consider aggregation of the item indexes *ijklm* to the stratum level *ijkl*. We first observe from proposition 2 that the transitivity of the item aggregate *ijklm* permits us to identify average price levels for the aggregate for each area in the region as

$$\ln \bar{p}^{a}_{ijklmn} = \sum_{n} \frac{1}{2}(w^{0}_{ijklmn} + w^{a}_{ijklmn})(\ln p^{a}_{ijklmn} - \ln p^{0}_{ijklmn}),$$

allowing us to rewrite the expression for the bilateral item index as

$$\ln T^{ab}_{ijklm} = \ln \bar{p}^{b}_{ijklm} - \ln \bar{p}^{a}_{ijklm}.$$

The bilateral index between areas a and b of the stratum aggregate *ijkl* over item groups *ijklm* is

$$\begin{aligned}\ln T^{ab}_{ijkl} &= \sum_{m} \frac{1}{2}(w^{a}_{ijklm} + w^{b}_{ijklm}) \ln T^{ab}_{ijklm} \\ &= \sum_{m} \frac{1}{2}(w^{a}_{ijklm} + w^{ab}_{ijklm})(\ln \bar{p}^{b}_{ijklm} - \ln \bar{p}^{a}_{ijklm}).\end{aligned}$$

Applying the proposition to the stratum level, the transitive bilateral index between areas a and b of the stratum aggregate *ijkl* over item groups *ijklm* is, therefore,

$$\begin{aligned}\ln T^{ab}_{ijkl} \equiv \sum_{m} \Bigg[&\frac{1}{2}(w^{0}_{ijklm} + w^{b}_{ijklm})(\ln \bar{p}^{b}_{ijklm} - \ln p^{0}_{ijklm}) \\ &- \frac{1}{2}(w^{0}_{ijklm} + w^{a}_{ijklm})(\ln \bar{p}^{a}_{ijklm} - \ln p^{0}_{ijklm}) \Bigg].\end{aligned}$$

We note that the expression for the transitive Törnqvist *item* index given above would have been obtained if the reference *specification* prices had been $\ln p^{0(m)}_{ijklmn} = \ln p^{0}_{ijklmn} + \ln p^{0}_{ijklm}$. Further, if the specification reference prices were so adjusted, the transitivity of the lower-level item indexes would continue to hold. Further still, because each level's log index is the difference between weighted log price relatives, those components constant within group cancel, leaving only those elements varying with members of the group. To confirm,

$$\begin{aligned}
\ln T^{ab}_{ijklm}
&\equiv \sum_n \Big[\frac{1}{2}(w^0_{ijklmn} + w^b_{ijklmn})(\ln \bar{p}^b_{ijklmn} - \ln p^0_{ijklmn} - \ln p^0_{ijklm}) \\
&\quad - \frac{1}{2}(w^0_{ijklmn} + w^a_{ijklmn})(\ln \bar{p}^a_{ijklmn} - \ln p^0_{ijklmn} - \ln p^0_{ijklm})\Big] \\
&= \sum_n \Big[\frac{1}{2}(w^0_{ijklmn} + w^b_{ijklmn})(\ln \bar{p}^b_{ijklmn} - \ln p^0_{ijklmn}) \\
&\quad - \frac{1}{2}(w^0_{ijklmn} + w^a_{ijklmn})(\ln \bar{p}^a_{ijklmn} - \ln p^0_{ijklmn})\Big] \\
&\quad - \sum_n \Big[\frac{1}{2}(w^0_{ijklmn} + w^b_{ijklmn})\ln p^0_{ijklm} - \frac{1}{2}(w^0_{ijklmn} + w^a_{ijklmn})\ln p^0_{ijklm}\Big] \\
&= \sum_n \Big[\frac{1}{2}(w^0_{ijklmn} + w^b_{ijklmn})(\ln \bar{p}^b_{ijklmn} - \ln p^0_{ijklmn}) \\
&\quad - \frac{1}{2}(w^0_{ijklmn} + w^a_{ijklmn})(\ln \bar{p}^a_{ijklmn} - \ln p^0_{ijklmn})\Big].
\end{aligned}$$

In effect, then, each level of aggregation adds a component to the reference price vector, and a system transitive at all levels of aggregation would therefore require specification reference prices of the form

$$\ln p^{0(ijklm)}_{ijklmn} = \ln p^0_{ijklmn} + \ln p^0_{ijklm} + \ln p^0_{ijkl} + \ln p^0_{ijk} + \ln p^0_{ij} + \ln p^0_{i}.$$

Finally, only the components of the reference price vector relevant to (within) a given aggregation level enter into that level's transitivity-adjustment equation. This permits a decomposition of the estimation procedure allowing the lowest aggregate reference shares and reference price components to be estimated first (using the regression equation presented in the proposition), followed by successively higher levels in turn. Again, the extension to quality-adjusted price indexes accounting for the differences in product characteristics across areas is straightforward.

3.3.3 More Than One Aggregation Tree

Statistical price series are often published on more than one aggregation scheme. For example, establishment data are often published on commodities and industries (producer price indexes) and occupations and industries (employment cost indexes). Multiple trees can be incorporated into the structure just elucidated by merging the trees and defining cells by crossing the classification strata in the two (or more) structures. There is a new consistency issue introduced, namely, that comparable aggregates formed from differing subaggregates in the distinct classification structures should be the same. Most obviously, the all-items price index on establishment data should be the same whether the subaggregates are industries or occupations/commodities. Similarly, the Industry Division 1 labor compensation index should be the same number, whether calculated as an aggregate of the two-digit industries or of major occupation groups within Division 1. Constraints of this type bring us much closer to imposing a de facto requirement of traditional consistency in aggregation on the data but are not equivalent to imposing the property unless each elementary price is contained in a distinct cross-cell of the two or more structures. In this paper, we consider only a one-commodity aggregation tree.

3.4 Estimation of the Reference Values for Shares, Prices, and Determinants of Quality

3.4.1 Adjusting for Quality from Place to Place

Data permitting, it is standard practice in constructing place-to-place price indexes to adjust for known price-determining specification characteristics using a regression of specification prices on measured characteristics and a set of dummy variables for locality. This country-product-dummy (CPD) approach is relatively simple and easily implemented. The most obvious way of controlling for quality in constructing a place-to-place index is to use the intercept plus coefficients on the area dummy variables as quality-adjusted price levels in the bilateral price index for the item group. In this section, we show that the use of the CPD model in this way is a special case of an exact Törnqvist index number that incorporates quality characteristics when there is a known hedonic function. The special case is that the hedonic function is the same from area to area, other than the intercept.

Suppose that the characteristics of specification n in area a are given by the vector x^{a}_{ijklmn} and that we define the set of dummy variables

$$L^{a} = \left[\text{For } b = 2, 3, \cdots, A, \quad \Delta^{ab} = \begin{cases} 1 \text{ if } b = a \\ 0 \text{ otherwise} \end{cases}\right].$$

A CPD regression would be run by fitting the following model:

$$\ln p^a_{ijklmn} = \alpha^0_{ijklm} + \alpha'_{ijklm} L^a + \beta'_{ijklm} x^a_{ijklmn} + \varepsilon^a_{ijklmn}.$$

As described above, the conventional technique is to use the estimates of the area dummy parameters $\hat{\alpha}_{ijklmn}$ as the item log prices to be used in further aggregation. Alternatively, from equation (13), the exact bilateral Törnqvist item index between areas a and b is

$$\ln T^{ab}_{ijklm} = -\sum_n \sum_z \frac{1}{2}(\beta^a_{ijklmnz} w^a_{ijklmn} + \beta^b_{ijklmnz} w^b_{ijklmn})(x^b_{ijklmnz} - x^a_{ijklmnz})$$
$$+ \sum_n \frac{1}{2}(w^a_{ijklmn} + w^b_{ijklmn})(\ln p^b_{ijklmn} - \ln p^a_{ijklmn}),$$

where the first term is a quality adjustment and the second is the familiar price index.

If the slopes are the same across areas, as in country-product-dummy models, the bilateral index reduces to

$$\ln T^{ab}_{ijklm} = \sum_n \frac{1}{2}(w^a_{ijklmn} + w^b_{ijklmn})\left(\ln p^b_{ijklmn} - \sum_z \beta_{ijklmnz} x^b_{ijklmnz}\right.$$
$$\left. - \left(\ln p^a_{ijklmn} - \sum_z \beta_{ijklmnz} x^a_{ijklmnz}\right)\right),$$

which is an index number of quality-adjusted specification prices. This is equivalent to the conventional practice of using the intercept estimates for each area as the quality-corrected area price level for the specifications within the item group since, from the CPD model, the area intercept coefficient for the item group can be expressed in terms of quality-corrected prices as

$$\alpha^0_{ijklm} + \alpha'_{ijklm} L^a = \ln p^a_{ijklmn} - \beta'_{ijklm} x^a_{ijklmn} - \varepsilon^a_{ijklmn}.$$

Although individual hedonic models by area are desirable, there may be insufficient data to obtain tight estimates of the coefficients or to identify the coefficients at all. In the first case, noisy coefficients can be estimated more accurately by blending them with a pooled regional regression. An example of this approach is set out in Randolph and Zieschang (1987) with application to a rent model for the CPI shelter component.

3.4.2 The EKS/CCD Approach

Caves, Christensen, and Diewert (1982a) show that application of the unweighted Eltetö, Köves, and Szulc approach to making a system of bilateral parities transitive is equivalent to choosing the reference shares and prices as

$$w^0_{ijklmn} = \frac{1}{A}\sum_a w^a_{ijklmn},$$

$$\ln p^0_{ijklmn} = \frac{1}{A}\sum_a \ln p^a_{ijklmn}.$$

A (preferable) version of the CCD formula would select the reference values as the weighted average across the area share in the next-higher-level aggregate, as in the following for aggregation of items to strata:

$$w^0_{ijklmn} = \sum_a s^a_{ijklm} w^a_{ijklmn},$$

$$\ln p^0_{ijklmn} = \sum_a s^a_{ijklm} \ln p^a_{ijklmn},$$

where

$$s^a_{ijklm} \equiv \frac{\sum_n p^a_{ijklmn} q^a_{ijklmn}}{\sum_a \sum_n p^a_{ijklmn} q^a_{ijklmn}}.$$

When quality-adjustment information is available, the reference hedonic prices (or coefficients) and item characteristics are determined (in weighted form) by

$$\beta^0_{ijklmnq} w^0_{ijklmn} = \sum_a s^a_{ijklm} \beta^a_{ijklmnq} w^a_{ijklmn},$$

$$x^0_{ijklmnq} = \sum_a s^a_{ijklm} x^a_{ijklmnq}.$$

3.4.3 A Regression Approach for Minimal Adjustment of the Data

An alternative to (or, as noted below, a likely superclass of) the EKS/CCD approach is to apply the proposition 1 transitivity condition directly. When this condition on the cross-weighted differences of log regional prices is not met, the data may be minimally adjusted to satisfy transitivity by fitting the following equation using least squares to obtain estimates $[\hat{w}^0_{ijklmn}, \hat{p}^0_{ijklmn}]$ for each specification n in item group $ijklm$:

$$w^a_{ijklmn} \ln p^b_{ijklmn} - w^b_{ijklmn} \ln p^a_{ijklmn} = w^0_{ijklmn}(\ln p^b_{ijklmn} - \ln p^a_{ijklmn}) - \ln p^0_{ijklmn}(w^b_{ijklmn} - w^a_{ijklmn}) + \varepsilon^{ab}_{ijklmn},$$

with the parameter restriction[2]

$$\sum_n w^0_{ijklmn} = 1.$$

Recalling that A is the number of areas in the region, there will be at most $A(A - 1)/2$ independent observations to estimate this equation for each specification *ijklmn*, and the model would be run as a stacked regression of specifications n within the item group *ijklm*.

2. This restriction is not required for transitivity, but it is required for aggregation consistency at the next level up and embodies an inherent property of the solution to the variant of the transitivity functional equation leading to the reference shares and prices form for the transitive system of bilateral Törnqvist index numbers.

In considering possible schemes for performing weighted estimation of the reference share and price parameters, each record could be weighted by the average importance of areas a and b at the next-higher-level (item) aggregate, that is, by

$$\sqrt{\frac{1}{2}(s^a_{ijklm} + s^b_{ijklm})}\,.$$

In this scheme, areas with higher overall shares for the item across the region would carry more weight in determining the estimated within-item specification reference shares and prices. This is reminiscent of, if distinct from, the weighting approach suggested by Selvanathan and Rao (1992) and is more transparent as to how a weighting methodology would actually work in a system of transitive Törnqvist parities—it affects the estimates of the reference shares and prices.[3]

When quality-adjustment information is available, the reference variables would be estimated in a way analogous to that for imposing transitivity in prices only as follows.

Estimate hedonic equations for each area as

$$\ln p^a_{ijklmn} = \alpha^a_{ijklm} + \beta^{a'}_{ijklm} x^a_{ijklmn} + \varepsilon^a_{ijklmn},$$

obtaining the estimates

$$\begin{bmatrix} \hat{\alpha}^a_{ijklm} \\ \hat{\beta}^a_{ijklm} \end{bmatrix}$$

for each area. Then, using least squares, estimate the vector

$$\begin{bmatrix} x^0_{ijklmnq} \\ -\beta^0_{ijklmnq} w^0_{ijklmn} \\ -\ln p^0_{ijklmn} \\ w^0_{ijklmn} \end{bmatrix}$$

by fitting the equation

$$w^a_{ijklm}\left(\ln p^b_{ijklmn} - \left(\sum_q \hat{\beta}^a_{ijklmnq} x^b_{ijklmn}\right)\right) - w^a_{ijklm}\left(\ln p^a_{ijklmn} - \left(\sum_q \hat{\beta}^b_{ijklmnq} x^a_{ijklmn}\right)\right)$$

$$= \sum_q [-\beta^0_{ijklmnq} w^0_{ijklmn}](x^b_{ijklmnq} - x^a_{ijklmnq}) + \sum_q x^0_{ijklmnq}(\hat{\beta}^b_{ijklmnq} w^b_{ijklm} - \hat{\beta}^a_{ijklmnq} w^a_{ijklm})$$

$$+ w^0_{ijklmn}(\ln p^b_{ijklmn} - \ln p^a_{ijklmn}) + [-\ln p^0_{ijklmn}](w^b_{ijklmn} - w^a_{ijklmn}) + \varepsilon^{ab}_{ijklmn}$$

3. Actually, there is probably a weighting scheme for the transitivity fitting equation that generates the EKS/CCD versions of the transitive Törnqvist system of parities, but it does not seem obvious how these weights would be determined.

with the parameter restriction

$$\sum_n w^0_{ijklmn} = 1.$$

There will be at most $A(A - 1)/2$ independent observations to estimate this equation for each specification *ijklmn*, and the model would be run as a stacked regression of specifications *n* within the item group *ijklm*. The observation weighting would follow the same scheme as in the simple case without specification characteristics and quality adjustment.

Notice that, if the hedonic slope coefficients are the same across areas for each specification characteristic, with the result that $\beta^a_{ijklmnq} = \beta^b_{ijklmnq} = \beta^0_{ijklmnq}$, then the estimating equation collapses to

$$w^a_{ijklm}\left(\ln p^b_{ijklmn} - \left(\sum_q \hat{\beta}^0_{ijklmnq} x^b_{ijklmn}\right)\right) - w^a_{ijklm}\left(\ln p^a_{ijklmn} - \left(\sum_q \hat{\beta}^0_{ijklmnq} x^a_{ijklmn}\right)\right)$$

$$= [w^0_{ijklmn}]\left(\ln p^b_{ijklmn} - \sum_q \hat{\beta}^0_{ijklmnq} x^b_{ijklmnq} - \left(\ln p^a_{ijklmn} - \sum_q \hat{\beta}^0_{ijklmnq} x^a_{ijklmnq}\right)\right)$$

$$+ \left[-\left(\ln p^0_{ijklmn} + \sum_q \hat{\beta}^0_{ijklmnq} x^0_{ijklmnq}\right)\right](w^b_{ijklmn} - w^a_{ijklmn}) + \varepsilon^{ab}_{ijklmn}.$$

In this case, the coefficient on the difference between the share vectors of the two areas is a *quality-adjusted reference price vector,* and no reference characteristics vector can be separately identified. It can, if desired, be independently determined as the EKS/CCD weighted average.

3.5 An Application to U.S. CPI Data

An empirical example of the methodology is provided by interarea prices for U.S. urban areas derived from hedonic regressions on data from the consumer price index. The CPI collects prices on a large sample of individual products, that is, on a probability sample of those specific products consumers are most likely to purchase in specific outlets in specific urban areas (see U.S. Department of Labor 1992). This approach results in a sample that is representative of household consumption choices but heterogeneous in nature. For example, the category for instant coffee consists of observations on instant coffee products of different sizes, brands, caffeine content, and other characteristics. In order to compare the prices of instant coffee across cities, these differences in characteristics must be explicitly accounted for, circumstances ideal for applying a hedonic regression approach. A recent major effort at the BLS has produced such interarea price indexes for most of the major categories of goods and services for forty-four U.S. urban areas; this effort is described in detail in Kokoski, Cardiff, and Moulton (1994). A more recent application of this approach to 1991 CPI data provided the data input for this example.

The CPI categories are organized in a hierarchical classification scheme as follows. The lowest level of aggregation, that at which individual prices are collected, is the entry-level item (ELI), designated by a five-digit code. The next highest level is the item stratum, designated by a four-digit code. Each stratum comprises one or more ELIs. The item strata are then aggregated into expenditure classes (ECs), designated by a two-digit code. These ECs may be organized into higher-level definitions, such as major groups, and then into the all-items CPI. For example, rice is ELI 01031, which is a member of the item stratum rice, pasta, and cornmeal (0103), which, in turn, is a member of EC 01, cereal and cereal products. Aggregating EC 01 and EC 02 provides the major group cereal and bakery products, which is part of the category food-at-home. The detailed classification structure for the entire CPI is provided in U.S. Department of Labor (1992).

For this example, we demonstrate the aggregation methodology on the food-at-home portion of the CPI. This group comprises eighty-eight ELIs, which are organized into eighteen ECs, and five major groups, providing the most detailed set of items in the CPI classification. For purposes of exposition, we will let the lowest level of aggregation, subscripted *ijklmn*, be represented by the ELIs (which, in many cases for food-at-home, map uniquely into item strata). The next highest level, subscripted *ijklm*, is represented by the expenditure class, and the ECs will be aggregated to a higher level, *ijkl*, the major groups. These major groups are then aggregated into an index for all food-at-home. This aggregation structure is presented in appendix B, table 3B.1. Table 3B.2 provides the expenditure shares for each major group as a component of all food-at-home and for each area as a proportion of all areas' total food-at-home.

The initial step is a log-linear hedonic regression, which was performed on each of these ELIs separately, as in equation (3) above:

$$\ln p^{a}_{ijklmn} = \alpha^{0}_{ijklm} + \alpha'_{ijklm} L^{a} + \beta'_{ijklm} x^{a}_{ijklmn} + \varepsilon^{a}_{ijklmn},$$

where $\ln p^{a}_{ijklmn}$ represents the log of the price of each item specification n in the mth ELI for the ath area, $x^{a}_{ijklmnz}$ are the variables defining the characteristics of each item specification, including the type of outlet where priced, and L^{a} is the dummy variable vector for area a. It has been shown (Summers 1973) that the exponentiated coefficients a_{ijklm} are bilateral price indexes for area a relative to the reference area (arbitrarily chosen as area $a = 1$). Space does not permit presentation of the results of the regressions for all eighty-eight ELIs, so the eight ELIs that compose fresh fruits and vegetables are provided as a representative example of the information available in the CPI data; these are presented in appendix C.

For exposition of the mechanics of the aggregation procedure, a simplified example for six ELIs and three areas is provided in appendix D. The six ELIs are apples, bananas, oranges, potatoes, lettuce, and tomatoes, which are aggregated into simplified hypothetical expenditure classes for fresh fruits and fresh vegetables. In table 3D.1 are presented the bilateral interarea indexes implied by the hedonic regression coefficients for each ELI. Table 3D.2 presents the expenditure shares that are used in the aggregation regression equation, and table 3D.3 provides the multilateral interarea index values that result from the aggregation. In this table, TORN*xy* is the index value comparing city *y* to city *x*. The ELIs are aggregated into ECs, and these ECs are then aggregated into a composite of the two (EC11 + EC12). The coefficient values from the aggregation regression equation are provided in tables 3D.4 and 3D.5, along with the adjusted R^2 values of the regression.

For comparison with the multilateral Törnqvist indexes in table 3D.3, a set of bilateral, and thus not necessarily transitive, parities was produced. These are provided in table 3D.6. This comparison shows, for this simplified example, the degree of empirical adjustment required to achieve the transitivity property. Recognizing that transitivity must be achieved at the cost of characteristicity, it is useful to assess this trade-off (see Dreschler 1973). In this case, the magnitude of the difference between the multilateral Törnqvist and its bilateral counterpart is less than 2 percent of the index value.

Appendix E contains results for the entire food-at-home group for all urban areas in the CPI. As in the example, the aggregation procedure is based on subsequent regressions at each level of aggregation. The adjusted R^2 values, given in table 3E.1, indicate that the equations fit well. (The value of 1.00 for EC08 occurs because there is only one ELI in that expenditure class.) Although the index values generated by the aggregation methodology are transitive, it is unwieldy to present the complete matrix of forty-four by forty-four area comparisons for each EC and major group. Thus, Philadelphia was arbitrarily chosen as the reference area for exposition of the results. The first set of index values, calculated by aggregating the ELIs into the expenditure classes, is presented in table 3E.2. The aggregation of these indexes into their respective major groups provides the index values in table 3E.3. The last column of this table provides the result of aggregating the five major groups into an all-food-at-home index. As expected, the indexes for Honolulu and Anchorage are well above the others, and, in general, the index values for the smaller urban areas are below that of the Philadelphia reference area, reflecting a priori expectations. As a result of the geometric averaging process, the variability of the index values is reduced at each higher level of aggregation.

3.6 Conclusion and an Extension

In this paper, we have considered the case of a single cross section of areas within which transitive bilateral quality-adjusted price comparisons are to be made between areas. We have also considered commodity aggregation within this framework, whereby transitivity is imposed while preserving a staged Törnqvist index aggregation rule. We have applied the technique to a small subset of the commodities priced in the U.S. consumer price index.

Our approach to transitivity has been a "minimum-data-adjustment" criterion with weighting specific to bilateral comparisons and therefore differs from other methods of imposing transitivity in a system of bilateral place-to-place Törnqvist index numbers. Although our method has some appeal because we can claim minimally to perturb the data in order to impose the transitive property with weighting sensitive to specific bilateral comparisons, the area expenditure-weighted sum of the log locality price levels will not necessarily be equal to zero, in contrast with the EKS/CCD approach, which satisfies this property by construction. The need for this property, as well as operational considerations such as ease of computation and calculation of measures of precision, would need to be weighed in deciding on an estimator for production of a regular statistical series of interarea price indexes. Before closing, we would like briefly to describe a promising avenue of research using this framework in a time-series context.

3.6.1 The Single Chain Link Case

It has been a problem in the interpretation of data from the International Comparisons Project that the change in the levels of real GDP implied by the international purchasing power parities from time to time has not been the same as the growth in national GDPs measured by direct deflation using a(n implicit) time-series GDP deflator. We consider here a remedy within the Törnqvist system of interarea and time-series index numbers by considering a system of area indexes that are transitive both among areas within the same time period and between areas from differing time periods. A direct implication of this is that the index change between two periods for a given area, say, a, can be expressed as the product of the relative level between two areas, say, a and b, in the first period, times the relative change in b between the two periods for any two areas a and b.

The Törnqvist item index $\ln T_{ijklm}^{ab,uv}$ between area a in time period u and area b in time period v, where $u, v \in \{t - 1, t\}$, is

$$\ln T_{ijklm}^{ab,uv} = \sum_{n} \frac{1}{2}(w_{ijklmn}^{au} + w_{ijklmn}^{bv})(\ln p_{ijklmn}^{bv} - \ln p_{ijklmn}^{au}).$$

It is straightforward to see that, for the system of between-area, between-period parities to be transitive, proposition 1 applies directly in this case, with reference share and price vectors determined for the union of the two time periods and collections of areas. If quality adjustments are possible using hedonic regressions, then proposition 2 can be applied to show the transitive form of the quality-adjusted system of parities as a function of a reference share, price and hedonic coefficients vectors across areas and time periods. We note below that, under international decentralization of compilation, the country hedonic regression coefficients would generally not be the same as in the CPD approach.

An additional comparison generally computed in this case is the change over time of the regional aggregate of areas. Examples of such indexes would be national consumer price and producer price and labor compensation indexes as composites of the subnational areas sampled to obtain the data. This index can be written as

$$\begin{aligned}\ln T^{R,uv}_{ijklm} &= \sum_a \frac{1}{2}(s^{au}_{ijklm} + s^{av}_{ijklm}) \ln T^{aa,uv}_{ijklm} \\ &= \sum_a \frac{1}{2}(s^{au}_{ijklm} + s^{av}_{ijklm}) \sum_n \frac{1}{2}(w^{au}_{ijklmn} + w^{av}_{ijklmn})(\ln p^{av}_{ijklmn} - \ln p^{au}_{ijklmn}).\end{aligned}$$

By period-to-period and interarea transitivity

$$\ln T^{R,uv}_{ijklm} = \sum_a \frac{1}{2}(s^{au}_{ijklm} + s^{av}_{ijklm})(\ln \bar{p}^{av}_{ijklmn} - \ln \bar{p}^{au}_{ijklmn}),$$

where

$$\ln \bar{p}^{au}_{ijklm} = \sum_n \frac{1}{2}(w^{00}_{ijklmn} + w^{au}_{ijklmn})(\ln p^{au}_{ijklmn} - \ln p^{00}_{ijklmn}).$$

The aggregate time-series index under period/area transitivity between the two periods is, therefore, a weighted average of the relative change in a set of area price levels, ensuring consistency between the levels within period across area and rates of change between periods within area.

3.6.2 Time-Series/Cross-Sectional Transitivity over Multiple Periods

Clearly, the single chain link, two-period case can be extended to the multiple-period case by pooling the data for multiple periods. A distinct advantage of the application of this procedure is that the problem of chain drift is eliminated over the multiple-period epoch being adjusted while maintaining much of the period specificity of the weight and price components of the Törnqvist index formula. The reason is that transitivity eliminates drift, which is usually defined as the persistent deviation of a direct index between nonadjacent periods as compared with the product of adjacent period chain links covering the multiple period interval. An issue to be resolved in applying this technique is that it refers to a moving window of a fixed time duration. Data passing outside the window would not exactly satisfy the transitive property. Choosing the window as a long-enough period could be expected to result in very slow change in the reference prices and shares, however, so that the effect could be minimized, at the cost of providing less of Drechsler's (1973) "characteristicity" for relatively recent time periods.

3.6.3 Decentralized Computing of International Parities While Controlling for the Quality of Goods Available in Different Countries

The methodology outlined here, which uses a hedonic, characteristics-based quality-adjustment procedure, permits decentralized, within-country estimation of the hedonic equation coefficients. This is especially attractive in view of the great and generally justified reluctance with which most statistical offices grant access to the micro-data sources of their price indexes. The prerequisite for this would be that a standard product classification would have to be adopted by all countries and also that, with each product class, a standard list of product characteristics or specification measures would have to be adopted. One such set of standards might be derived by merging the U.S. CPI specification file, listing the characteristics measures for some 365 product categories, with a standard international commodity classification, such as the central product classification or CPC of the United Nations, itself a superset of the now standard harmonized classification for internationally traded commodities.

A compilation strategy such as this for the ICP would have a distinct advantage over the current approach of pricing a long, detailed list of narrowly specified items. The number of product strata required would be smaller, and the countries could use the estimates for their own, internal quality-adjustment needs for time-series and within-country geographic comparisons.

Appendix A
Proofs of Propositions

Proof of Proposition 1

The proof of this proposition follows methods used in, for example, Aczel (1966) and Eichhorn (1978). First, we establish the following solution of the *transitivity (functional) equation* for all single-valued functions g of two vectors of identical dimension in an argument set D that satisfy an identity condition $g(x, x) = 0$:

$$g(x, y) + g(y, z) = g(x, z) \quad \text{and} \quad g(x, x) = 0, \quad \forall (x, y, z) \in D,$$

if and only if

$$g(x, y) = h(y) - h(x).$$

Let $y = y^0$. Then, for all x and z in the domain of g,

$$g(x, y^0) + g(y^0, z) = r(x) + h(z) = g(x, z).$$

Substituting this back into the transitivity equation,

$$r(x) + h(y) + r(y) + h(z) = g(x, z).$$

By identity

$$g(y, y) = r(y) + h(y) = 0,$$

and, hence,

$$r(y) = -h(y).$$

We can now express $g(x, z)$ in terms of h as

$$g(x, z) = h(z) - h(y),$$

yielding the desired result.

From this, transitivity of the Törnqvist bilateral relative requires that, for all a, b,

$$\ln T^{ab} = h(\vec{w}^b, \vec{p}^b) - h(\vec{w}^a, \vec{p}^a).$$

Expanding the bilateral relative expression, we have

$$\begin{aligned} \ln T^{ab} &\equiv \sum_i \frac{1}{2}(w_i^a + w_i^b)(\ln p_i^b - \ln p_i^a) \\ &= \frac{1}{2}\sum_i (w_i^b \ln p_i^b - w_i^b \ln p_i^b + w_i^a \ln p_i^b - w_i^b \ln p_i^a) \\ &= h(\vec{w}^b, \vec{p}^b) - h(\vec{w}^a, \vec{p}^a). \end{aligned}$$

We set $b = 0$ and solve for $h(\overrightarrow{w^a}, \overrightarrow{p^a})$ in terms of reference area 0 as

$$\begin{aligned} h(\vec{w}^a, \vec{p}^a) = {} & \left[h(\vec{w}^0, \vec{p}^0) - \frac{1}{2}\sum_i w_i^0 p_i^0\right] + \frac{1}{2}\sum_i w_i^0 \ln p_i^a - \frac{1}{2}\sum_i \ln p_i^0 w_i^a \\ & + \frac{1}{2}\sum_i w_i^a p_i^a . \end{aligned}$$

Substituting this into the expanded equation for the transitive bilateral log parity, multiplying through by two, and subtracting $w_i^b \ln p_i^b - w_i^a \ln p_i^a$ inside the summations from both sides, we have

$$\sum_i (w_i^a \ln p_i^b - w_i^b \ln p_i^a) = \sum_i w_i^0(\ln p_i^b - \ln p_i^a) - \sum_i \ln p_i^0(w_i^b - w_i^a).$$

The expression for the transitive Törnqvist bilateral parity obtains by substituting this expression for the cross-product between the area weights and the prices into the expanded expression for the parity, adding and subtracting the term $\Sigma_i w_i^0 \ln p_i^0$, and collecting terms.

Q.E.D.

Proof of Proposition 2

The proof of proposition 2 follows very closely that of proposition 1. It is easy to see from this that the price level for each area now has a price- and quality-adjustment component.

Appendix B
Structure and Expenditure Shares of Division "Food-at-Home"

Table 3B.1 CPI Classification Structure for Food-at-Home as Aggregated for Interarea Indexes

Group 1: Cereal and bakery products	
EC01 Cereal and cereal products	1011 Flour
	1012 Prepared flour mixes
	1021 Cereal
	1031 Rice
	1032 Macaroni, similar products, and cornmeal
EC02 Bakery products	2011 White bread
	2021 Bread other than white
	2022 Rolls, biscuits, and muffins (excluding frozen)
	2041 Cakes and cupcakes (except frozen)
	2042 Cookies
	2061 Crackers
	2062 Bread and cracker products
	2063 Sweetrolls, coffee cake, and doughnuts (excluding frozen)
	2064 Frozen bakery products and frozen/refrigerated doughs and batters
	2065 Pies, tarts, turnovers (excluding frozen)
Group 2: Meat, poultry, fish, and eggs	
EC03 Beef and veal	3011 Ground beef
	3021 Chuck roast
	3031 Round roast
	3041 Other roasts (excluding chuck and round)
	3042 Other steak (excluding round and sirloin)
	3043 Other beef
	3051 Round steak
	3061 Sirloin steak
EC04 Pork	4011 Bacon
	4021 Pork chops
	4031 Ham (excluding canned)
	4032 Canned ham
	4041 Pork roast, picnics, other pork
	4042 Pork sausage
EC05 Other meats	5011 Frankfurters
	5012 Bologna, liverwurst, salami
	5013 Other lunchmeats (excluding bologna, liverwurst, salami)
	5014 Lamb, organ meats, and game
EC06 Poultry	6011 Fresh whole chicken
	6021 Fresh and frozen chicken parts
	6031 Other poultry

Table 3B.1 (continued)

EC07 Fish and seafood	7011 Canned fish or seafood 7021 Shellfish (excluding canned) 7022 Fish (excluding canned)
EC08 Eggs	8011 Eggs
Group 3: Dairy products	
EC09 Fresh milk and cream	9011 Fresh whole milk 9021 Other fresh milk and cream
EC10 Processed dairy products	10011 Butter 10012 Other dairy products 10021 Cheese 10041 Ice cream and related products
Group 4: Fruits and vegetables	
EC11 Fresh fruits[a]	11011 Apples 11021 Bananas 11031 Oranges 11041 Other fresh fruits
EC12 Fresh vegetables[a]	12011 Potatoes 12021 Lettuce 12031 Tomatoes 12041 Other fresh vegetables
EC13 Fruit juices and frozen fruit	13011 Frozen orange juice 13012 Other frozen fruits and fruit juices 13013 Fresh, canned, or bottled fruit juices 13031 Canned and dried fruits
EC14 Processed vegetables	14011 Frozen vegetables 14021 Canned beans other than lima beans 14022 Canned cut corn 14023 Other processed vegetables
Group 5: Other foods-at-home	
EC15 Sugar and sweets	15011 Candy and chewing gum 15012 Other sweets (excluding candy and chewing gum) 15021 Sugar and artificial sweeteners
EC16 Fats and oils	16011 Margarine 16012 Other fats and oils 16013 Nondairy cream substitutes 16014 Peanut butter
EC17 Nonalcoholic beverages	17011 Cola drinks 17012 Carbonated drinks other than cola 17031 Roasted coffee 17032 Instant and freeze-dried coffee 17051 Noncarbonated fruit-flavored drinks 17052 Tea 17053 Other noncarbonated drinks

(continued)

Table 3B.1 (continued)

EC18 Other prepared food	18011 Canned and packaged soup 18021 Frozen prepared meals 18022 Frozen prepared food other than meals 18031 Potato chips and other snacks 18032 Nuts 18041 Salt and other seasonings and spices 18042 Olives, pickles, relishes 18043 Sauces and gravies 18044 Other condiments (excluding olives, pickles, relishes) 18061 Canned or packaged salads and desserts 18062 Baby food 18063 Other canned or packaged prepared foods

[a]Items for which detailed calculations are shown in apps. C and D.

Table 3B.2 **Expenditure Shares for Food-at-Home, 1991**

A. Shares of major groups in food-at-home	
Cereal and bakery products	.13876
Meat, poultry, fish, and eggs	.27262
Dairy products	.13245
Fruits and vegetables	.18144
Other food-at-home	.27473
B. Food-at-home shares of each area in national food-at-home	
Northeast region	.211814
Philadelphia	.025485
Boston	.016779
Pittsburgh	.011474
Buffalo	.008681
New York City	.027687
New York–Connecticut suburbs	.021090
New Jersey suburbs	.029169
Northeast, B PSUs	.033929
Northeast, C PSUs	.027482
Northeast, D PSUs	.010038
North Central region	.224254
Chicago	.040356
Detroit	.020599
St. Louis	.011198
Cleveland	.013992
Minneapolis	.010670
Milwaukee	.014031
Cincinnati	.009500
Kansas City	.008020
North Central, B PSUs	.028207
North Central, C PSUs	.038570
North Central, D PSUs	.029111

Table 3B.2 (continued)

South region	.283384
Washington, D.C.	.017232
Dallas	.019052
Baltimore	.013888
Houston	.018058
Atlanta	.010608
Miami	.013914
Tampa	.013365
New Orleans	.005904
South, B PSUs	.076899
South, C PSUs	.061897
South, D PSUs	.032567
West region	.280549
Los Angeles County	.061876
Greater Los Angeles	.061876
San Francisco	.035026
Seattle	.014656
San Diego	.012288
Portland, Oreg.	.013951
Honolulu	.002924
Anchorage	.001103
Denver	.008912
West, B PSUs	.026148
West, C PSUs	.025536
West, D PSUs	.016253

Note: PSU = primary sampling unit.

Appendix C
Hedonic Regression Results for Fresh Fruits and Vegetables

Table 3C.1 Hedonic Regression Results for Fresh Fruit

Variable	Coefficient			
	11011 Apples	11021 Bananas	11031 Oranges	11041 Other Fresh Fruit
Mean of dependent variable: log price	−2.8886	−3.5118	−2.8848	−2.7285
Adjusted R^2	.3329	.3314	.3403	.5932
Sample size	9,423	6,791	12,610	26,069
Area				
Philadelphia-Wilmington-Trenton, PA-DE-NJ-MD	REF	REF	REF	REF
Boston-Lawrence-Salem, MA-NH	−.11302*	−.01941	.06634	−.03617
Pittsburgh–Beaver Valley, PA	−.07054	−.18766*	.03833	−.10519*
Buffalo–Niagara Falls, NY	−.20282*	−.01421	−.00287	−.20930*
New York City	.06071*	.01522	.13458*	−.01446
New York–Connecticut suburbs	−.05384*	−.02748	−.02681	−.02877
New Jersey suburbs	−.02993	−.00491	−.12221*	−.05622*
Northeast region, B size PSUs	−.04942	−.05294	.01369	−.03246
Northeast region, C size PSUs	−.09681*	−.13774*	−.03222	−.13600*
Northeast region, D size PSUs	−.08744*	.06198	.18581*	.06941*
Chicago–Gary–Lake County, IL-IN-WI	.06099*	−.02261	.23166*	−.00033
Detroit–Ann Arbor, MI	−.04564	−.22780*	.01161	−.16026*
St. Louis–East St. Louis, MO-IL	−.02697	.04932	.06972	−.02078
Cleveland-Akron-Lorain, OH	−.16546*	−.14305*	.02411	−.01462
Minneapolis–St. Paul, MN-WI	−.00094	−.23664*	.07582	−.09512*

Milwaukee, WI	.01006	−.11808*	.01591	−.04899
Cincinnati–Hamilton, OH-KY-IN	−.08656	−.18154*	.01295	−.00455
Kansas City, MO–Kansas City, KS	−.01494	−.10477*	.44553*	−.05681
North Central region, B size PSUs	.01421	−.18972*	.19250*	−.03791
North Central region, C size PSUs	−.05190	−.23027*	.09794*	−.09401*
North Central region, D size PSUs	−.15361*	−.08938*	−.11770*	−.20600*
Washington, DC–MD–VA	−.01403	.01819	−.00686	.08002*
Dallas–Fort Worth, TX	−.05908*	−.25484*	−.14995*	−.03552
Baltimore, MD	−.04353	−.05988	−.02344	.03504
Houston-Galveston-Brazoria, TX	−.11871*	−.12477*	−.22791*	−.18318*
Atlanta, GA	−.00259	−.30016*	.11588*	.01898
Miami–Fort Lauderdale, FL	−.03459	−.50646*	−.3650*	−.31861*
Tampa–St. Petersburg–Clearwater, FL	.05230	−.31348*	−.65656*	−.07724*
New Orleans, LA	−.02976	−.00361	−.02918	−.00581
South region, B size PSUs	−.01420	−.20297*	.01788	−.07499*
South region, C size PSUs	−.10217*	−.22290*	−.08665*	−.10318*
South region, D size PSUs	−.04820*	−.04390	−.06157*	−.08453*
Los Angeles County, CA	−.17093*	−.02534	−.08845*	−.12754*
Greater Los Angeles, CA	−.20649*	−.10411*	−.07800*	−.07642*
San Francisco–Oakland–San Jose, CA	−.19800*	−.08328*	.06305	−.02541
Seattle-Tacoma, WA	.19215*	.12307*	.05389	.15074*
San Diego, CA	−.17869*	−.17004*	−.04144*	−.12074*
Portland-Vancouver, OR-WA	−.17181*	−.12146*	.10993	.01383
Honolulu, HI	.00481	.55898*	.22252*	.11750*
Anchorage, AK	−.10670	.43803*	.35183*	.12539*
Denver-Boulder, CO	−.03535	.08929	.21348*	.02031
West region, B size PSUs	−.11932*	−.05745	.06924	.02278
West region, C size PSUs	−.12937*	−.09631*	.04585	−.05945
West region, D size PSUs	−.07509*	−.27655*	−.02118	−.08199*

(continued)

Table 3C.1 (continued)

Variable	Coefficient 11011 Apples	11021 Bananas	11031 Oranges	11041 Other Fresh Fruit
Rotation group				
Same sample as previous month	REF	REF	REF	REF
New sample	−.03224*	−.01012	−.04340*	.04082*
Month of collection (11 months of data)				
January	REF	REF	REF	REF
February	.03022*	.09711*	.08286*	−.04646*
March	.04515*	.27093*	.12605*	−.06279*
April	.05623*	.22156*	.14488*	−.01630
May	.10186*	.23920*	.14072*	−.01466*
June	.15730*	.15363*	.25007*	−.05689*
July	.19332*	.12148*	.29341*	−.14868*
August	.20836*	−.08925*	.34215*	−.26625*
September	.17460*	−.04995*	.39708*	−.23102*
October	.01422	−.14127*	.23935*	−.18405*
November	.03296*	−.05280*	−.04562*	−.13335*
Outlet type				
Chain grocery	REF	REF	REF	REF
Independent grocery stores	−.07482*	−.04932*	−.10642*	−.04624*
Full service department stores	−.11568	−.23667	−.00639	.24117*
Produce market	−.18148*	−.13124*	−.15877*	−.14419*
Convenience stores	.05972	.39717*	.01025	−.00807
Commodity oriented outlet not elsewhere classified	−.41615*	−.23223*	−.61553*	−.00207
Outlet not elsewhere classified	−.41217*	−.60950*	−.95862*	−.72177*

Package type				
Packaging: loose	−.07769			
Packaging multi-pack	−.33425*		−.19648*	.00157
Packaging: single item, individually wrapped	−.43312*		−.06804	.04563*
Other	REF		REF	REF
Package size				
0–10 pounds	REF			
Above 10 pounds	.01042			
Size represents: weighed one multipack	−.09754*		−.11644*	−.39775*
Size represents: weight labeled	−.03956	−.03506*	−.10879*	−.33027*
Size represents: weighed one bunch		REF		
Size: weigh 2 items	−.05256		−.22778*	−.31620*
Size: other			REF	REF
Grade				
Store seconds or other than first quality		.06409		
First quality or class		.01581		
U.S. fancy			.01961*	
U.S. extra fancy	.01337			.04234*
Other grade/grade not available	REF	REF	REF	REF
Variety				
Delicious	.00968			
Golden delicious	−.03918			
Red delicious	−.04676			
Other delicious	[a]			
Granny Smith	.03385*			
Gravenstein	−.17583*			
Jonathan	−.17352*			
McIntosh	−.05280*			
Rome Beauty (Red Rome)	−.03041			

(continued)

Table 3C.1 (continued)

Variable	Coefficient 11011 Apples	11021 Bananas	11031 Oranges	11041 Other Fresh Fruit
Stayman	−.11321*			
Winesap	−.06369*			
York (York Imperial)	.74756*			
Navel			.29912*	
Temple			−.11816*	
Valencia			.14359*	
Tangelo			.28032*	
Tangerine			.44601*	
Avocados				.19300*
Berries				.28592*
Blueberries				.25617*
Cranberries				−.02272
Raspberries				.99209*
Strawberries				−.30166*
Cherries (sweet/tart)				.28835*
Grapefruit				−.90631*
Pink grapefruit				−.03686
Red (ruby) grapefruit				−.06126
White (yellow) grapefruit				−.12095

Grapes				−.18645*
Red (flame) seedless grapes				−.00638
Emperor or tokay grapes				−.05662
Rebier grapes				.04156
Concord grapes				−.00026
Thompson seedless				.05003
Lemons				−.26529*
Limes				−.28323*
Melons				−.82440*
Watermelon				−.71551*
Cantaloupe melons				−.19234*
Honeydew melons				−.01389
Casaba melons				−.07582
Crenshaw melons				.29155*
Persian melons				.26967*
Santa Claus melons				−.08431
Peaches				−.44508*
Pears				−.19322*
Anjou pears				−.46646*
Bartlett pears				−.45497*
Bosc pears				−.34758*
Seckel pears				.01742
Pineapples				−1.02458*
Plums				−.21393*
Other	REF	REF	REF	REF

Note: REF = reference.

[a]There were insufficient records with this characteristic to include it in the model.

*Statistically significant at the 5 percent level.

Table 3C.2 **Hedonic Regression Results for Fresh Vegetables**

	Coefficient			
	12011 Potatoes	12021 Lettuce	12031 Tomatoes	12041 Other Fresh Vegetables
Mean of dependent variable: log price	−3.7367	−3.1222	−2.7011	−3.0502
Adjusted R^2	.6228	.5385	.1991	.4872
Sample size	6,770	6,764	6,769	14,241
Variable				
Area				
Boston-Lawrence-Salem, MA-NH	−.06387	−.29268*	.01678	−.18035*
Pittsburgh–Beaver Valley, PA	−.23146*	−.25850*	−.05929	−.19013*
Buffalo–Niagara Falls, NY	−.07529	−.23543*	−.13822*	.16605*
New York City	−.11552*	−.04737	−.02569	.08782*
New York–Connecticut suburbs	−.08816*	−.14394*	−.13951*	−.11219*
New Jersey suburbs	−.07603*	−.12854*	−.00121	−.07710*
Northeast region, B size PSUs	−.02074	−.15127*	−.06827*	−.09467*
Northeast region, C size PSUs	−.08409*	−.13739*	−.13375*	−.23278*
Northeast region, D size PSUs	−.00174	−.05323	.19425*	−.12643*
Chicago–Gary–Lake County, IL-IN-WI	.15525*	−.08079*	−.05506	.07671*
Detroit–Ann Arbor, MI	−.20893*	−.25233*	−.12991*	−.23765*
St. Louis–East St. Louis, MO-IL	.17335*	−.10762	−.24150*	.04041
Cleveland-Akron-Lorain, OH	−.15117	−.36251*	−.33654*	−.25811*
Minneapolis–St. Paul, MN-WI	−.44411*	−.37324*	−.19602*	−.06638*
Milwaukee, WI	−.18660	−.28823*	−.30839*	−.10271*
Cincinnati–Hamilton, OH-KY-IN	−.10455	.04950	−.04369	−.04813
Kansas City, MO–Kansas City, KS	−.12530*	−.19183*	−.15001	−.01079
North Central region, B size PSUs	−.16171*	−.25840*	−.12224*	−.18928*

North Central region, C size PSUs	−.18821*	−.27277*	−.22069*	−.10230*
North Central region, D size PSUs	−.35833*	−.31699*	−.24259*	−.33089*
Washington, DC–MD–VA	.03654	−.05585	.00536	.05914
Dallas–Fort Worth, TX	−.18926*	−.10340*	−.27754*	−.03080
Baltimore, MD	−.14338*	−.15619	−.09699	−.13070*
Houston-Galveston-Brazoria, TX	.04671	−.04106	−.1059*	−.21465*
Atlanta, GA	−.05600	−.22134*	−.05248	−.16305*
Miami–Fort Lauderdale, FL	−.15124*	−.27921*	−.50285*	−.15624*
Tampa–St. Petersburg–Clearwater, FL	−.04608*	−.33078*	−.16919*	−.16947*
New Orleans, LA	−.31242*	−.37557*	−.14505*	−.12714*
South region, B size PSUs	−.16423*	−.15786*	−.20441*	−.11132*
South region, C size PSUs	−.11572*	−.16706*	−.28840*	−.15629*
South region, D size PSUs	−.16923*	−.05474	−.21121*	−.06746*
Los Angeles County, CA	−.06874	−.42383*	−.32307*	−.30328*
Greater Los Angeles, CA	−.02597	−.46923*	−.39316*	−.30388*
San Francisco–Oakland–San Jose, CA	−.03646*	−.42209*	−.25084*	−.25710*
Seattle-Tacoma, WA	−.21794*	−.31179*	−.29454*	.28713*
San Diego, CA	−.26552*	−.63145*	−.52545*	−.41637*
Portland-Vancouver, OR-WA	−.19012*	−.36026*	−.17875*	−.41273*
Honolulu, HI	.57928*	.11241*	.04541	.52045*
Anchorage, AK	.31145*	.24406	.13065	.34271*
Denver-Boulder, CO	.07387	−.01414	−.11228*	.04067
West region, B size PSUs	−.27332*	−.39816*	−.31270*	−.38406*
West region, C size PSUs	−.15643*	−.30573*	−.28317*	−.13431*
West region, D size PSUs	−.28844*	−.43150*	−.46059*	−.21889*
Rotation group				
Same sample as previous month	REF	REF	REF	REF
New sample	−.02658	−.00552	−.02653	−.07898

(continued)

Table 3C.2 (continued)

	Coefficient			
	12011 Potatoes	12021 Lettuce	12031 Tomatoes	12041 Other Fresh Vegetables
Month of collection (11 months of data)				
January	REF	REF	REF	REF
February	.01009	−.17981*	−.06562*	−.04138*
March	−.00178	−.27424*	.03005	−.06361*
April	.02363	−.19154*	.24313*	.04834*
May	.06245*	−.10586	.38750*	−.03054
June	.16776*	−.02842*	.54470*	.01922
July	.17522*	−.29127*	.20950*	−.07667*
August	.11864*	−.32081*	−.19194*	−.14672*
September	.01893	−.29942*	−.18685*	−.17824*
October	−.07231*	−.28473*	−.25130*	−.18042*
November	−.09234	.02991	−.12365*	−.10289*
Outlet type				
Chain grocery	REF	REF	REF	REF
Full service department stores	−.61027*	.12446	−.19366	.02692
Independent grocery stores	−.02169	−.02528	−.05455*	−.03887
Produce market	−.12607*	−.08024*	−.25519*	−.12645*
Convenience stores	.27450*	.30320*	.57844*	.21838*
Commodity oriented outlet not elsewhere classified	−.39201*	−.49838*	−.73586*	−.35484*
Outlet not elsewhere classified	−.31633*	−.66851*	−.69563*	−.55401*

Package type				
Packaging: loose	−.78494*			
Trimmed				−.02206
Packaging: single item, individually wrapped			.48149*	
Packaging: multipack		−.12828	.16951*	.36462*
Packaging: multipack, weight: 0–9.999 lb.	.35237*			
Packaging: multipack, weight: greater than or equal to 1 lb.	−.51911*			
Other	REF	REF	REF	REF
Package size				
Size represents: weighed one multipack		.17900*		−.44977*
Weighed 2 potatoes	−.32016*		−.14385*	
Weight labeled	−.10378			
Other	REF	REF	REF	REF
Variety				
White potato	−.11197			
Round or long russet	−.12073*			
Round or long white	−.08151*			
Round red	.13440*			
Baking potato	.07752*			
Yam	−.25983*			
Sweet potato/yam	.28265			
Sweet potato	−.45422*			
Unable to determine variety	−.40993*			

(continued)

Table 3C.2 (continued)

	Coefficient			
	12011 Potatoes	12021 Lettuce	12031 Tomatoes	12041 Other Fresh Vegetables
Bibb		.59326*		
Boston		.13274*		
Butterhead		.91953*		
Cos/Romaine		.36190*		
Green leaf		.60410*		
Red leaf		.64647*		
Unspecified variety			−.05647*	
Field grown/vine-ripened			−.30496*	
Hot house or greenhouse			−.26962*	
Unable to determine type			−.26013*	
Radishes with tops				−.01645
Radishes without tops				−.76580*
Yellow corn				−.50855*
White corn				−.13763*
Artichokes				.56677*
Asparagus				.58202*
Bean sprouts				.04937
Miniature carrots				.11278
Green snap beans				−.01606

Pole beans				−.05602
Yellow wax beans				.42235*
Lima beans				.60288*
Domestic (green) cabbage				−1.01275*
Savoy (crinkled leaf) cabbage				−.68162*
Chinese (celery) cabbage				−.12622
Hearts of celery				−.02623
Yellow onions				−.95681
White onions				−.38899*
Pickling cucumbers				−.04579
Spaghetti squash				−.49575*
Yellow straightneck squash				−.08479
Yellow crookneck squash				.06090
Butternut squash				−.31871*
Acorn squash				−.29270*
Zucchini (Italian) squash				−.11483
Green peppers				.20108*
Regular mushrooms				−.27425*
Spanish onion				−.52453*
Red onion				−.11692*
Other	REF	REF	REF	REF

Note: REF = reference.

*Statistically significant at the 5 percent level.

Appendix D
Sample Index Calculation for Fruits and Vegetables and Three Areas

Table 3D.1 CPD Results for Bilateral Relatives for Fruits and Vegetables Entry-Level Items (ELIs)

ELI	Description	AREA1 (PHILA)	AREA2 (BOSTON)	AREA3 (PITTSBG)
11011	Apples	1.0000	.89618	.96884
11021	Bananas	1.0000	.98656	.83609
11031	Oranges	1.0000	1.04715	1.05667
12011	Potatoes	1.0000	.95105	.81712
12021	Lettuce	1.0000	.78312	.75877
12031	Tomatoes	1.0000	1.03136	.95217

Table 3D.2 Expenditures Shares within and across Areas

Share Type	AREA1 (PHILA)	AREA2 (BOSTON)	AREA3 (PITTSBG)
	ELI Shares by Area		
W11011	.39638	.40456	.52877
W11021	.33412	.27270	.31936
W11031	.26950	.32273	.15187
W12011	.41063	.36953	.35684
W12021	.29234	.34244	.37069
W12031	.29704	.28803	.27247
	Expenditure Class Shares by Area		
S(EC11)	.46262	.33328	.20410
S(EC12)	.42243	.38425	.19331

Note: ELI = entry-level items.

Table 3D.3 Multilateral Törnqvist Indexes

Item	TORN12	TORN13	TORN21	TORN23	TORN31	TORN32
EC11	.9614	.9455	1.0402	.9835	1.0576	1.0167
EC12	.9140	.8345	1.0941	.9131	1.1983	1.0952
EC11+EC12	.9382	.8897	1.0659	.9483	1.1240	1.0545

Note: EC = expenditure class.

Table 3D.4 **Estimated Reference Shares and Prices at the ELI Level**

EC/ELI	Variable	Coefficient Estimate
Fruits[a]		
Apples	W0 (11011)	.4391
Bananas	W0 (11021)	.3094
Oranges	W0 (11031)	.2515
Apples	P0 (11011)	−.0458
Bananas	P0 (11021)	−.0585
Oranges	P0 (11031)	.0318
Vegetables[b]		
Potatoes	W0 (12011)	.3803
Lettuce	W0 (12021)	.3331
Tomatoes	W0 (12031)	.2866
Potatoes	P0 (12011)	−.0780
Lettuce	P0 (12021)	−.1671
Tomatoes	P0 (12031)	−.0041

Note: EC = expenditure class. ELI = entry-level item.

[a]Adjusted R^2 = 0.9708.

[b]Adjusted R^2 = 0.9980.

Table 3D.5 **Estimated Reference Expenditure Shares and Prices at the Expenditure Class**

EC	Variable	Coefficient Estimate
Fruits[a]	W0 (EC11)	.5041
	P0 (EC11)	−.0005
Vegetables[a]	W0 (EC12)	.4959
	P0 (EC12)	−.0006

Note: EC = expenditure class.

[a]Adjusted R^2 = 0.9947.

Table 3D.6 **Unadjusted Bilateral Törnqvist Indexes**

EC	TORN12	TORN13	TORN21	TORN23	TORN31	TORN32
EC11	.96622	.94033	1.03496	.98959	1.06346	1.01052
EC12	.91564	.83279	1.09213	.91505	1.20078	1.09284

Note: EC = expenditure class.

Appendix E
Results for Division Food-at-Home and 44 CPI Areas

Table 3E.1 Adjusted R^2 of Transitivity Regressions at Each Aggregation Level

Group	Adjusted R^2	Definition	Number of Subaggregate Items
		Expenditure Class (subaggregate item = entry-level item [ELI])	
EC1	.8333	Cereals and cereal products	5
EC2	.9266	Bakery products	10
EC3	.9494	Beef and veal	8
EC4	.9395	Pork	6
EC5	.9717	Other meats	4
EC6	.9737	Poultry	3
EC7	.8940	Fish and seafood	3
EC8	1.0000	Eggs	1
EC9	.9673	Fresh milk and cream	2
EC10	.9819	Processed dairy products	4
EC11	.9832	Fresh fruits	4
EC12	.9918	Fresh vegetables	4
EC13	.9567	Processed fruits	4
EC14	.9682	Processed vegetables	4
EC15	.9791	Sugar and sweets	3
EC16	.9722	Fats and oils	4
EC17	.9443	Nonalcoholic beverages	7
EC18	.9444	Other prepared foods	12
		Major Group (subaggregate item = expenditure class [EC])	
CERBAK	.9936	Cereal and bakery products (EC01–EC02)	2
MPFE	.9839	Meat, poultry, fish, and eggs (EC03–EC08)	6
DAIRY	.9952	Dairy products (EC09–EC10)	2
FRTVEG	.9902	Fruits and vegetables (EC11–EC14)	4
OTHER	.9721	Other food at home (EC15–EC18)	4
		Division (subaggregate item = major group)	
FOOD	.9934	All food at home	5

Table 3E.2 **Index Values Aggregated by Expenditure Class, Reference Area Philadelphia**

Area	EC1	EC2	EC3	EC4	EC5	EC6	EC7	EC8	EC9	EC10	EC11	EC12	EC13	EC14	EC15	EC16	EC17	EC18
Philadelphia	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Boston	.957	.974	1.012	.889	.899	1.137	.977	.904	.939	.852	.983	.840	.854	.927	.798	.966	.950	.759
Pittsburgh	.864	.889	.897	.923	.924	.865	.942	.740	.976	.800	.924	.823	.816	.849	.978	.946	.891	.877
Buffalo	.852	.957	.987	.964	.975	.929	.829	.789	.791	.759	.877	1.034	.715	.886	.914	1.010	1.089	.793
New York City	.956	.969	1.006	1.050	1.105	.985	1.061	1.063	1.074	1.001	1.022	1.021	1.036	.931	1.150	1.127	1.081	.987
New York-Connecticut suburbs	.908	1.032	1.075	1.020	1.040	1.025	.820	1.051	1.090	1.006	.968	.885	.861	.856	1.015	.924	.976	.966
New Jersey suburbs	.894	1.105	1.096	.992	1.056	.995	.900	1.063	1.134	.982	.950	.909	.834	.933	1.308	.957	.905	.866
Northeast, B PSUs	.811	1.017	1.017	.991	.967	.946	1.020	.930	.998	.897	.981	.924	.828	.893	1.033	.920	.960	.824
Northeast, C PSUs	.928	1.025	1.013	1.034	.952	.785	.990	.787	.897	.867	.907	.821	.780	.819	1.201	.962	1.088	.852
Northeast, D PSUs	.817	1.052	.976	.997	.850	1.000	.956	.863	.916	.789	1.056	.965	.847	.784	1.142	.959	.993	.944
Chicago	.847	.984	1.017	.968	1.011	.923	.905	.822	1.010	.958	1.043	1.059	.871	.874	1.035	.979	.948	.876
Detroit	.883	.852	.996	.995	.947	1.041	.965	.809	1.039	.881	.888	.806	.823	.843	.992	.959	1.045	.870
St. Louis	1.031	.988	.992	1.005	1.047	.894	.949	.845	1.143	.806	.999	1.009	.877	.798	.870	.981	1.026	.981
Cleveland	.742	.899	1.001	1.005	.918	.887	.859	.760	1.058	.904	.977	.797	.903	.857	.784	.971	1.051	.900
Minneapolis	1.045	.734	.987	1.063	.987	.914	1.056	.712	.975	.655	.923	.829	.810	.825	1.006	.964	.952	.880
Milwaukee	.834	.714	1.061	1.123	.872	.863	.815	.690	.993	.919	.979	.855	.770	.740	.610	.920	.899	.831
Cincinnati	.810	1.148	1.022	1.076	.926	1.137	1.060	.786	.983	.996	.965	.960	1.026	.964	1.298	.907	.887	.797
Kansas City	1.054	1.024	.967	.906	1.050	1.424	.838	.762	.983	.697	1.016	.911	.818	.990	1.167	1.050	.829	.868
North Central, B PSUs	.844	.841	.940	.966	.980	.889	.796	.742	.934	.869	.979	.842	.891	.794	1.093	1.006	1.002	.814
North Central, C PSUs	.827	1.000	.946	.940	.951	.852	.990	.805	1.030	.808	.930	.861	.836	.866	1.133	.938	.877	.815
North Central, D PSUs	.798	.896	.923	.878	.886	.775	.991	.755	.961	.782	.860	.750	.846	.858	1.009	.894	.989	.880
Washington, D.C.	.935	1.047	1.062	1.027	1.062	.922	1.138	.870	1.043	.994	1.055	1.033	.940	.934	1.033	1.093	.876	.976
Dallas	.890	.776	.942	.956	.954	.834	.977	.756	.995	.948	.916	.886	.888	.845	1.135	.955	.933	.927
Baltimore	.847	1.005	1.000	.964	.971	.849	1.030	.876	1.064	.986	.991	.861	.952	1.054	.930	.990	.890	.989

(continued)

Table 3E.2 (continued)

Area	EC1	EC2	EC3	EC4	EC5	EC6	EC7	EC8	EC9	EC10	EC11	EC12	EC13	EC14	EC15	EC16	EC17	EC18
Houston	.906	.906	1.018	1.078	1.101	1.140	.801	.867	1.354	.970	.830	.859	.840	.865	1.012	1.027	.943	.977
Atlanta	.983	1.017	.996	1.171	1.109	.855	.823	.729	1.011	1.007	.969	.853	.888	1.033	1.280	1.117	1.045	.822
Miami	.947	.735	1.000	.882	.924	.710	.940	.862	1.255	.957	.745	.798	.807	.665	1.571	.875	.941	.928
Tampa	.913	1.002	.970	.882	.918	.782	.927	.767	1.075	.707	.858	.859	.790	.861	.881	.831	.830	.802
New Orleans	1.039	1.175	.973	.995	1.004	.859	.875	.750	1.263	.822	.989	.824	.924	.886	.729	1.014	.883	.945
South, B PSUs	.845	.844	1.035	1.020	.973	.851	.845	.789	1.125	.850	.933	.872	.815	.832	.861	.886	.875	.855
South, C PSUs	.898	.810	.970	.988	.913	.863	.789	.775	1.148	.836	.884	.845	.854	.830	.985	.909	.913	.855
South, D PSUs	.965	.887	1.029	1.009	.960	.832	.815	.850	1.089	.848	.930	.907	.897	.886	.879	.971	.952	.916
Los Angeles County	.931	.847	1.031	.995	.962	1.065	1.073	1.506	1.036	.939	.911	.755	.898	.878	1.049	1.050	.998	.936
Greater Los Angeles	.938	1.041	.966	.965	.964	.948	.845	1.384	1.003	.930	.899	.738	.886	.882	1.243	.999	.939	.877
San Francisco	.995	1.052	1.071	1.097	1.092	1.086	1.036	1.281	.930	1.027	.939	.774	.904	1.075	1.290	1.015	1.068	.916
Seattle	1.065	.984	.927	.904	1.000	.905	1.075	.885	1.054	.956	1.169	1.064	.940	.930	.804	1.079	1.343	1.034
San Diego	.826	1.137	1.052	1.058	.954	1.368	.802	1.814	1.092	.979	.865	.643	1.144	.844	1.619	.915	.772	.859
Portland, Oreg.	.945	.761	1.011	.960	1.016	1.012	.904	.845	.852	.969	.964	.691	.881	.899	1.363	.995	1.144	.992
Honolulu	1.333	1.795	1.342	1.172	1.305	1.410	1.349	1.398	1.763	1.179	1.197	1.390	1.218	1.219	1.020	1.454	1.054	1.360
Anchorage	1.091	1.042	1.031	1.114	1.172	1.586	.833	1.374	1.439	1.007	1.178	1.322	1.008	1.003	1.374	1.087	1.201	.947
Denver	.922	.987	1.099	.976	.702	.918	.986	.759	1.056	.880	1.048	1.016	.929	.879	1.330	.940	.821	1.031
West, B PSUs	.853	1.104	.916	.944	.904	.925	.888	1.052	.939	.878	.975	.713	1.023	.945	.755	.950	.843	.862
West, C PSUs	.973	.815	.962	.963	.932	.977	.853	.914	.995	.897	.951	.823	.935	.820	.978	.984	.976	.894
West, D PSUs	1.076	.931	.961	.991	.987	.855	1.011	.846	1.121	.943	.893	.729	.878	.990	.772	1.004	.863	.860

Table 3E.3 Index Values by Major Group and All Food-at-Home, Reference Area Philadelphia

Area	CERBAK	MPFE	DAIRY	FRTVEG	OTHER	FOOD
Philadelphia	1.000	1.000	1.000	1.000	1.000	1.000
Boston	.969	1.007	.888	.901	.789	.888
Pittsburgh	.881	.890	.877	.856	.897	.872
Buffalo	.924	.947	.770	.874	.842	.863
New York City	.965	1.018	1.032	1.010	1.025	1.007
New York-Connecticut suburbs	.997	1.022	1.043	.901	.970	.941
New Jersey suburbs	1.044	1.024	1.049	.907	.917	.946
Northeast, B PSUs	.955	.983	.939	.913	.863	.915
Northeast, C PSUs	.994	.921	.879	.843	.892	.877
Northeast, D PSUs	.976	.975	.844	.931	.970	.933
Chicago	.941	.965	.977	.978	.907	.962
Detroit	.860	1.001	.950	.841	.901	.865
St. Louis	1.000	.960	.954	.931	.973	.951
Cleveland	.849	.944	.971	.885	.902	.891
Minneapolis	.798	.970	.788	.853	.907	.845
Milwaukee	.747	.963	.951	.855	.809	.843
Cincinnati	1.040	1.053	.993	.977	.858	.966
Kansas City	1.032	1.068	.819	.934	.912	.927
North Central, B PSUs	.842	.919	.896	.886	.865	.878
North Central, C PSUs	.943	.910	.904	.878	.860	.889
North Central, D PSUs	.864	.858	.860	.823	.902	.846
Washington, D.C.	1.011	1.011	1.018	1.002	.991	1.004
Dallas	.805	.904	.966	.890	.954	.897
Baltimore	.952	.940	1.022	.954	.983	.964
Houston	.904	1.051	1.138	.848	.987	.907
Atlanta	1.006	.959	1.002	.924	.896	.941
Miami	.790	.861	1.088	.762	.975	.830
Tampa	.973	.874	.868	.846	.816	.862
New Orleans	1.137	.929	1.010	.909	.927	.957
South, B PSUs	.844	.952	.970	.871	.862	.878
South, C PSUs	.834	.918	.973	.858	.878	.871
South, D PSUs	.910	.938	.952	.910	.923	.918
Los Angeles County	.871	1.045	.981	.855	.964	.888
Greater Los Angeles	1.009	.960	.962	.844	.926	.894
San Francisco	1.034	1.086	.973	.898	.966	.934
Seattle	1.012	.929	.999	1.042	1.020	1.036
San Diego	1.027	1.135	1.027	.832	.920	.891
Portland, Oreg.	.816	.991	.907	.848	1.036	.874
Honolulu	1.668	1.334	1.416	1.264	1.321	1.344
Anchorage	1.055	1.207	1.192	1.144	1.008	1.112
Denver	.967	.964	.957	.982	1.048	.990
West, B PSUs	.957	.925	.903	.910	.859	.909
West, C PSUs	.863	.954	.939	.885	.917	.894
West, D PSUs	.978	.927	1.021	.856	.863	.894

References

Aczel, J. 1966. *Lectures on functional equations and their applications.* New York: Academic.

Caves, W., L. Christensen, and W. Diewert. 1982a. The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica* 50:1393–1414.

———. 1982b. Multilateral comparisons of output, input, and productivity using superlative index numbers. *Economic Journal* 92:73–86.

Cuthbert, J., and M. Cuthbert. 1988. On aggregation methods of purchasing power parities. Working Paper no. 56. Organization for Economic Cooperation and Development, Department of Economics and Statistics.

Diewert, W. 1976. Exact and superlative index numbers. *Journal of Econometrics* 4: 114–45.

———. 1986. Micro economic approaches to the theory of international comparisons. Technical Paper no. 53. Cambridge, Mass.: National Bureau of Economic Research.

———. 1987. Index numbers. In *The new Palgrave dictionary of economics,* ed. J. Eatwell, M. Milgate, and P. Newman. New York: Stockton.

Drechsler, L. 1973. Weighting of index numbers in multilateral international comparisons. *Review of Income and Wealth,* ser. 19, no. 1:17–34.

Eichhorn, W. 1978. *Functional equations in economics.* Reading, Mass.: Addison-Wesley.

Fixler, D., and K. Zieschang. 1992. Incorporating ancillary measures of process and quality change into a superlative productivity index. *Journal of Productivity Analysis* 2, no. 2:245–67.

Kokoski, Mary, Pat Cardiff, and Brent Moulton. 1994. Interarea price indices for consumer goods and services: An hedonic approach using CPI data. Working Paper no. 256. Washington, D.C.: Bureau of Labor Statistics, July.

Moulton, B. R. 1995. Interarea indexes of the cost of shelter using hedonic quality adjustment techniques. *Journal of Econometrics* 68, no. 1:181–204.

Randolph, W., and K. Zieschang. 1987. Aggregation consistent restriction based improvement of local area estimators. In *American Statistical Association: Proceedings of the Section on Business and Economic Statistics.* Alexandria, Va.: American Statistical Association.

Selvanathan, E., and D. Prasada Rao. 1992. An econometric approach to the construction of generalized Theil-Törnqvist indices for multilateral comparisons. *Journal of Econometrics* 54:335–46.

Summers, R. 1973. International comparisons with incomplete data. *Review of Income and Wealth,* ser. 19, no. 1 (March): 1–16.

U.S. Department of Labor. Bureau of Labor Statistics. 1992. *BLS Handbook of Methods.* Bulletin 2414. Washington, D.C., September.

Zieschang, K. 1985. Output price measurement when output characteristics are endogenous. Working Paper no. 150. Washington, D.C.: Bureau of Labor Statistics. Revised, 1987.

———. 1988. The characteristics approach to the problem of new and disappearing goods in price indexes. Working Paper no. 183. Washington, D.C.: Bureau of Labor Statistics.

Comment Paul Pieper

A major gap in the U.S. system of economic statistics is the absence of any regional or area price indexes. While the Bureau of Labor Statistics publishes consumer price indexes for major U.S. cities, these indexes do not permit comparisons between cities, only comparisons over time at the same city. Area price indexes are likely to be in great demand by both academics and the general public. This paper makes several contributions, including incorporating hedonic quality adjustments into Törnqvist bilateral indexes, developing a multilateral Törnqvist index that minimally adjusts the Törnqvist system of bilateral index comparisons and aggregating the multilateral Törnqvist index so that it is transitive in both the index subcomponents and the index total.

I agree with the authors' focus on a multilateral index. A system of bilateral Törnqvist indexes will be unwieldy if there are more than a handful of areas. For example, if there are forty-four different areas, as in the authors' empirical work, there will be 946 (44 × 43/2) possible bilateral price comparisons. This would entail significantly more reporting expenses than a single index with forty-four entries. In addition, bilateral Törnqvist indexes are not necessarily transitive. Finally, while a bilateral price index is appropriate for some purposes, such as comparing the real wage of job offers in two different cities, in many cases it is necessary to make multilateral comparisons. This is especially true in academic work, where regional or area price indexes are most likely to be used to deflate cross-sectional data.

Given that a multilateral index is preferable to a bilateral index, how should it be constructed? The authors propose a multilateral Törnqvist index constructed using reference price and share vectors estimated from a regression of the cross-weighted differences of log area prices against the area differences of log prices and the area share differences. The authors argue that this index will minimally perturb the system of bilateral Törnqvist index comparisons, but, unfortunately, it is difficult to see from the paper why this is the case. Their point can be more easily seen if one starts by minimizing the squared deviations (weighted if desired) between the bilateral Törnqvist index (eq. [9]) and the multilateral Törnqvist index (eq. [11]). Expanding the two equations and simplifying yields the regression used in the paper.

Minimizing the squared deviations from the bilateral Törnqvist index comparisons is on the surface a reasonable objective for a multilateral index, but at what cost is it achieved? The paper is largely silent on this issue except for a passing reference in the concluding section. An evaluation of the merits of the authors' proposed index requires a discussion of the advantages and disadvantages of their index in relation to the Caves, Christensen, and Diewert (CCD) method or some other competing alternative. The authors' index has the disadvantage of significantly greater computational cost relative to the CCD index

Paul Pieper is associate professor of economics at the University of Illinois at Chicago.

since a regression is necessary for each category in each level of aggregation to determine reference shares and prices. Their index by construction has the advantage of minimizing squared deviations from the bilateral Törnqvist index, but the importance of this effect is unclear. Tables 3D.3 and 3D.6 of the paper provide a comparison of the multilateral Törnqvist index with the bilateral Törnqvist index, but, for this to be put in perspective, a similar comparison needs to be made with other indexes.

Table 3F.1 amends the authors' bilateral price comparisons to include two other index types: an unweighted CCD index and a CCD index where reference shares are weighted by area shares.[1] The three areas are Philadelphia (Phi), Boston (Bos), and Pittsburgh (Pit). The differences between the authors' index (KMZ) and the CCD indexes are trivial. The largest difference between the authors' index and the unweighted CCD index in the six comparisons is only 0.00057. The unweighted CCD index and the KMZ index are identical when rounded to three decimal places, which is the likely number of decimal places to be reported by the BLS. This example is not conclusive since it includes only three areas and two product categories, but it raises doubts as to whether the reduction in deviations from the bilateral Törnqvist index is worth the additional computational cost of the authors' proposed index.

A major part of both the theoretical and the empirical sections of the paper concerns the use of hedonic price indexes to adjust for quality differences in products across areas. However, with the major exception of housing, most products in the United States are homogeneous across areas. Hedonic quality adjustments are necessary in the authors' data set not because the goods themselves differ across regions but because the product definitions are imprecise. Thus, it would be possible to control for quality differences directly, without the use of hedonic regressions, by narrowly defining the good, for example, Red Delicious apples, unpackaged and sold at a chain grocery store, versus just apples. The advantages of a narrow definition are a better control for quality differences and lower computational costs, whereas the hedonic method has the advantage of allowing a greater coverage of products. It is probably not possible to construct area price indexes using narrow product definitions with the authors' data set because there are likely to be only a few observations in a given area at the narrowest level of product definition. However, it is unclear whether the authors' preference for hedonic quality adjustments is based on principle or is necessitated by their data. Some discussion of this issue in the paper would be useful.

Area price indexes must also deal with thorny issues of area differences in nonmarket consumption owing to weather or other factors. Do the extra expenses for warm clothing and heating in Northern cities represent an increase in consumption or simply an increase in costs? One could argue that these

1. The weights should be based on the area's share in the next-higher-level aggregate, i.e., fruits or vegetables. In the absence of this information, I used the food-at-home shares of each area (table 3B.2) as weights.

Table 3F.1 Comparison of Törnqvist Indexes

	Index Type			
	Bilateral Törnqvist (1)	KMZ (2)	Unweighted CCD (3)	Weighted CCD (4)
Index level:				
Fruits				
Phi-Bos	.96622	.96136	.96086	.96275
Phi-Pit	.94033	.94553	.94558	.94525
Bos-Pit	.98959	.98353	.98410	.98179
Vegetables				
Phi-Bos	.91564	.91397	.91379	.91445
Phi-Pit	.83279	.83453	.83448	.83437
Bos-Pit	.91505	.91308	.91321	.91242
Deviation from bilateral Törnqvist index:				
Fruits				
Phi-Bos	.0	−.00486	−.00536	−.00347
Phi-Pit	.0	.00520	.00525	.00492
Bos-Pit	.0	−.00606	−.00549	−.00780
Mean absolute deviation	.0	.00537	.00537	.00540
Vegetables				
Phi-Bos	.0	−.00167	−.00185	−.00119
Phi-Pit	.0	−.00174	.00169	.00158
Bos-Pit	.0	−.00197	−.00184	−.00263
Mean absolute deviation	.0	.00179	.00179	.00180

Note: Abbreviations are defined in the text.

same heating services are provided Southern residents for free by nature and that their price vector should show a zero price. A similar argument could be made for the extra expenses incurred for security in a crime-ridden city as opposed to a crime-free city. These issues are unlikely to be confronted directly by a government-produced index. The practice of using only market consumption and prices implicitly means that any "necessary" consumption purchases for heating, security, or the like will be measured as an increase in quantity rather than an increase in price. However, it is possible that differences in weather-related consumption items may in practice account for the bulk of actual area price differences.

While the authors' work on improving the construction of an area price index is laudable, I hope that the search for index perfection does not delay the introduction of area price indexes by the statistical agencies. Since the U.S. government does not presently produce any area price indexes, academics must use either undeflated data or crude proxies such as median housing prices or average wages. Seen in this light, virtually any well-defined method of indexation, whether based on the authors' minimum-adjustment criteria, the CCD method, or even a fixed basket of goods, would be a major improvement.