Quality Adjustment and Hedonics: An Attempt at a Unified Approach

Erwin Diewert,[1]
Discussion Paper 18-10,
Vancouver School of Economics,
University of British Columbia,
Vancouver, B.C., Canada, V6T 1L4.

November 28, 2018.

Abstract

The paper takes a consumer demand perspective to the problem of adjusting product prices for quality change. The various approaches to the problem of quality adjustment can be seen as special cases of the general framework. The special cases include the use of inflation adjusted carry forward and carry backward prices, the use of hedonic regressions and the estimation of Hicksian reservation prices.

Keywords

Quality adjustment, hedonic regressions, reservation prices, consumer theory, time product dummy regressions, scanner data.

JEL Classification Numbers

C43, C81, E31.

[1] W. Erwin Diewert: Vancouver School of Economics, University of British Columbia, Vancouver B.C., Canada, V6T 1Z1 and the School of Economics, UNSW Sydney, NSW 2052, Australia (erwin.diewert@ubc.ca). The author thanks Jan de Haan and Jerry Hausman for helpful comments and gratefully acknowledges the financial support of a SSHRC grant.

1. Introduction

Note that the paper is incomplete!

2. A Framework for Evaluating Quality Change In the Scanner Data Context

In this section, we attempt to provide a framework for the construction of consumer price and quantity indexes in the scanner data context using the economic approach to index number theory. We assume that transactions data for the sales or purchases of N products over T time periods are available.[2] The N products will typically be a group of related products so that the goal is the construction of price and quantity indexes at the first stage of aggregation. The transactions data are aggregated over time within each period. Let $p^t \equiv [p_{t1},...,p_{tN}]$ and $q^t \equiv [q_{t1},...,q_{tN}]$ denote the price and quantity vectors for time periods t = 1,...,T. The period t quantity for product n, q_{tn}, is equal to total purchases of product n by purchasers or to the sales of product n by the outlet (or group of outlets) for period t, while the period t price for product n, p_{tn}, is equal to the value of sales (or purchases) of product n in period t, v_{tn}, divided by the corresponding total quantity sold (or purchased), q_{tn}. Thus $p_{tn} \equiv v_{tn}/q_{tn}$ is the *unit value price* for product n in period t for t = 1,...,T and n = 1,...,N. In this section, we assume that all prices, quantities and values are positive; in subsequent sections, this assumption will be relaxed.

Let $q \equiv [q_1,...,q_N]$ be a generic quantity vector. In order to compare various methods for comparing the value of alternative combinations of the N products, it is necessary that a *valuation function* or *aggregator function*, Q(q), exist. This function allows us to value alternative combinations of products; if $Q(q^2) > Q(q^1)$, then purchasers of the products place a higher utility value on the vector of purchases q^2 than they place on the vector of purchases q^1. The function Q(q) can also act as an *aggregate quantity level* for the vector of purchases, q. Thus $Q(q^t)$ can be interpreted as an aggregate quantity level for the period t vector of purchases, q^t, and the ratios, $Q(q^t)/Q(q^1)$, t = 1,..,T, can be interpreted as *fixed base quantity indexes* covering periods 1 to T.

In the following analysis, we assume that Q(q) has the following properties: (i) Q(q) > 0 if $q >> 0_N$;[3] (ii) Q(q) is nondecreasing in its components; (iii) $Q(\lambda q) = \lambda Q(q)$ for $q \geq 0_N$ and $\lambda \geq 0$; (iv) Q(q) is a continuous concave function over the nonnegative orthant. Assumption (iii), linear homogeneity of Q(q), is a somewhat restrictive assumption. However, this assumption is required to ensure that the aggregate price level, P(p,q), that

[2] The data could be price and quantity (or value and quantity) on sales of the N products from a retail outlet (or group of outlets in the same region) or it could be price and quantity data for the purchases of the N products by a group of similar households.

[3] Notation: $q >> 0_N$ means each component of q is positive, $q \geq 0_N$ means each component of q is nonnegative and $q > 0_N$ means $q \geq 0_N$ but $q \neq 0_N$,

corresponds to $Q(q)$ does not depend on the scale of q.[4] Property (iv) will ensure that the first order necessary conditions for the budget constrained maximization of $Q(q)$ are also sufficient.

Let $p \equiv [p_1,...,p_N] > 0_N$ and $q \equiv [q_1,...,q_N] > 0_N$ with $p \cdot q \equiv \Sigma_{n=1}^N p_n q_n > 0$. Then the *aggregate price level*, $P(p,q)$ that corresponds to the aggregate quantity level $Q(q)$ is defined as follows:

$$(1)\ P(p,q) \equiv p \cdot q/Q(q).$$

Thus the implicit price level that is generated by the generic price and quantity vectors, p and q, is equal to the value of purchases, $p \cdot q$, deflated by the aggregate quantity level, $Q(q)$. Note that using these definitions, the product of the aggregate price and quantity levels equals the value of purchases during the period, $p \cdot q$.

Once the functional form for the aggregator function $Q(q)$ is known, then the *aggregate quantity level for period t*, Q^t, can be calculated in the obvious manner:

$$(2)\ Q^t \equiv Q(q^t); \qquad t = 1,...,T.$$

Using definition (1), the corresponding period t aggregate price level, P^t, can be calculated as follows:

$$(3)\ P^t \equiv p^t \cdot q^t/Q(q^t); \qquad t = 1,...,T.$$

Note that if $Q(q)$ turns out to be a linear aggregator function, so that $Q(q^t) \equiv \alpha \cdot q^t = \Sigma_{n=1}^N \alpha_n q_{tn}$, then the corresponding period t price level P^t is equal to $p^t \cdot q^t/\alpha \cdot q^t$, which is a *quality adjusted unit value price level*.

In order to make further progress, it is necessary to make some additional assumptions. The two additional assumptions are: (v) $Q(q)$ is once differentiable with respect to the components of q and (vi) the observed strictly positive quantity vector for period t, $q^t >> 0_N$, is a solution to the following period t constrained maximization problem defined by: [5]

$$(4)\ \max_q \{Q(q) : p^t \cdot q = v^t ; q \geq 0_N\}; \qquad t = 1,...,T.$$

The first order conditions for solving (4) for period t are the following conditions:[6]

---

[4] $P(p,q) \equiv p \cdot q/Q(q)$ where $p \cdot q \equiv \Sigma_{n=1}^N p_n q_n$. Thus using property (iii) of $Q(q)$, we have $P(p,\lambda q) = p \cdot \lambda q/Q(\lambda q) = P(p,q)$.

[5] The theory that follows dates back to Konüs and Byushgens (1926). This approach blends standard consumer demand theory based on the maximization of a linearly homogeneous utility function with index number theory. It was further developed by Shephard (1953) (in the context of a cost minimization framework), Samuelson and Swamy (1974) and Diewert (1976). What is new in the present paper is the application of this theory to hedonic regression models.

[6] Using the assumption of concavity of $Q(q)$ and the assumption that $q^t >> 0_N$, these conditions are also sufficient to solve (4). Notation: $\nabla_q Q(q) \equiv [\partial Q(q)/\partial q_1,..., [\partial Q(q)/\partial q_N]$.

$$(5)\ \nabla_q Q(q^t) = \lambda_t p^t\ ; \qquad t = 1,...,T;$$
$$(6)\quad p^t \cdot q^t = v^t\ ; \qquad t = 1,...,T.$$

Since Q(q) is assumed to be linearly homogeneous with respect to q, Euler's Theorem on homogeneous functions implies that the following equations hold:

$$(7)\ q^t \cdot \nabla_q Q(q^t) = Q(q^t)\ ; \qquad t = 1,...,T.$$

Take the inner product of both sides of equations (5) with q^t and use the resulting equations along with equations (7) to solve for the Lagrange multipliers, λ_t:

$$(8)\ \lambda_t = Q(q^t)/p^t \cdot q^t \qquad t = 1,...,T$$
$$= 1/P^t \qquad \text{using definitions (3).}$$

Thus if we assume utility maximizing behavior on the part of purchasers of the N products using the collective utility function Q(q) that satisfies the above regularity conditions, then the period t quantity aggregate is $Q^t \equiv Q(q^t)$ and the companion period t price level defined as $P^t \equiv p^t \cdot q^t/Q^t$ is equal to $1/\lambda_t$ where λ_t is the Lagrange multiplier for problem t in the constrained utility maximization problems (4) and where q^t and λ_t solve equations (5) and (6) for period t. Equations (8) also imply that the product of P^t and Q^t is exactly equal to observed period t expenditure v_t; i.e., we have

$$(9)\ P^t Q^t = p^t \cdot q^t = v_t\ ; \qquad t = 1,...,T.$$

Substitute equations (8) into equations (5) and after a bit of rearrangement, the following *fundamental equations* are obtained:[7]

$$(10)\ p^t = P^t \nabla_q Q(q^t)\ ; \qquad t = 1,...,T.$$

In the following section, we will assume that the aggregator function, Q(q) is a linear function and we will show how this assumption along with equations (9) for the case where T = 2 and N = 3 can lead to a simple well known method for quality adjustment that does not involve any econometric estimation of the parameters of the linear function. In subsequent sections, equations (10) will be utilized in the hedonic regression context and finally, in the final sections of the paper, the assumption that Q(q) is a linear function will be relaxed.

3. A Nonstochastic Method for Quality Adjustment: A Simple Model

---

[7] Multiply the right hand side of equation t in (10) by $1 = Q^t/Q(q^t)$ and use $P^t Q^t = v_t$ to obtain the following system of equations: $p^t = v_t \nabla_q Q(q^t)/Q(q^t)$ for t = 1,...,T. For each t, this system of equations is the consumer's system of *inverse demand functions*, that give the period t prices that rationalize the observed period t demands q^t as functions of q^t and period t expenditure v_t. Konüs and Byushgens (1926) obtained a system of equations that are equivalent to this system of inverse demand functions. Linear homogeneity of the utility function is required to obtain these equations and the equivalent equations defined by (9) and (10).

A major problem that arises when statistical agencies use scanner data to construct an elementary index is that some products are sold or purchased in one period but not in a subsequent period. Conversely, new products appear in the present period which were not present in previous periods. How should price and quantity indexes be constructed under these circumstances? Equations (10) in the previous section can be used to provide an answer to this question.

Consider the special case where the number of periods T is equal to 2 and the number of products in scope for the elementary index is N equal to 3. Product 1 is present in both periods, product 2 is present in period 1 but not in period 2 (a disappearing product) and product 3 is not present in period 1 but is present in period 3 (a new product).[8] We assume that purchasers of the three products behave as if they collectively maximized the following linear aggregator function:

$$(11)\ Q(q_1,q_2,q_3) \equiv \alpha_1 q_1 + \alpha_2 q_2 + \alpha_3 q_3$$

where the α_n are positive constants. Under these assumptions, equations (10) written out in scalar form become the following equations:[9]

$$(12)\ p_{tn} = P^t \alpha_n\ ; \qquad\qquad n = 1,2,3;\ t = 1,2.$$

Equations (12) are 6 equations in the 5 parameters P^1 and P^2 (which can be interpreted as *aggregate price levels* for periods 1 and 2) and α_1, α_2 and α_3, which can be interpreted as *quality adjustment factors* for the 3 products; i.e., each α_n measures the relative usefulness of an additional unit of product n to purchasers of the 3 products. However, product 3 is not observed in the marketplace during period 1 and product 2 is not observed in the marketplace in period 2 and so there are only 4 equations in (12) to determine 5 parameters. However, the P^t and the α_n cannot all be identified using observable data; i.e., if P^1, P^2, α_1, α_2 and α_3 satisfy equations (12) and λ is any positive number, then λP^1, λP^2, $\lambda^{-1}\alpha_1$, $\lambda^{-1}\alpha_2$ and $\lambda^{-1}\alpha_3$ will also satisfy equations (12). Thus it is necessary to place a normalization (like $P^1 = 1$ or $\alpha_1 = 1$) on the 5 parameters which appear in equations (12) in order to obtain a unique solution. In the index number context, it is natural to set the price level for period 1 equal to unity and so we impose the following normalization on the 5 unknown parameters which appear in equations (12):

$$(13)\ P^1 = 1.$$

---

[8] The "new" product may not be a truly new product; it may be the case that product 3 was temporarily not available in period 1. Similarly, product 2 may not permanently disappear in period 2; it may reappear in a subsequent period.

[9] This is a special case of the Time Product Dummy regression model which will be studied in more detail in subsequent sections. Thus equations (12), which are the inverse consumer demand functions that result from the maximization of a linear utility function, lead directly to a particular hedonic regression model. It is this result which allows us to claim that our present approach is a way of reconciling hedonic regression models with classical consumer demand theory.

The 4 equations in (12) which involve observed prices and the single equation (13) are 5 equations in 5 unknowns. The unique solution to these equations is:

(14) $P^1 = 1$; $P^2 = p_{21}/p_{11}$; $\alpha_1 = p_{11}$; $\alpha_2 = p_{12}$; $\alpha_3 = p_{23}/(p_{21}/p_{11}) = p_{23}/P^2$.

Note that the resulting *price index*, P^2/P^1, is equal to p_{21}/p_{11}, the price ratio for the commodity that is present in both periods. Thus the price index for this very simple model turns out to be a *maximum overlap price index.*[10]

Once the P^t and α_n have been determined, equations (12) for the missing products can be used to define the following *imputed prices* p_{tn}^* for commodity 3 in period 1 and product 2 in period 2:

(15) $p_{13}^* \equiv P^1\alpha_3 = p_{23}/(P^2/P^1)$; $p_{22}^* \equiv P^2\alpha_2 = (p_{21}/p_{11})p_{12} = (P^2/P^1)p_{12}$.

These imputed prices can be interpreted as Hicksian (1940; 12) *reservation prices*;[11] i.e., they are the lowest possible prices that would deter purchasers from purchasing the products during periods if the unavailable products hypothetically became available.[12]

Note that $p_{13}^* = p_{23}/(P^2/P^1)$ is an *inflation adjusted carry backward price*; i.e., the observed price for product 3 in period 2, p_{23}, is divided by the maximum overlap price index P^2/P^1 in order to obtain a "reasonable" valuation for a unit of product 3 in period 1. Similarly, $p_{22}^* = (P^2/P^1)p_{12}$ is an *inflation adjusted carry forward price* for product 2 in period 2; i.e., the observed price for product 2 in period 1, p_{12}, is multiplied by the maximum overlap price index P^2/P^1 in order to obtain a "reasonable" valuation for a unit of product 2 in period 2.[13]

Note that the above algebra can be implemented without a knowledge of quantities sold or purchased. Assuming that quantity information is available, we now consider how companion quantity levels, Q^1 and Q^2, for the price levels, P^1 and P^2, can be determined. Note that $q_{13} = 0$ and $q_{22} = 0$ since consumers cannot purchase products that are not available. Use the imputed prices defined by (15) to obtain complete price vectors for each period; i.e., define the period 1 complete price vector by $p^1 \equiv [p_{11}, p_{12}, p_{13}^*]$ and the complete period 2 price vector by $p^2 \equiv [p_{21}, p_{22}^*, p_{23}]$. The corresponding complete quantity vectors are by $q^1 \equiv [q_{11}, q_{12}, 0]$ and $q^2 \equiv [q_{21}, 0, q_{23}]$. The period t aggregate quantity level Q^t can be calculated directly using only information on q^t and the vector of

---

[10] Keynes (1930; 94) was an early author that advocated this method for dealing with new goods by restricting attention to the goods that were present in both periods being compared. He called his suggested method the *highest common factor method*. Marshall (1887; 373) implicitly endorsed this method. Triplett (2004; 18) called it the *overlapping link method*.

[11] Hicks (1940 dealt only with the case of new goods; Hofsten (1952; 95-97) extended his approach to cover the case of disappearing goods as well.

[12] Strictly speaking, it would be necessary to add a tiny amount to these prices to deter consumers from purchasing these products if they were made available.

[13] The use of carry forward and backward prices to estimate missing prices is widespread in statistical agencies. For additional materials on this method for estimating missing prices, see Triplett (2004), de Haan and Krsinich (2012) and Diewert, Fox and Schreyer (2017).

quality adjustment factors, $\alpha \equiv [\alpha_1, \alpha_2, \alpha_3]$, or indirectly by deflating period t expenditure $v_t \equiv p^t \cdot q^t$ by the estimated period t price level, P^t. Thus we have the following two possible methods for constructing the Q^t:

$$(16)\ Q^t \equiv \alpha \cdot q^t \ ; \text{ or } Q^t \equiv p^t \cdot q^t / P^t \ ; \qquad t = 1,2.$$

However, using the complete price vectors p^t with imputed prices filling in for the missing prices, equations (12) hold exactly and thus it is straightforward to show that $Q^t = \alpha \cdot q^t = p^t \cdot q^t / P^t$ for t = 1,2. Thus it does not matter whether we use the direct or indirect method for calculating the quantity levels; both methods give the same answer in this simple model.[14]

A problem with this simple model is that there is only one product that is present in both periods. In the following section, we generalize the present model to allow for multiple overlapping products.

4. A Nonstochastic Method for Quality Adjustment: A More Complex Model

In order to generalize the very simple model for dealing with new and disappearing products that was presented in the previous section, it is first necessary to develop another application of the fundamental equations (10) in section 2.

Define the aggregator function Q(q) as follows:

$$(17)\ Q_{KBF}(q^*) \equiv [q^* \cdot Aq^*]^{1/2}$$

where q^* is defined as the N dimensional quantity vector $[q_1^*,...,q_N^*]$ and $A \equiv [a_{ij}]$ is an N by N symmetric matrix of parameters which satisfies certain regularity conditions.[15] Suppose further that that the observed price and quantity vectors for periods 1 and 2 are the positive price and quantity vectors, $p^{t*} \equiv [p_{t1}^*,...,p_{tN}^*]$ and $q^{t*} \equiv [q_{t1}^*,...,q_{tN}^*]$ for t = 1,2. We assume that q^{t*} solves $\max_q \{Q(q) : p^{t*} \cdot q = v^{t*} \ ; q \geq 0_N\}$ for t = 1,2 where $v^{t*} \equiv p^{t*} \cdot q^{t*}$ is observed expenditure on the N products for periods t = 1,2. The inverse demand functions (10) that correspond to this particular aggregator function are the following ones:

$$(18)\ p^{t*} = P^{t*}\, \nabla_q Q_{KBF}(q^{t*}) = P^t\, [q^{t*} \cdot Aq^{t*}]^{-1/2}\, Aq^{t*} \ ; \qquad t = 1,2.$$

Using the framework described in section 2 above, the period t aggregate quantity level for the present model is $Q^{t*} \equiv [q^{t*} \cdot Aq^{t*}]^{1/2}$ and the corresponding period t price level is P^{t*}

[14] In subsequent sections when we no longer assume that equations (12) hold exactly, then the direct and indirect methods for calculating the Q^t can differ.

[15] A is assumed to have one positive eigenvalue with a corresponding strictly positive eigenvector and N−1 negative or zero eigenvalues. This functional form was introduced into the economics literature by Konüs and Byushgens (1926) who showed its connection to the Fisher (1922) ideal index. This explains why $Q(q^*)$ is labeled as $Q_{KBF}(q^*)$. For further discussion of the regularity conditions on $Q_{KBF}(q^*)$, see Diewert (1976) and Diewert and Hill (2010).

$\equiv p^{t*}\cdot q^{t*}/Q^{t*}$ for t = 1,2. The *Fisher* (1922) *ideal quantity index* is a function of the observable price and quantity data and is defined as follows:;

(19) $Q_F(p^{1*},p^{2*},q^{1*},q^{2*}) \equiv [p^{1*}\cdot q^{2*}\ p^{2*}\cdot q^{2*}/p^{1*}\cdot q^{1*}\ p^{2*}\cdot q^{1*}]^{1/2}$.

Use equations (18) to eliminate p^{1*} and p^{2*} from the right hand side of (19). We find that

(20) $(p^{1*}\cdot q^{2*}\ p^{2*}\cdot q^{2*})/(p^{1*}\cdot q^{1*}\ p^{2*}\cdot q^{1*}) = q^{2*}\cdot Aq^{2*}/q^{1*}\cdot Aq^{1*}$.

Take positive square roots on both sides of (20). Using definitions (17) and (19), the resulting equation is:

(21) $Q_{KBF}(q^{2*})/Q_{KBF}(q^{1*}) = Q_F(p^{1*},p^{2*},q^{1*},q^{2*})$.

Thus $Q^{2*}/Q^{1*} = Q_{KBF}(q^{2*})/Q_{KBF}(q^{1*})$ is equal to the Fisher quantity index $Q_F(p^{1*},p^{2*},q^{1*},q^{2*})$, which can be calculated using observable price and quantity data for the two periods. We know from section 2 that

(22) $P^{t*}Q^{t*} = p^{t*}\cdot q^{t*}$; t = 1,2.

Now make the normalization $P^{1*} = 1$. Using this normalization and equations (21) and (22), the aggregate price and quantity levels for the two periods can be defined in terms of observable data as follows:

(23) $P^{1*} \equiv 1$; $Q^{1*} \equiv p^{1*}\cdot q^{1*}$; $Q^{2*} \equiv Q^{1*}\ Q_F(p^{1*},p^{2*},q^{1*},q^{2*})$; $P^{2*} \equiv p^{1*}\cdot q^{1*}/Q^{2*}$.

The above results can be combined with the 3 product model that was described in the previous section: relabel the above aggregate data as a composite product 1 so that the new product 1 that corresponds to the first product in section 3 has prices and quantities defined as $p_{t1} \equiv P^{t*}$ and $q_{t1} \equiv Q^{t*}$ for t = 1,2. Products 2 and 3 are a disappearing product and a new product respectively as in section 3 above. The aggregate price levels for the two periods (which use all N+2 products) are P^1 and P^2 and the new α_n parameters are defined by the following counterparts to equations (14) above:

(24) $P^1 = 1$; $P^2 = P^{2*}/P^{1*} = P_F(p^{1*},p^{2*},q^{1*},q^{2*})$; $\alpha_1 = 1$; $\alpha_2 = p_{12}$; $\alpha_3 = p_{23}/(P^{2*}/P^{1*})$

where $P^{2*}/P^{1*} \equiv [v^{2*}/v^{1*}]/[Q^{2*}/Q^{1*}] \equiv P_F(p^{1*},p^{2*},q^{1*},q^{2*})$ is the Fisher (1922) ideal price index that compares the prices of the N products that are present in both periods, p^{1*}, p^{2*}, for the two periods under consideration. The imputed prices for the missing products, $p_{13}{}^*$ and $p_{22}{}^*$, are obtained by using equations (15) for our present model:

(25) $p_{13}{}^* \equiv p_{23}/P_F(p^{1*},p^{2*},q^{1*},q^{2*})$; $p_{22}{}^* \equiv P_F(p^{1*},p^{2*},q^{1*},q^{2*})p_{12}$.

Comparing (24) and (25) with the corresponding equations (14) and (15) for the 3 product model, it can be seen that the price ratio for product 1 that was present in both periods, p_{21}/p_{11}, is replaced by the Fisher index $P_F(p^{1*},p^{2*},q^{1*},q^{2*})$ which is now defined over the set of products that are present in both periods. The type of inflation adjusted carry backward price p_{13}^* and the inflation adjusted carry forward price p_{22}^* defined by (25) are widely used by statistical agencies to estimate missing prices but usually using Laspeyres or Paasche indexes in place of the Fisher price index.[16]

The aggregator function that is consistent with the new model with N continuing products, one disappearing product and one new product is defined as follows:

$$(26)\ Q(q_1^*,...,q_N^*,q_2,q_3) \equiv \alpha_1 Q_{KBF}(q^*) + \alpha_2 q_2 + \alpha_3 q_3$$

where $Q_{KBF}(q^*)$ is the KBF aggregator function defined by (17) and α_1 is set equal to 1.[17] Note that the model defined by (26) is restrictive from the economic perspective because the additive nature of definition (26) implies that the composite first commodity is perfectly substitutable (after quality adjustment) with the new and disappearing commodities (which are also perfect substitutes for each other after quality adjustment). However, if the products under consideration are highly substitutable for each other, the implicit assumption of perfect substitutes for missing products may be acceptable. Moreover, the advantage of this form of quality adjustment is that it is relatively easy to explain to the public and it is fairly straightforward to implement.

The restriction that there is only one new product and one disappearing product is readily relaxed. The overall price index will continue to be $P_F(p^{1*},p^{2*},q^{1*},q^{2*})$ and counterparts to equations (25) can be used to generate imputed prices for the missing products.

We turn now to applications of the basic framework explained in section 2 where conditions (10) only hold approximately rather than exactly.

5. Time Product Dummy Regressions: The Case of No Missing Observations

In this section, it is assumed that price and quantity data for N products are available for T periods. Let $p^t \equiv [p_{t1},...,p_{tN}]$ and $q^t \equiv [q_{t1},...,q_{tN}]$ denote the price and quantity vectors for time periods t = 1,...,T. Initially, it is assumed that there are no missing prices or quantities so that all NT prices and quantities are positive. We assume that the quantity aggregator function Q(q) is the following linear function:

$$(27)\ Q(q) = Q(q_1,q_2,...,q_N) \equiv \Sigma_{n=1}^N \alpha_n p_n = \alpha\cdot q$$

where the α_n are positive parameters, which can be interpreted as quality adjustment factors. Under the assumption of maximizing behavior on the part of purchasers of the N commodities, assumption (27) applied to equations (10) imply that the following NT equations should hold exactly:

---

[16] Note that the aggregate price index that is generated by this model is $P_F(p^{1*},p^{2*},q^{1*},q^{2*})$ which does not use the unmatched prices for the two periods.

[17] It is not necessary to use the KBF aggregator function in the above model; any aggregator function that has an exact index number associated with it will work. See Diewert (1976) for examples of exact index number formulae.

(28) $p_{tn} = \pi_t\alpha_n$; $\qquad$ n = 1,...,N; t = 1,...,T

where we have redefined the period t price levels P^t in equations (10) as the parameters π_t for t = 1,...,T.

Note that equations (28) form the basis for the *time dummy hedonic regression model* which is due to Court (1939).[18] Note that these equations are a special case of the model of consumer behavior that was explained in section 2 above.

At this point, it is necessary to point out that our consumer theory derivation of equations (28) is not accepted by all economists. Rosen (1974), Triplett (1987) and Pakes (2001)[19] have argued for a more general approach to the derivation of hedonic regression models that is based on supply conditions as well as on demand conditions. The present approach is obviously based on consumer demands and preferences only. This consumer oriented approach was endorsed by Griliches (1971; 14-15), Muellbauer (1974; 988) and Diewert (2003a) (2003b).[20] Of course, the separability assumptions which justify the present consumer approach are quite restrictive but nevertheless, it is useful to imbed hedonic regression models in a traditional consumer demand setting.

Empirically, equations (28) are unlikely to hold exactly. Thus we assume that the exact model defined by (28) holds only to some degree of approximation and so error terms, e_{tn}, are added to the right hand sides of equations (28). The unknown parameters, $\pi \equiv [\pi_1,...,\pi_T]$ and $\alpha \equiv [\alpha_1,...,\alpha_N]$, will be estimated as solutions to the following (nonlinear) least squares minimization problem:

(29) $\min_{\alpha,\pi} \Sigma_{n=1}^{N} \Sigma_{t=1}^{T} [p_{tn} - \pi_t\alpha_n]^2$.

[18] This was Court's (1939; 109-111) hedonic suggestion number two. He transformed the underlying equations (28) by taking logarithms of both sides of these equations (which will be done below). He chose to transform the prices by the log transformation because the resulting regression model fit his data on automobiles better. Diewert (2003b) also recommended the log transformation on the grounds that multiplicative errors were more plausible than additive errors.

[19] "The derivatives of a hedonic price function should not be interpreted as either willingness to pay derivatives or cost derivatives; rather they are formed from a complex equilibrium process." Ariel Pakes (2001; 14).

[20] Diewert (2003b; 97) justified the consumer demand approach as follows: "After all, the purpose of the hedonic exercise is to find how demanders (and not suppliers) of the product value alternative models in a given period. Thus for the present purpose, it is the preferences of consumers that should be decisive, and not the technology and market power of producers. The situation is similar to ordinary general equilibrium theory where an equilibrium price and quantity for each commodity is determined by the interaction of consumer preferences and producer's technology sets and market power. However, there is a big branch of applied econometrics that ignores this complex interaction and simply uses information on the prices that consumers face, the quantities that they demand and perhaps demographic information in order to estimate systems of consumer demand functions. Then these estimated demand functions are used to form estimates of consumer utility functions and these functions are often used in applied welfare economics. What producers are doing is entirely irrelevant to these exercises in applied econometrics with the exception of the prices that they are offering to sell at. In other words, we do not need information on producer marginal costs and markups in order to estimate consumer preferences: all we need are selling prices."

Our approach to the specification of the error terms will not be very precise. Throughout the paper, we will obtain estimators for the aggregate price levels π_t and the quality adjustment parameters α_n as solutions to least squares minimization problems like those defined by (29) or as solutions to weighted least squares minimization problems that will be considered in subsequent sections. Our focus will not be on the distributional aspects of our estimators; rather, our focus will be on the *axiomatic* or *test properties* of the price levels that are solutions to the various least squares minimization problems.[21] Basically, the approach taken here is a descriptive statistics approach: we consider simple models that aggregate price and quantity information for a given period over a set of specified commodities into scalar measures of aggregate price and quantity that summarize the detailed price and quantity information in a "sensible" way.[22]

The first order necessary (and sufficient) conditions for $\pi \equiv [\pi_1,...,\pi_T]$ and $\alpha \equiv [\alpha_1,...,\alpha_N]$ to solve the minimization problem defined by (29) are equivalent to the following N + T equations:

$$(30)\ \alpha_n = \Sigma_{t=1}^{T}\, \pi_t p_{tn}/\Sigma_{t=1}^{T}\, \pi_t^2 \qquad n = 1,...,N$$
$$= \Sigma_{t=1}^{T}\, \pi_t^2\, (p_{tn}/\pi_t)/\Sigma_{t=1}^{T}\, \pi_t^2\ ;$$
$$(31)\ \pi_t = \Sigma_{n=1}^{N}\, \alpha_n p_{tn}/\Sigma_{t=1}^{T}\, \alpha_n^2 \qquad t = 1,...,T$$
$$= \Sigma_{n=1}^{N}\, \alpha_n^2 (p_{tn}/\alpha_n)/\Sigma_{t=1}^{T}\, \alpha_n^2.$$

Solutions to the two sets of equations can readily be obtained by iterating between the two sets of equations. Thus set $\alpha^{(1)} = 1_N$ (a vector of ones of dimension N) in equations (31) and calculate the resulting $\pi^{(1)} = [\pi_1^{(1)},...,\pi_T^{(1)}]$. Then substitute $\pi^{(1)}$ into the right hand sides of equations (30) to calculate $\alpha^{(2)} \equiv [\alpha_1^{(2)},...,\alpha_T^{(2)}]$. And so on until convergence is achieved.

If $\pi^* \equiv [\pi_1^*,...,\pi_T^*]$ and $\alpha^* \equiv [\alpha_1^*,...,\alpha_N^*]$ is a solution to (30) and (31), then $\lambda\pi^*$ and $\lambda^{-1}\alpha^*$ is also a solution for any $\lambda > 0$. Thus to obtain a unique solution we impose the normalization $\pi_1^* = 1$. Then $1,\pi_2^*,...,\pi_T^*$ is the sequence of fixed base index numbers that is generated by the least squares minimization problem defined by (29).

If quantity data are available, then using the general methodology that was outlined in section 2, aggregate quantity levels for the t periods can be obtained as $Q^{t*} \equiv \alpha^*{\cdot}q^t = \Sigma_{n=1}^{N}\, \alpha_n^* q_{tn}$ for t = 1,...,T. Estimated aggregate price levels can be obtained directly from the solution to (29); i.e., set $P^{t*} = \pi_t^*$ for t = 1,...,T. Alternative price levels can be indirectly obtained as $P^{t**} \equiv p^t{\cdot}q^t/Q^{t*} = p^t{\cdot}q^t/\alpha^*{\cdot}q^t$ for t = 1,...,T. If the optimized objective function in (29) is 0 (so that all errors $e_{tn}^* \equiv p_{tn} - \pi_t^*\alpha_n^*$ equal 0 for t = 1,...,T and n = 1,...,N), then

[21] For rigorous econometric approaches to the stochastic approach to index number theory, see Rao and Hajargasht (2016) and Gorajek (2018). These papers consider many transformations of the fundamental hedonic equations (28) and many methods for constructing averages of prices.
[22] Our approach is broadly similar to Theil's (1967; 136-137) approach to index number theory.

P^{t*} will equal P^{t**} for all t. However, usually nonzero errors will occur and so a choice between the two sets of estimators must be made.[23]

From (30), it can be seen that α_n^*, the quality adjustment parameter for product n, is a weighted average of the T inflation adjusted prices for product n, the p_{tn}/π_t^*, where the weight for p_{tn}/π_t^* is $\pi_t^{*2}/\Sigma_{\tau=1}^T \pi_\tau^{*2}$. This means that the weight for p_{tn}/π_t^* will be very high for periods t where general inflation is high, which seems rather arbitrary. From (31), it can be seen that π_t^*, the period t price level (and fixed base price index), is weighted average of the N quality adjusted prices for period t, the p_{tn}/α_t^*, where the weight for p_{tn}/α_n^* is $\alpha_n^{*2}/\Sigma_{i=1}^N \alpha_i^{*2}$. It is a positive that π_t^* is a weighted average of the quality adjusted prices for period t but the quadratic nature of the weights is not an attractive feature.

In addition to having unattractive weighting properties, the estimates generated by solving the least squares minimization problem (29) suffer from a fatal flaw: *the estimates are not invariant to changes in the units of measurement*. In order to remedy this defect, we turn to an alternative error specification.

Instead of adding approximation errors to the exact equations (28), we could append multiplicative approximation errors. Thus the exact equations become $p_{tn} = \pi_t\alpha_n e_{tn}$ for n = 1,...,N and t = 1,...,T. Upon taking logarithms of both sides of these equations, we obtain the following system of estimating equations:

$$(32)\ \ln p_{tn} = \ln\pi_t + \ln\alpha_n + \ln e_{tn}\ ; \qquad n = 1,...,N;\ t = 1,...,T$$
$$= \rho_t + \beta_n + \varepsilon_{tn}$$

where $\rho_t \equiv \ln\pi_t$ for t = 1,...,T and $\beta_n \equiv \ln\alpha_n$ for n = 1,...,N. The model defined by (32) is the basic *Time Product Dummy regression model* with no missing observations.[24] Now choose the ρ_t and β_n to minimize the sum of squared residuals, $\Sigma_{n=1}^N \Sigma_{t=1}^T \varepsilon_{tn}^2$. Thus let $\rho \equiv [\rho_1,...,\rho_T]$ and $\beta \equiv [\beta_1,...,\beta_N]$ be a solution to the following least squares minimization problem:

$$(33)\ \min_{\rho,\beta} \Sigma_{n=1}^N \Sigma_{t=1}^T [\ln p_{tn} - \rho_t - \beta_n]^2 .$$

The first order necessary conditions for $\rho_1,...,\rho_T$ and $\beta_1,...,\beta_N$ to solve (33) are the following T + N equations:

$$(34)\ N\rho_t + \Sigma_{n=1}^N \beta_n = \Sigma_{n=1}^N \ln p_{tn}\ ; \qquad t = 1,...,T;$$
$$(35)\ \Sigma_{t=1}^T \rho_t + T\beta_n = \Sigma_{t=1}^T \ln p_{tn}\ ; \qquad n = 1,...,N.$$

[23] Usually, the direct estimates for the price levels will be used in hedonic regression studies; i.e., the $P^{t*} = \pi_t^*$ estimates will be used. For statistical agencies, an advantage of the direct estimates is that they can be calculated without the use of quantity information.

[24] A generalized version of this model (with missing observations) was proposed by Summers (1973) in the international comparison context where it is known as the Country Product Dummy regression model. A weighted version of this model (with missing observations) was proposed by Aizcorbe, Corrado and Doms (2000).

Replace the ρ_t and β_n in equations (34) and (35) by $\ln\pi_t$ and $\ln\alpha_n$ respectively for t = 1,...,T and n = 1,...,N. After some rearrangement, the resulting equations become:

(36) $\pi_t = \Pi_{n=1}^{N} (p_{tn}/\alpha_n)^{1/N}$; t = 1,...,T;
(37) $\alpha_n = \Pi_{t=1}^{T} (p_{tn}/\pi_t)^{1/T}$; n = 1,...,N.

Thus the period t aggregate price level, π_t, is equal to the geometric average of the N quality adjusted prices for period t, p_{t1}/α_1, ..., p_{tN}/α_N, while the quality adjustment factor for product n, α_n, is equal to the geometric average of the T inflation adjusted prices for product n, p_{1n}/π_1, ..., p_{Tn}/π_T. These estimators look very reasonable (if quantity weights are not available).

Solutions to (36) and (37) can readily be obtained by iterating between the two sets of equations. Thus set $\alpha^{(1)} = 1_N$ (a vector of ones of dimension N) in equations (36) and calculate the resulting $\pi^{(1)} = [\pi_1^{(1)},...,\pi_T^{(1)}]$. Then substitute $\pi^{(1)}$ into the right hand sides of equations (37) to calculate $\alpha^{(2)} \equiv [\alpha_1^{(2)},...,\alpha_T^{(2)}]$. And so on until convergence is achieved. Alternatively, equations (34) and (35) are linear in the unknown parameters and can be solved (after normalizing one parameter) by a simple matrix inversion. A final method of obtaining a solution to (34) and (35) is to apply a simple linear regression model to equations (32).[25]

If $\pi^* \equiv [\pi_1^*,...,\pi_T^*]$ and $\alpha^* \equiv [\alpha_1^*,...,\alpha_N^*]$ is a solution to (36) and (37), then $\lambda\pi^*$ and $\lambda^{-1}\alpha^*$ is also a solution for any $\lambda > 0$. Thus to obtain a unique solution we impose the normalization $\pi_1^* = 1$ (which corresponds to $\rho_1 = 0$). Then $1,\pi_2^*,...,\pi_T^*$ is the sequence of fixed base index numbers that is generated by the least squares minimization problem defined by (33).

Once we have the unique solution $1,\pi_2^*,...,\pi_T^*$ for the T price levels that are generated by the (33), the *price index* between period t relative to period s can be defined as π_t^*/π_s^*. Using equations (36) for π_t^* and π_s^*, we have the following expression for the price index:

$$(38)\quad \pi_t^*/\pi_s^* = \Pi_{n=1}^{N} (p_{tn}/\alpha_n^*)^{1/N} / \Pi_{n=1}^{N} (p_{sn}/\alpha_n^*)^{1/N}$$
$$= \Pi_{n=1}^{N} (p_{tn}/p_{sn})^{1/N}.$$

Thus if there are no missing observations, the Time Product Dummy price indexes between any two periods in the window of T period under consideration is equal to the *Jevons index* between the two periods (the simple geometric mean of the price ratios, p_{tn}/p_{sn}).[26] This is a somewhat disappointing result since an equally weighted average of the price ratios is not necessarily a representative average of the prices; i.e., unimportant products to purchasers (in the sense that they spend very little on these products) are

[25] Again we require one normalization on the parameters such as $\rho_1 = 0$.
[26] This result is a special case of a more general result obtained by Triplett and McDonald (1977; 150).

given the same weight in the Jevons measure of inflation between the two periods as is given to high expenditure products.[27]

If there are no missing observations, then it can be seen using equations (37) that the ratio of the quality adjustment factor for product n relative to product m is equal to the following sensible expression:

$$(39)\ \alpha_n^*/\alpha_m^* = \Pi_{t=1}^T (p_{tn}/\pi_t^*)^{1/T} / \Pi_{t=1}^T (p_{tm}/\pi_t^*)^{1/T}$$
$$= \Pi_{t=1}^T (p_{tn}/p_{tm})^{1/T}.$$

If quantity data are available, then aggregate quantity levels for the t periods can be obtained as $Q^{t*} \equiv \alpha^*\cdot q^t = \Sigma_{n=1}^N \alpha_n^* q_{tn}$ for t = 1,...,T. Estimated aggregate price levels can be obtained directly from the solution to (33); i.e., set $P^{t*} = \pi_t^*$ for t = 1,...,T. Alternative price levels can be obtained indirectly as $P^{t**} \equiv p^t\cdot q^t/Q^{t*} = p^t\cdot q^t/\alpha^*\cdot q^t$ for t = 1,...,T. If the optimized objective function in (33) is 0 (so that all errors $\varepsilon_{tn}^* \equiv \ln p_{tn} - \rho_t^* - \beta_n^*$ equal 0 for t = 1,...,T and n = 1,...,N), then P^{t*} will equal P^{t**} for all t. If the estimated residuals are not all equal to 0, then the two estimates for the period t price level P^t will differ. The two estimates for P^t will generate different estimates for the companion aggregate quantity levels.

Note that the underlying exact model ($p_{tn} = \pi_t\alpha_n$ for t and n) is the same for both least squares minimization problems, (29) and (33). However, different error specifications and different transformations of both sides of the equations $p_{tn} = \pi_t\alpha_n$ can lead to very different estimators for the π_t and α_n. Our strategy in this section and in the following sections will be to choose specifications of the least squares minimization problem that lead to estimators for the price levels π_t that have good axiomatic properties.[28] From this perspective, it is clear that (33) leads to “better” estimates than (29).

In the following section, we allow for missing observations.

6. Time Product Dummy Regressions: The Case of Missing Observations

In this section, the least squares minimization problem defined by (33) is generalized to allow for missing observations. In order to make this generalization, it is first necessary to make some definitions. As in the previous section, there are N products and T time periods but not all products are purchased (or sold) in all time periods. For each period t, define the set of products n that are present in period t as $S(t) \equiv \{n: p_{tn} > 0\}$ for t = 1,2,...,T. It is assumed that these sets are not empty; i.e., at least one product is purchased in each period. For each product n, define the set of periods t where product n is present as $S^*(n) \equiv \{t: p_{tn} > 0\}$. Again, assume that these sets are not empty; i.e., each product is sold in at least one time period. Define the integers N(t) and T(n) as follows:

---

[27] However, if quantity data are not available, the Jevons index has the strongest axiomatic properties compared to its competitors; see the ILO, Eurostat, IMF, OECD, UNECE and the World Bank (2004).

[28] From the perspective of the economic approach to index number theory that was outlined in section 2, problems (29) and (33) have the same economic justification.

(40) $N(t) \equiv \Sigma_{n\in S(t)}\, 1$; $t = 1,...,T$;
(41) $T(n) \equiv \Sigma_{t\in S^*(n)}\, 1$; $n = 1,...,N$.

If all N products are present in period t, then N(t) = N; if product n is present in all T periods, then T(n) = T.

Using the notation that was defined in the previous section, the counterpart to (33) when there are missing products is the following least squares minimization problem:

(42) $\min_{\rho,\beta} \Sigma_{t=1}^{T} \Sigma_{n\in S(t)} [\ln p_{tn} - \rho_t - \beta_n]^2 = \min_{\rho,\alpha} \Sigma_{n=1}^{N} \Sigma_{t\in S^*(n)} [\ln p_{tn} - \rho_t - \beta_n]^2$.

Note that there are two equivalent ways of writing the least squares minimization problem.[29] The first order necessary conditions for $\rho_1,...,\rho_T$ and $\beta_1,...,\beta_N$ to solve (42) are the following counterparts to (34) and (35):

(43) $\Sigma_{n\in S(t)} [\rho_t + \beta_n] = \Sigma_{n\in S(t)} \ln p_{tn}$; $t = 1,...,T$;
(44) $\Sigma_{t\in S^*(n)} [\rho_t + \beta_n] = \Sigma_{t\in S^*(n)} \ln p_{tn}$; $n = 1,...,N$.

As in the previous section, let $\rho_t \equiv \ln\pi_t$ for t = 1,...,T and let $\beta_n \equiv \ln\alpha_n$ for n = 1,...,N. Substitute these definitions into equations (43) and (44). After some rearrangement and using definitions (40) and (41), equations (43) and (44) become the following ones:

(45) $\pi_t = \Pi_{n\in S(t)} [p_{tn}/\alpha_n]^{1/N(t)}$; $t = 1,...,T$;
(46) $\alpha_n = \Pi_{t\in S^*(n)} [p_{tn}/\pi_t]^{1/T(n)}$; $n = 1,...,N$.

The same iterative procedure that was explained in the previous section will work to generate a solution to equations (45) and (46).[30] As was the case in the previous section, solutions to (45) and (46) are not unique; if π^*, β^* is a solution to (45) and (46), then $\lambda\pi^*$ and $\lambda^{-1}\alpha^*$ is also a solution for any $\lambda > 0$. Thus to obtain a unique solution we impose the normalization $\pi_1^* = 1$ (which corresponds to $\rho_1 = 0$). Then $1,\pi_2^*,...,\pi_T^*$ is the sequence of (normalized) price levels that is generated by the least squares minimization problem defined by (42). In this case, $\pi_t^* = \Pi_{n\in S(t)}[p_{tn}/\alpha_n^*]^{1/N(t)}$ is the equally weighted geometric mean of all of the quality adjusted prices for the products that are available in period t for t = 2,3,...,T and the quality adjustment factors are normalized so that $\pi_1^* = \Pi_{n\in S(1)} [p_{1n}/\alpha_n^*]^{1/N(1)} = 1$. From (46), we can deduce that α_n^* will be larger for products that are relatively expensive and will be smaller for cheaper products.

[29] The first expression is used when (42) is differentiated with respect to π_t and the second expression is used when differentiating (42) with respect to β_n.
[30] Of course, it is not necessary to use the iterative procedure to find a solution to equations (43) and (44). After setting $\rho_1 = 0$ and dropping the first equation in (43), matrix algebra can be used to find a solution to the remaining equations. Alternatively, after setting $\rho_1 = 0$, use the equations $\ln p_{tn} = \rho_t + \beta_n + \varepsilon_{tn}$ for t = 1,...,T and $n\in S(t)$ to set up a linear regression model with time and product dummy variables and use a standard ordinary least squares econometric software package to obtain the solution $\rho_1^* = 0$, $\rho_2^*,...,\rho_T^*$, $\beta_1^*,...,\beta_N^*$ to (42).

Once we have the unique solution $1,\pi_2^*,...,\pi_T^*$ for the T price levels that are generated by the (42), the *price index* between period t relative to period r can be defined as π_t^*/π_r^*. Using equations (45) and (46), we have the following expressions for π_t^*/π_r^* and α_n^*/α_m^*:

(47) $\pi_t^*/\pi_r^* = \Pi_{n\in S(t)}[p_{tn}/\alpha_n^*]^{1/N(t)} / \Pi_{n\in S(r)}[p_{rn}/\alpha_n^*]^{1/N(r)}$; $1 \leq t, r \leq T$;
(48) $\alpha_n^*/\alpha_m^* = \Pi_{t\in S^*(n)}[p_{tn}/\pi_t^*]^{1/T(n)} / \Pi_{t\in S^*(m)}[p_{tm}/\pi_t^*]^{1/T(m)}$; $1 \leq n, m \leq N$.

Note that in general, the quality adjustment factors α_n^* do not cancel out for the indexes π_t^*/π_r^* defined by (47) as they did in the previous section. However, these price indexes do have some good axiomatic properties.[31] If the set of available products is the same in periods r and t, then the quality adjustment factors do cancel and the price index for period t relative to period r is $\pi_t^*/\pi_r^* = \Pi_{n\in S(t)}[p_{tn}/p_{rn}]^{1/N(t)}$, which is the Jevons index between periods r and t. Again, while this index is a excellent one if quantity information is not available, it is not satisfactory when quantity information is available due to its equal weighting of economically important and unimportant price ratios.

There is another unsatisfactory property of the estimated price levels that are generated by solving the time product dummy hedonic model that is defined by (42): a product that is available only in one period out of the T periods has no influence on the aggregate price levels π_t^*.[32] To see this, consider equations (43) and (44) and suppose that product n^* was only available in period t^*.[33] Equation n^* in N equations in (44) becomes the equation: $[\rho_{t^*} + \beta_{n^*}] = \ln p_{t^*n^*}$. Thus once ρ_{t^*} has been determined, β_{n^*} can be defined as $\beta_{n^*} \equiv \ln p_{t^*n^*} - \rho_{t^*}$. Subtract the equation $[\rho_{t^*} + \beta_{n^*}] = \ln p_{t^*n^*}$ from equation t^* and the resulting equations in (43) can be written as equations (49) below. Dropping equation n^* in equations (44) leads to equations (50) below:

(49) $\Sigma_{n\in S(t), n\neq n^*} [\rho_t + \beta_n] = \Sigma_{n\in S(t), n\neq n^*} \ln p_{tn}$; $t = 1,...,T$;
(50) $\Sigma_{t\in S^*(n)} [\rho_t + \beta_n] = \Sigma_{t\in S^*(n)} \ln p_{tn}$; $n = 1,...,n^*-1,n^*+1,...,N$.

Equations (49) and (50) are T+N−1 equations which do not involve $p_{t^*n^*}$. After making the normalization $\rho_1^* = 0$, these equations can be solved for $\rho_2^*,..., \rho_T^*, \beta_1^*,..., \beta_{n^*-1}^*, \beta_{n^*+1}^*,..., \beta_N^*$. Now define $\beta_{n^*}^* \equiv \ln p_{t^*n^*} - \rho_{t^*}$ and we have the (normalized) solution for (42). Since the ρ_t^* do not involve $p_{t^*n^*}$, the resulting $\pi_t^* \equiv \exp[\rho_t^*]$ for t = 1,...,T also do not depend on the isolated price $p_{t^*n^*}$. This proof can be repeated for any number of isolated prices. This property of the time product dummy model is unfortunate because it means that when a new product enters the marketplace in period T, it has no influence on the price levels $1,\pi_2^*,...,\pi_T^*$ that are generated by solving the least squares minimization problem defined by (42).

[31] The index π_t^*/π_r^* satisfies the identity test (if prices are the same in periods r and t, then the index is equal to 1) and it is invariant to changes in the units of measurement. It is also homogeneous of degree one in the prices of period t and homogeneous of degree minus one in the prices of period r.
[32] This property of the Time Product Dummy model was first noticed by Diewert (2004) (in the context of the Country Product Dummy model).
[33] We assume that products other than product n* are available in period t^*.

If quantity data are available, then aggregate quantity levels for the t periods can be obtained as $Q^{t*} \equiv \alpha^*\cdot q^t = \Sigma_{n=1}^N \alpha_n^* q_{tn}$ for t = 1,...,T.[34] Estimated aggregate price levels can be obtained directly from the solution to (42); i.e., set $P^{t*} = \pi_t^*$ for t = 1,...,T. Alternative price levels can be obtained indirectly as $P^{t**} \equiv p^t\cdot q^t/Q^{t*} = p^t\cdot q^t/\alpha^*\cdot q^t$ for t = 1,...,T. If the optimized objective function in (42) is 0, so that all errors $\varepsilon_{tn}^* \equiv \ln p_{tn} - \rho_t^* - \beta_n^*$ equal 0 for t = 1,...,T and $n \in S(t)$, then P^{t*} will equal P^{t**} for all t. If the estimated residuals are not all equal to 0, then the two estimates for the period t price level P^t will differ. The two estimates for P^t will generate different estimates for the companion aggregate quantity levels.

7. Weighted Time Product Dummy Regressions: The Bilateral Case

A major problem with the indexes discussed in the previous 2 sections is the fact that they do not weight the individual product prices by their economic importance. The first serious index number economist to stress the importance of weighting was Walsh (1901).[35] Keynes was quick to follow up on the importance of weighting[36] and Fisher emphatically endorsed weighting.[37] Griliches also endorsed weighting in the hedonic regression context.[38]

---

[34] Note that each $\alpha_n^* > 0$ since $\alpha_n^* \equiv \exp[\beta_n^*]$ for n = 1,...,N.

[35] See Walsh (1901). This book laid the groundwork for the test or axiomatic approach to index number theory which was further developed by Fisher (1922). In his second book on index number theory, Walsh made the case for weighting by economic importance as follows: "It might seem at first sight as if simply every price quotation were a single item, and since every commodity (any kind of commodity) has one price-quotation attached to it, it would seem as if price-variations of every kind of commodity were the single item in question. This is the way the question struck the first inquirers into price-variations, wherefore they used simple averaging with even weighting. But a price-quotation is the quotation of the price of a generic name for many articles; and one such generic name covers a few articles, and another covers many. … A single price-quotation, therefore, may be the quotation of the price of a hundred, a thousand, or a million dollar's worths, of the articles that make up the commodity named. Its weight in the averaging, therefore, ought to be according to these money-unit's worth." Correa Moylan Walsh (1921; 82-83).

[36] "It is also clear that the so-called unweighted index numbers, usually employed by practical statisticians, are the worst of all and are liable to large errors which could have been easily avoided." J.M. Keynes (1983; 79). This paper was written in 1909 and won the Cambridge University Adam Smith Prize for that year. Keynes (1930; 76-77) again stressed the importance of weighting in a later paper which drew heavily on his 1909 paper.

[37] "It has already been observed that the purpose of any index number is to strike a .fair average of the price movements or movements of other groups of magnitudes. At first a simple average seemed fair, just because it treated all terms alike. And, in the absence of any knowledge of the relative importance of the various commodities included in the average, the simple average is fair. But it was early recognized that there are enormous differences in importance. Everyone knows that pork is more important than coffee and wheat than quinine. Thus the quest for fairness led to the introduction of weighting." Irving Fisher (1922; 43).

[38] "But even here, we should use a weighted regression approach, since we are interested in an estimate of a weighted average of the pure price change, rather than just an unweighted average over all possible models, no matter how peculiar or rare." Zvi Griliches (1971; 8).

In this section, we will discuss some alternative methods for weighting by economic importance in the context of a bilateral time product dummy regression model.[39] Initially, we will assume that there are no missing observations.

Recall the least squares minimization problem defined by (33) in section 5 above. The squared residuals, $[\ln p_{tn} - \rho_t - \beta_n]^2$, appear in this problem without any weighting. Thus products which have a high volume of sales in any period are given the same weight in the least squares minimization problem as products which have very few sales. In order to take economic importance into account, for the case of 2 time periods, replace (33) by the following *weighted least squares minimization problem*:

$$(51)\ \min_{\rho,\beta} \Sigma_{n=1}^{N} q_{1n}[\ln p_{1n} - \beta_n]^2 + \Sigma_{n=1}^{N} q_{2n}[\ln p_{2n} - \rho_2 - \beta_n]^2$$

where we have set $\rho_1 = 0$. The squared error for product n in period t is repeated q_{tn} times to reflect the sales of the product in period t. Thus the new problem (51) takes into account the popularity of each product.

The first order necessary conditions for the minimization problem defined by (51) are the following N + 1 equations:

$$(52)\ (q_{1n}+q_{2n})\beta_n = q_{1n}\ln p_{1n} + q_{2n}(\ln p_{2n} - \rho_2)\ ; \qquad n = 1,\ldots,N;$$
$$(53)\ (\Sigma_{n=1}^{N} q_{2n})\rho_2 = (\Sigma_{n=1}^{N} q_{2n})(\ln p_{2n} - \beta_n).$$

The solution to (52) and (53) is the following one:[40]

$$(54)\ \rho_2^* \equiv \Sigma_{n=1}^{N} q_{1n}q_{2n}(q_{1n} + q_{2n})^{-1} \ln(p_{2n}/p_{1n})/\Sigma_{i=1}^{N} q_{1i}q_{2i}(q_{1i} + q_{2i})^{-1};$$
$$(55)\ \beta_n^* \equiv q_{1n}(q_{1n} + q_{2n})^{-1} \ln(p_{1n}) + q_{2n}(q_{1n} + q_{2n})^{-1} \ln(p_{2n}/\pi_2^*)\ ; \qquad n = 1,\ldots,N$$

where $\pi_2^* \equiv \exp[\rho_2^*]$. Note that the weight for the term $\ln(p_{2n}/p_{1n})$ in (54) can be written as follows:

$$(56)\ q_n^* \equiv \Sigma_{n=1}^{N} q_{1n}q_{2n}(q_{1n} + q_{2n})^{-1} /\Sigma_{i=1}^{N} q_{1i}q_{2i}(q_{1i} + q_{2i})^{-1}\ ; \qquad n = 1,\ldots,N$$
$$= h(q_{1n},q_{2n})/\Sigma_{i=1}^{N} h(q_{1i},q_{2i})$$

where $h(a,b) \equiv 2ab/(a+b) = [½\, a^{-1} + ½\, b^{-1}]^{-1}$ is the *harmonic mean* of a and b.[41]

Note that the q_n^* sum to 1 and thus ρ_2^* is a weighted average of the logarithmic price ratios $\ln(p_{2n}/p_{1n})$. Using $\pi_2^* = \exp[\rho_2^*]$ and $\pi_1^* = \exp[\rho_1^*] = \exp[0] =1$, the bilateral price index that is generated by the solution to (51) is

---

[39] The approach taken in this section is based on Rao (1995) (2004) (2005) and Diewert (2003b), (2005a) (2005b). Diewert (2005a) considered all four forms of weighting that will be discussed in this section while Rao (1995) (2005) discussed mainly the third form of weighting.

[40] See Diewert (2005a).

[41] h(a,b) is well defined by ab/(a+b) if a and b are nonnegative and at least one of these numbers is positive. In order to write h(a,b) as $[½\, a^{-1} + ½\, b^{-1}]^{-1}$, we require a > 0 and b > 0.

(57) $\pi_2^*/\pi^* = \exp[\rho_2^*] = \exp[\Sigma_{n=1}^N q_n^* \ln(p_{2n}/p_{1n})]$.

Thus π_2^*/π^* is a weighted geometric mean of the price ratios p_{2n}/p_{1n} with weights q_n^* defined by (56). Although this seems to be a reasonable bilateral index number formula, it must be rejected for practical use on the grounds that *the index is not invariant to changes in the units of measurement*.

Since values are invariant to changes in the units of measurement, the lack of invariance problem could be solved if we replace the quantity weights in (51) with expenditure or sales weights.[42] This leads to the following weighted least squares minimization problem where the weights v_{tn} are defined as $p_{tn}q_{tn}$ for t = 1,2 and n = 1,...,N:

(58) $\min_{\rho,\beta} \Sigma_{n=1}^N v_{1n}[\ln p_{1n} - \beta_n]^2 + \Sigma_{n=1}^N v_{2n}[\ln p_{2n} - \rho_2 - \beta_n]^2$.

It can be seen that problem (58) has exactly the same mathematical form as problem (51) except that v_{tn} has replaced q_{tn} and so the solutions (54) and (55) will be valid in the present context if v_{tn} replaces q_{tn} in these formulae. Thus the solution to (58) is:

(59) $\rho_2^* \equiv \Sigma_{n=1}^N v_{1n}v_{2n}(v_{1n} + v_{2n})^{-1} \ln(p_{2n}/p_{1n})/\Sigma_{i=1}^N v_{1i}v_{2i}(v_{1i} + v_{2i})^{-1}$;
(60) $\beta_n^* \equiv v_{1n}(v_{1n} + v_{2n})^{-1} \ln(p_{1n}) + v_{2n}(v_{1n} + v_{2n})^{-1} \ln(p_{2n}/\pi_2^*)$; n = 1,...,N

where $\pi_2^* \equiv \exp[\rho_2^*]$.

The resulting price index, $\pi_2^*/\pi_1^* = \pi_2^* = \exp[\rho_2^*]$ is indeed invariant to changes in the units of measurement. However, if we regard π_2^* as a function of the price and quantity vectors for the two periods, say $P(p^1,p^2,q^1,q^2)$, then another problem emerges for the price index defined by the solution to (58): $P(p^1,p^2,q^1,q^2)$ is not homogeneous of degree 0 in the components of q^1 or in the components of q^2. These properties are important because it is desirable that the companion implicit quantity index defined as $Q(p^1,p^2,q^1,q^2) \equiv [p^2\cdot q^2/p^1\cdot q^1]/P(p^1,p^2,q^1,q^2)$ be homogeneous of degree 1 in the components of q^2 and homogeneous of degree minus 1 in the components of q^1.[43] We also want $P(p^1,p^2,q^1,q^2)$ to be homogeneous of degree 1 in the components of p^2 and homogeneous of degree minus 1 in the components of p^1 and these properties are also not satisfied. Thus we conclude that the solution to the weighted least squares problem defined by (58) does not generate a satisfactory price index formula.

---

[42] "But on what principle shall we weight the terms? Arthur Young's guess and other guesses at weighting represent, consciously or un consciously, the idea that relative money values of the various commodities should determine their weights. A value is, of course, the product of a price per unit, multiplied by the number of units taken. Such values afford the only common measure for comparing the streams of commodities produced, exchanged, or consumed, and afford almost the only basis of weighting which has ever been seriously proposed." Irving Fisher (1922; 45).

[43] Thus we want Q to have the following properties: $Q(p^1,p^2,q^1,\lambda q^2) = \lambda Q(p^1,p^2,q^1,q^2)$ and $Q(p^1,p^2,\lambda q^1,q^2) = \lambda^{-1}Q(p^1,p^2,q^1,q^2)$ for all $\lambda > 0$. For a list of desirable properties or tests for bilateral price indexes of the form $P(p^1,p^2,q^1,q^2)$, see Diewert (1992) or the ILO, Eurostat, IMF, OECD, UNECE and the World Bank (2004).

The above deficiencies can be remedied if the *expenditure amounts* v_{tn} in (58) are replaced by *expenditure shares,* s_{tn}, where $v_t \equiv \Sigma_{n=1}^N v_{tn}$ for t = 1,2 and $s_{tn} \equiv v_{tn}/v_t$ for t = 1,2 and n = 1,...,N. This replacement leads to the following weighted least squares minimization problem:[44]

$$(61)\ \min_{\rho,\,\beta} \Sigma_{n=1}^N s_{1n}[\ln p_{1n} - \beta_n]^2 + \Sigma_{n=1}^N s_{2n}[\ln p_{2n} - \rho_2 - \beta_n]^2.$$

Again, it can be seen that problem (61) has exactly the same mathematical form as problem (51) except that s_{tn} has replaced q_{tn} and so the solutions (54) and (55) will be valid in the present context if s_{tn} replaces q_{tn} in these formulae. Thus the solution to (61) is:

$$(62)\ \rho_2^* \equiv \Sigma_{n=1}^N s_{1n}s_{2n}(s_{1n} + s_{2n})^{-1} \ln(p_{2n}/p_{1n})/\Sigma_{i=1}^N s_{1i}s_{2i}(s_{1i} + s_{2i})^{-1};$$

$$(63)\ \beta_n^* \equiv s_{1n}(s_{1n} + s_{2n})^{-1} \ln(p_{1n}) + s_{2n}(s_{1n} + s_{2n})^{-1} \ln(p_{2n}/\pi_2^*)\ ; \qquad n = 1,...,N$$

where $\pi_2^* \equiv \exp[\rho_2^*]$. Define the *normalized harmonic mean share weights* as follows:

$$(63)\ s_n^* \equiv h(s_{1n},s_{2n})/\Sigma_{i=1}^N h(s_{1i},s_{2i}); \qquad n = 1,...,N.$$

Then the weighted time product dummy bilateral price index, $P_{WTPD}(p^1,p^2,q^1,q^2) \equiv \pi_2^*/\pi_1^* = \pi_2^*$, has the following logarithm:

$$(64)\ \ln P_{WTPD}(p^1,p^2,q^1,q^2) \equiv \Sigma_{n=1}^N s_n^* \ln(p_{2n}/p_{1n}).$$

Thus $P_{WTPD}(p^1,p^2,q^1,q^2)$ is equal to a share weighted geometric mean of the price ratios, p_{2n}/p_{1n}, where the weight for ratio n is the average expenditure share s_n^* defined by (64). This index is a satisfactory one from the viewpoint of the test approach to index number theory. It can be shown that $P_{WTPD}(p^1,p^2,q^1,q^2)$ satisfies the following tests: (i) the identity test; i.e., $P_{WTPD}(p^1,p^2,q^1,q^2) = 1$ if $p^1 = p^2$; (ii) time reversal; i.e., $P_{WTPD}(p^2,p^1,q^2,q^1) = 1/P_{WTPD}(p^1,p^2,q^1,q^2)$;[45] (iii) homogeneity of degree 1 in period 2 prices; i.e., $P_{WTPD}(p^1,\lambda p^2,q^1,q^2) = \lambda P_{WTPD}(p^1,p^2,q^1,q^2)$; (iii) homogeneity of degree −1 in period 1 prices; i.e., $P_{WTPD}(\lambda p^1,p^2,q^1,q^2) = \lambda^{-1} P_{WTPD}(p^1,p^2,q^1,q^2)$; (iv) homogeneity of degree 0 in period 1 quantities; i.e., $P_{WTPD}(p^1,p^2,\lambda q^1,q^2) = \lambda P_{WTPD}(p^1,p^2,q^1,q^2)$; (v) homogeneity of degree 0 in period 2 quantities; i.e., $P_{WTPD}(p^1,p^2,q^1,\lambda q^2) = P_{WTPD}(p^1,p^2,q^1,q^2)$; (vi) invariance to changes in the units of measurement; (vii) $\min_n\{p_{2n}/p_{1n} : n = 1,...,N\} \leq$

[44] Note that the minimization problem defined by (61) is equivalent to the problem of minimizing $\Sigma_{n=1}^N e_{1n}^2 + \Sigma_{n=1}^N e_{2n}^2$ with respect to $\rho_2, \beta_1, ..., \beta_N$ where the error terms e_{tn} are defined by the equations $s_{1n}^{1/2}\ln p_{1n} = s_{1n}^{1/2}\beta_n + e_{1n}$ for n = 1,...,N and $s_{2n}^{1/2}\ln p_{2n} = s_{2n}^{1/2}\rho_2 + s_{2n}^{1/2}\beta_n + e_{2n}$ for n = 1,...,N. Thus the solution to (61) can be found by running a linear regression using the above two sets of estimating equations. The numerical equivalence of the least squares estimates obtained by repeating multiple observations or by using the square root of the weight transformation was noticed long ago as the following quotation indicates: "It is evident that an observation of weight w enters into the equations exactly as if it were w separate observations each of weight unity. The best practical method of accounting for the weight is, however, to prepare the equations of condition by multiplying each equation throughout by the square root of its weight." E. T. Whittaker and G. Robinson (1940; 224).

[45] See Diewert (2003b) (2005b).

$P_{WTPD}(p^1,p^2,q^1,q^2) \le \max_n\{p_{2n}/p_{1n} : n = 1,...,N\}$; and (viii) invariance to the ordering of the products. Moreover, Diewert (2005b; 564) showed that $P_{WTPD}(p^1,p^2,q^1,q^2)$ approximated the superlative Törnqvist-Theil index to the second order around an equal price and quantity point where $p^1 = p^2$ and $q^1 = q^2$.[46] Thus if changes in prices and quantities going from one period to the next are not too large, P_{WTPD} should be close to the superlative Fisher (1922) and Törnqvist-Theil indexes.[47]

Recall the results from section 5 above for the unweighted time product dummy model. From equation (38), it can be seen that the unweighted bilateral time product dummy regression model generated the Jevons index as the solution to the unweighted least squares minimization problem that is a counterpart to the weighted problem defined by (61) above. Thus appropriate weighting of the squared errors has changed the solution index dramatically: the index defined by (64) weights products by their economic importance and has good test properties whereas the Jevons index can generate very problematic results due to its lack of weighting according to economic importance. Note that both models have the same underlying structure; i.e., they assume that p_{tn} is approximately equal to $\pi_t\alpha_n$ for t = 1,2 and n = 1,...,N. *Thus weighting by economic importance has converted a least squares minimization problem that generates a rather poor price index into a problem that generates a rather good index.*

There is one more weighting scheme that generates an even better index in the bilateral context where we are running a time product dummy hedonic regression using the price and quantity data for only two periods. Consider the following weighted least squares minimization problem:

$$(65)\ \min_{\rho,\beta} \Sigma_{n=1}^N (½)(s_{1n}+s_{2n})[\ln p_{1n} - \beta_n]^2 + \Sigma_{n=1}^N (½)(s_{1n}+s_{2n})[\ln p_{2n} - \rho_2 - \beta_n]^2.$$

As usual, it can be seen that problem (65) has exactly the same mathematical form as problem (51) except that $(½)(s_{1n}+s_{2n})$ has replaced q_{tn} and so the solutions (54) and (55) will be valid in the present context if $(½)(s_{1n}+s_{2n})$ replaces q_{tn} in these formulae. Thus the solution to (65) simplifies to the following solution:

$$(66)\ \rho_2^* \equiv \Sigma_{n=1}^N (½)(s_{1n}+s_{2n})\ln(p_{2n}/p_{1n});$$
$$(67)\ \beta_n^* \equiv (½)\ln(p_{1n}) + (½)\ln(p_{2n}/\pi_2^*)\ ; \qquad n = 1,...,N$$

where $\pi_2^* \equiv \exp[\rho_2^*]$ and $\pi_1^* \equiv \exp[\rho_1^*] = \exp[0] = 1$ since we have set $\rho_1^* = 0$. Thus the bilateral index number formula which emerges from the solution to (65) is $\pi_2^*/\pi_1^* = \exp[\Sigma_{n=1}^N (½)(s_{1n}+s_{2n})\ln(p_{2n}/p_{1n})] \equiv P_T(p^1,p^2,q^1,q^2)$, which is the Törnqvist (1936)[48] Theil

[46] Diewert (2005a; 564) noted this result. Thus P_{WTPD} is a pseudo-superlative index. For the definition of a superlative index, see Diewert (1976) and for the definition of a pseudo-superlative index, see Diewert (1978).
[47] However, with large changes in price and quantities going from period 1 to 2, P_{WTPD} will tend to lie below its superlative counterparts; see Diewert (2018; 53) and the example in Diewert and Fox (2017; 24).
[48] See Törnqvist and Törnqvist (1937) for the actual formula. Diewert (1976) showed that this index number formula was superlative.

(1967; 137-138) bilateral index number formula.[49] Thus the use of the weights in (65) has generated an even better bilateral index number formula than the formula which resulted from the use of the weights in (61).[50] This result reinforces the case for using appropriately weighted versions of the basic time product dummy hedonic regression model.[51]

The aggregate quantity levels for the t periods can be obtained as $Q^{t*} \equiv \alpha^*\cdot q^t = \Sigma_{n=1}^N \alpha_n^* q_{tn}$ for t = 1,2 where the α_n^* are defined as the exponentials of the β_n^* defined by (67). Estimated aggregate price levels can be obtained directly from the solution to (65); i.e., set $P^{t*} = \pi_t^*$ for t = 1,2.[52] Alternative price levels can be obtained indirectly as $P^{t**} \equiv p^t\cdot q^t/Q^{t*} = p^t\cdot q^t/\alpha^*\cdot q^t$ for t = 1,2. If the optimized objective function in (65) is 0, so that all errors equal 0, then P^{t*} will equal P^{t**} for t = 1,2. If the estimated residuals are not all equal to 0, then the two estimates for the period t price level P^t will differ and the alternative estimates for P^t will generate different estimates for the companion aggregate quantity levels.

8. Weighted Time Product Dummy Regressions: The Bilateral Case with Missing Observations

In this section, we will generalize the last two models in the previous section to cover the case where there are missing observations.[53] Thus we assume that there are products that are missing in period 2 that were present in period 1 and some new products that appear in period 2. As in section 6 above, S(t) denotes the set of products n that are present in period t for t = 1,2. It is assumed that $S(1)\cap S(2)$ is not the empty set; i.e., there are one or more products that are present in both periods. The new weighted least squares minimization problem that generalizes (61) is the following minimization problem:[54]

$$(68)\ \min_{\rho,\beta} \Sigma_{n\in S(1)}\ s_{1n}[\ln p_{1n} - \beta_n]^2 + \Sigma_{n\in S(2)}\ s_{2n}[\ln p_{2n} - \rho_2 - \beta_n]^2.$$

The first order conditions for $\rho_2^*, \beta_1^*,..., \beta_N^*$ to solve (68) are equivalent to the following equations:

---

49 Theil's (1967; 137-138) derivation of $P_T(p^1,p^2,q^1,q^2)$ is equivalent to solving the following least squares minimization problem: $\min_\rho \Sigma_{n=1}^N (½)(s_{1n}+s_{2n})[\ln(p_{2n}/p_{1n}) - \rho]^2$ where $\rho \equiv \ln P_T(p^1,p^2,q^1,q^2)$. Thus the β_n do not appear in this minimization problem.

50 Diewert (1992; 223) lists the tests that $P_T(p^1,p^2,q^1,q^2)$ satisfies.

51 Note that the bilateral regression model defined by the minimization problem (61) is readily generalized to the case of T periods whereas the bilateral regression model defined by the minimization problem (65) cannot be generalized to the case of T periods.

52 In this case, alternative period t quantity levels are defined as $Q^{1**} \equiv p^1\cdot q^1$ and $Q^{2**} \equiv p^2\cdot q^2/\pi_2^* = [v_2/v_1]/P_T(p^1,p^2,q^1,q^2)$. If the squared errors in (65) are all 0, then the alternative quantity estimates are equal to each other and the model $\ln p_{tn} = \rho_t + \beta_n$ holds exactly for each tn and prices are proportional across the two periods; i.e., we have $p^t = \pi_t^*\alpha^*$ for t = 1,2 where $\alpha^* \equiv [\alpha_1^*,...,\alpha_N^*]$. In the case where the squared errors are nonzero, the π_t^*,Q^{t**} aggregates are preferred since $P_T(p^1,p^2,q^1,q^2)$ is a superlative index and thus has a stronger economic justification.

53 The results in this section are closely related to the results derived by de Haan (2004) and de Haan and Krsinich (2014). However, our method of derivation is somewhat different.

54 This form of weighting was suggested by Rao (1995) (2004) (2005), Diewert (2004) (2005a) and Haan (2004).

(69) $\Sigma_{n\in S(2)}\ s_{2n}[\ln p_{2n} - \rho_2^* - \beta_n^*] = 0$;
(70) $(s_{1n} + s_{2n})\beta_n^* = s_{1n}\ln p_{1n} + s_{2n}[\ln p_{1n} - \rho_2^*]$; $\qquad n\in S(1)\cap S(2)$;
(71) $\qquad \beta_n^* = \ln p_{1n}$; $\qquad n\in S(1),\ n\notin S(2)$;
(72) $\qquad \beta_n^* = \ln p_{2n} - \rho_2^*$; $\qquad n\in S(2),\ n\notin S(1)$.

Define the intersection set of products S^* as follows:

(73) $S^* \equiv S(1)\cap S(2)$.

Substituting equations (72) into equation (69) leads to the following equation:

(74) $\Sigma_{n\in S^*}\ s_{2n}[\ln p_{2n} - \rho_2^* - \beta_n^*] = 0$.

Consider the following least squares minimization problem that is defined over the set of products that are present in both periods:

(75) $\min_{\rho,\ \beta,\ n\in S^*} \Sigma_{n\in S^*}\ s_{1n}[\ln p_{1n} - \beta_n]^2 + \Sigma_{n\in S^*}\ s_{2n}[\ln p_{2n} - \rho_2 - \beta_n]^2$.

The first order conditions for this problem are (74) and (70). Once we find the solution to this problem, define β_n^* for the products that are not present in both periods by equations (71) and (72). This augmented solution will solve problem (68). The solution to (75) can be found by adapting the solution to (61) to the current situation. Recall equations (62) and (63) from the previous section. Replacing the entire set of product indices n = 1,...,N by the intersection set S^* defined by (73) leads to the following solution to (75):

(76) $\rho_2^* \equiv [\Sigma_{n\in S^*}\ s_{1n}s_{2n}(s_{1n} + s_{2n})^{-1}\ln(p_{2n}/p_{1n})]/[\Sigma_{i\in S^*}\ s_{1i}s_{2i}(s_{1i} + s_{2i})^{-1}]$;
(77) $\beta_n^* \equiv s_{1n}(s_{1n} + s_{2n})^{-1}\ln(p_{1n}) + s_{2n}(s_{1n} + s_{2n})^{-1}\ln(p_{2n}/\pi_2^*)$; $\qquad n\in S^*$

where $\pi_2^* \equiv \exp[\rho_2^*]$. Define the *normalized harmonic mean share weights* for the always present products as follows:

(78) $s_n^* \equiv h(s_{1n},s_{2n})/\Sigma_{i\in S^*} h(s_{1i},s_{2i})$; $\qquad n\in S^*$.

Denote the period t price and quantity vectors that include only always present products by p^{t*} and q^{t*} respectively for t = 1,2. Then the *weighted time product dummy bilateral price index with missing observations*, $P_{WTPDM}(p^{1*},p^{2*},q^{1*},q^{2*}) \equiv \pi_2^*/\pi_1^* = \pi_2^*$, has the following logarithm:

(78) $\ln P_{WTPDM}(p^{1*},p^{2*},q^{1*},q^{2*}) \equiv \Sigma_{n\in S^*}\ s_n^*\ln(p_{2n}/p_{1n})$.

To complete the solution to (68), use equations (71) and (72) to define the β_n^* for the new and missing products in period 2. Thus P_{WTPDM} *uses only the matched prices* to define the price index that results from solving the weighted least squares minimization problem with missing observations that is defined by (68).

As usual, the hedonic regression model that is generated by solving (68) can be used to impute reservation prices for missing observations. Thus define $\alpha_n^* \equiv \exp[\beta_n^*]$ for n = 1,...,N. Then the missing prices p_{tn}^* can be defined as follows:

$$(79)\ p_{2n}^* \equiv \pi_2^*\alpha_n^* = \pi_2^* p_{1n} \qquad n\in S(1),\ n\notin S(2);$$
$$(80)\ p_{1n}^* \equiv \pi_1^*\alpha_n^* = p_{2n}/\pi_2^* \qquad n\in S(2),\ n\notin S(1).$$

Thus the missing prices for period 2, p_{2n}^*, are the corresponding *inflation adjusted carry forward prices* from period 1, p_{1n} times π_2^* and the missing prices for period 1, p_{1n}^*, are the corresponding *inflation adjusted carry backward prices* from period 2, p_{2n} deflated by π_2^*, where π_2^* is the weighted time product dummy price index $P_{WTPDM}(p^{1*},p^{2*},q^{1*},q^{2*})$ defined by (78). Note the similarity of equations (79) and (80) the nonstochastic imputed prices defined by equations (25) in section 4.

Estimated aggregate price levels can be obtained directly from the solution to (75); i.e., set $P^{1*} = 1$ and $P^{2*} = \pi_2^*$. The corresponding quantity levels are defined as $Q^{1*} \equiv p^1\cdot q^1$ and $Q^{2*} \equiv p^2\cdot q^2/\pi_2^*$.[55] Alternative price and quantity levels can be obtained as $Q^{t**} \equiv \alpha^*\cdot q^t$ and $P^{t**} \equiv p^t\cdot q^t/Q^{t**}$ for t = 1,2. If the optimized objective function in (75) is 0, so that all errors equal 0, then P^{t*} will equal P^{t**} for all t. If the estimated residuals are not all equal to 0, then the two estimates for the period t price level P^t will differ and the alternative estimates for P^t will generate different estimates for the companion aggregate quantity levels.

The above analysis is not quite the end of the story. The expenditure shares s_{1n} and s_{2n} which appear in (75) are not the expenditure shares that characterize the always present products; they are the original expenditure shares defined over all N products. It is of interest to compare $P_{WTPDM}(p^{1*},p^{2*},q^{1*},q^{2*})$ defined by (78) with the "true" weighted time product dummy index, $P_{WTPDM}(p^{1*},p^{2*},q^{1*},q^{2*})$, that is defined over the common set of products, S^*.[56]

Define $v_t^* \equiv \Sigma_{n\in S^*} v_{tn}$ as the total expenditure on always present products for t = 1,2 and define the corresponding *restricted expenditure shares* as:

$$(81)\ s_{tn}^* \equiv v_{tn}/v_t^*\ ; \qquad t = 1,2;\ n\in S^*.$$

The restricted share version of (75) is the following weighted least squares minimization problem:

$$(82)\ \min_{\rho,\,\beta,\,n\in S^*} \Sigma_{n\in S^*}\ s_{1n}^*[\ln p_{1n} - \beta_n]^2 + \Sigma_{n\in S^*}\ s_{2n}^*[\ln p_{2n} - \rho_2 - \beta_n]^2.$$

---

[55] Use the imputed prices defined by (78) and (79) for the missing prices and set the missing quantity levels equal to 0. The resulting N dimensional vectors are denoted by p^t and q^t for t = 1,2 in the definitions which follow.

[56] Recall that S^* is defined by (73) and p^{t*} and q^{t*} are the period t price and quantity vectors that include only products that are present in both periods.

The ρ_2 solution to (82) is the following one:

$$(83)\ \rho_2^{**} \equiv [\Sigma_{n\in S^*}\ s_{1n}^*s_{2n}^*(s_{1n}^* + s_{2n}^*)^{-1}\ln(p_{2n}/p_{1n})]/[\Sigma_{i\in S^*}\ s_{1i}^*s_{2i}^*(s_{1i}^* + s_{2i}^*)^{-1}]$$
$$= [\Sigma_{n\in S^*}\ h(s_{1n}^*,s_{2n}^*)\ln(p_{2n}/p_{1n})]/[\Sigma_{i\in S^*}\ h(s_{1i}^*,s_{2i}^*)]$$

where $h(s_{1n}^*,s_{2n}^*)$ is the harmonic mean of the restricted shares s_{1n}^* and s_{2n}^*. The relationship between the *true shares*, the s_{tn}, and the *restricted shares*, the s_{tn}^*, for the always present products is given by the following equations:

$$(84)\ s_{tn} \equiv v_{tn}/v_t = [v_{tn}/v_t^*][v_t^*/v_t] = s_{tn}^*f_t\ ; \qquad t = 1,2\ ;\ n\in S^*$$

where the *fraction* of expenditures on always available commodities compared to expenditures on all commodities during period t is $f_t \equiv v_t^*/v_t$ for t = 1,2. Using equations (84) and (76), it can be seen that the logarithm of $P_{WTPDM}(p^{1*},p^{2*},q^{1*},q^{2*})$ defined by (78) is equal to the following expression:

$$(85)\ \rho_2^* \equiv [\Sigma_{n\in S^*}\ h(s_{1n},s_{2n})\ln(p_{2n}/p_{1n})]/[\Sigma_{i\in S^*}\ h(s_{1i},s_{2i})]$$
$$= [\Sigma_{n\in S^*}\ h(f_1s_{1n}^*,f_2s_{2n}^*)\ln(p_{2n}/p_{1n})]/[\Sigma_{i\in S^*}\ h(f_1s_{1i}^*,f_2s_{2i}^*)].$$

If either: (i) $p_{2n} = \lambda p_{1n}$ for all $n\in S^*$ so that we have price proportionality for the always present products or (ii) $f_1 = f_2$ so that the ratio of expenditures on always present products to total expenditure in each period is constant across the two periods, then $\rho_2^{**} = \rho_2^*$. However, if these conditions are not satisfied and there is considerable variation in prices and quantities across periods, then ρ_2^{**} could differ substantially from ρ_2^*. Since neither index is superlative, it is difficult to recommend one of these indexes over the other as the "optimal" carry forward and backward inflation rate which could be used to fill in missing prices.

We conclude this section by considering the weighting scheme suggested by de Haan (2004). The following weighted least squares minimization problem is his generalization of (65) to the case of missing observations:

$$(86)\ \min_{\rho,\beta}\ \Sigma_{n\in S^*}(½)(s_{1n}+s_{2n})[\ln p_{1n} - \beta_n]^2 + \Sigma_{n\in S(1),\, n\notin S(2)}(½)s_{1n}[\ln p_{1n} - \beta_n]^2$$
$$+ \Sigma_{n\in S^*}(½)(s_{1n}+s_{2n})[\ln p_{2n} - \rho_2 - \beta_n]^2 + \Sigma_{n\in S(2),\, n\notin S(1)}(½)s_{2n}[\ln p_{2n} - \rho_2 - \beta_n]^2.$$

The first order conditions that a solution to (86) satisfies simplify to the following conditions:

$$(87)\ \Sigma_{n\in S^*}(s_{1n}+s_{2n})\rho_2^* + \Sigma_{n\in S(2),\, n\notin S(1)}\ s_{2n}\rho_2^*$$
$$= \Sigma_{n\in S^*}(s_{1n}+s_{2n})[\ln p_{2n} - \beta_n^*] + \Sigma_{n\in S(2),\, n\notin S(1)}\ s_{2n}[\ln p_{2n} - \beta_n^*]\ ;$$
$$(88)\ 2\beta_n^* = \ln p_{1n} + \ln p_{2n} - \rho_2^*\ ; \qquad n\in S^*;$$
$$(89)\ \beta_n^* = \ln p_{1n}\ ; \qquad n\in S(1),\ n\notin S(2);$$
$$(90)\ \beta_n^* = \ln p_{2n} - \rho_2^*\ ; \qquad n\in S(2),\ n\notin S(1).$$

Substitute equations (88) and (90) into equation (87). The resulting equation simplifies to the following equation:

$$(91)\ \rho_2^* = \Sigma_{n\in S^*} \tfrac{1}{2}(s_{1n}+s_{2n})\ln(p_{2n}/p_{1n})/\Sigma_{i\in S^*} \tfrac{1}{2}(s_{1i}+s_{2i})$$
$$= \Sigma_{n\in S^*} \tfrac{1}{2}(f_1 s_{1n}^*+f_2 s_{2n}^*)\ln(p_{2n}/p_{1n})/\Sigma_{i\in S^*} \tfrac{1}{2}(f_1 s_{1i}^*+f_2 s_{2i}^*)$$

where the second equation follows using equations (84). Recall that the s_{tn} are the true expenditure shares and the s_{tn}^* are restricted expenditure shares that sum to one when only always present products are in scope. Once ρ_2^* has been determined, the β_n^* are determined using equations (88)-(90). As usual, define $\pi_1^* \equiv 1$ and $\pi_2^* \equiv \exp[\rho_2^*]$ where ρ_2^* is defined by (91). Then the *Haan weighted time product dummy bilateral price index with missing observations*, $P_{WTPDH}(p^{1*},p^{2*},q^{1*},q^{2*}) \equiv \pi_2^*/\pi_1^* = \pi_2^*$, has the following logarithm:

$$(92)\ \ln P_{WTPDH}(p^{1*},p^{2*},q^{1*},q^{2*}) \equiv \Sigma_{n\in S^*}\ s_n^* \ln(p_{2n}/p_{1n})$$

where the shares s_n^* are defined as follows:

$$(93)\ s_n^* \equiv \tfrac{1}{2}(f_1 s_{1n}^*+f_2 s_{2n}^*)/\Sigma_{i\in S^*} \tfrac{1}{2}(f_1 s_{1i}^*+f_2 s_{2i}^*)\ ; \qquad n\in S^*.$$

If either: (i) $p_{2n} = \lambda p_{1n}$ for all $n\in S^*$ so that we have price proportionality for the always present products or (ii) $f_1 = f_2$ so that the ratio of expenditures on always present products to total expenditure in each period is constant across the two periods, then $P_{WTPDH}(p^{1*},p^{2*},q^{1*},q^{2*}) = P_T(p^{1*},p^{2*},q^{1*},q^{2*})$, the Törnqvist Theil index defined over the always present products.[57] Note that P_{WTPDHK} *uses only the matched prices* to define the price index that results from solving the weighted least squares minimization problem with missing observations that is defined by (86).

As usual, the hedonic regression model that is generated by solving (68) can be used to impute reservation prices for missing observations. Thus define $\alpha_n^* \equiv \exp[\beta_n^*]$ for n = 1,...,N. Then the missing prices p_{tn}^* can be defined as follows:

$$(94)\ p_{2n}^* \equiv \pi_2^*\alpha_n^* = \pi_2^* p_{1n} \qquad n\in S(1),\ n\notin S(2);$$
$$(95)\ p_{1n}^* \equiv \pi_1^*\alpha_n^* = p_{2n}/\pi_2^* \qquad n\in S(2),\ n\notin S(1).$$

Thus the missing prices for period 2, the p_{2n}^* defined by (94), are the corresponding *inflation adjusted carry forward prices* from period 1, p_{1n} times π_2^* and the missing prices for period 1, p_{1n}^* defined by (95), are the corresponding *inflation adjusted carry backward prices* from period 2, p_{2n} deflated by π_2^*, where π_2^* is the de Haan weighted time product dummy price index $P_{WTPDH}(p^{1*},p^{2*},q^{1*},q^{2*})$ defined by (92). Again, note the similarity of equations (94) and (95) to the nonstochastic imputed prices defined by equations (25) in section 4.

[57] This result was first derived by de Haan and Krsinich (2014; 347).

Comparing $P_{WTPDM}(p^{1*},p^{2*},q^{1*},q^{2*})$ defined by (78) and $P_{WTPDH}(p^{1*},p^{2*},q^{1*},q^{2*})$ defined by (92), the latter index is preferred due to the fact that it collapses to the superlative Törnqvist Theil index $P_T(p^{1*},p^{2*},q^{1*},q^{2*})$ if $f_1 = f_2$. If $f_1 = f_2$, $P_{WTPDM}(p^{1*},p^{2*},q^{1*},q^{2*})$ does not collapse to $P_T(p^{1*},p^{2*},q^{1*},q^{2*})$.

Finally, de Haan and Krsinich (2014; 345) derived an interesting property of the Haan index defined by (92): define the missing shares and quantities as follows:

(96) $s_{2n} \equiv 0$ and $q_{2n} \equiv 0$; $\quad n\in S(1), n\notin S(2)$;
(97) $s_{1n} \equiv 0$ and $q_{2n} \equiv 0$; $\quad n\in S(2), n\notin S(1)$.

Define the missing prices by (94) and (95). Denote the now complete price and quantity vectors for period t as p^t and q^t for t = 1,2. Calculate the Törnqvist Theil index over the complete data as $P_T(p^1,p^2,q^1,q^2)$. It turns out that $P_T(p^1,p^2,q^1,q^2) = P_{WTPDH}(p^{1*},p^{2*},q^{1*},q^{2*})$ where the logarithm of the latter index is defined by (91).

To prove this result, note that for $n\in S(2)$, $n\notin S(1)$, we have $s_{1n} = 0$ using (97). Thus we can add $\Sigma_{n\in S(2),n\notin S(1)}\, s_{1n}\rho_2^*$ to the left hand side of (87) and add $\Sigma_{n\in S(2),n\notin S(1)}\, s_{1n}[\ln p_{2n} - \beta_n^*]$ to the right hand side of (87). The resulting equation is:

$$(98)\ \Sigma_{n\in S^*}(s_{1n}+s_{2n})\rho_2^* + \Sigma_{n\in S(2),\, n\notin S(1)}(s_{1n} + s_{2n})\rho_2^* = \Sigma_{n\in S^*}(s_{1n}+s_{2n})[\ln p_{2n} - \beta_n^*] + \Sigma_{n\in S(2),\, n\notin S(1)}(s_{1n}+s_{2n})[\ln p_{2n} - \beta_n^*]\ .$$

Upon taking logarithms of both sides of equations (94), it can be seen that we can add $\Sigma_{n\in S(1),n\notin S(2)}(s_{1n}+s_{2n})\rho_2^*$ to the left hand side of (87) and add $\Sigma_{n\in S(2),n\notin S(1)}(s_{1n}+s_{2n})[\ln p_{2n}^* - \beta_n^*]$ to the right hand side of (98). The resulting equation is:

$$(99)\ \Sigma_{n=1}^{N}(s_{1n}+s_{2n})\rho_2^* = \Sigma_{n=1}^{N}(s_{1n}+s_{2n})[\ln p_{2n} - \beta_n^*]$$

where p_{2n}^* is denoted as p_{2n} for $n\in S(1)$ and $n\notin S(2)$ in (99). Our next task is to eliminate the sum $\Sigma_{n=1}^{N}(s_{1n}+s_{2n})\beta_n^*$ from the right hand side of (99). Using equations (88), it is straightforward to derive the following equation:

$$(100)\ \Sigma_{n\in S^*}(s_{1n}+s_{2n})\beta_n^* = ½\Sigma_{n\in S^*}(s_{1n}+s_{2n})[\ln p_{1n} + \ln p_{2n} - \rho_2^*].$$

Upon taking logarithms of both sides of equations (94), we obtain the equations $\beta_n^* = \ln p_{2n}^* - \rho_2^*$ for $n\in S(1)$, $n\notin S(2)$. Using these equations and equations (89), it can be seen that we obtain the equations $2\beta_n^* = [\ln p_{1n} + \ln p_{2n}^* - \rho_2^*]$ for $n\in S(1)$, $n\notin S(2)$ and these equations imply the following equation:

$$(101)\ \Sigma_{n\in S(1),n\notin S(2)}(s_{1n}+s_{2n})\beta_n^* = ½\Sigma_{n\in S(1),n\notin S(2)}(s_{1n}+s_{2n})[\ln p_{1n} + \ln p_{2n}^* - \rho_2^*].$$

Upon taking logarithms of both sides of equations (95), we obtain the equations $\beta_n^* = \ln p_{1n}^*$ for $n\in S(2)$, $n\notin S(1)$. Using these equations and equations (90), it can be seen that we obtain the equations $2\beta_n^* = [\ln p_{1n}^* + \ln p_{2n} - \rho_2^*]$ for $n \in S(2)$, $n \notin S(1)$ and these equations imply the following equation:

(102) $\Sigma_{n\in S(2),n\notin S(1)} (s_{1n}+s_{2n})\beta_n^* = \frac{1}{2}\Sigma_{n\in S(2),n\notin S(1)} (s_{1n}+s_{2n})[\ln p_{1n}^* + \ln p_{2n} - \rho_2^*]$.

Equations (100)-(102) imply that the following equation holds:

(103) $\Sigma_{n=1}^N (s_{1n}+s_{2n})\beta_n^* = \frac{1}{2}\Sigma_{n=1}^N (s_{1n}+s_{2n})[\ln p_{1n} + \ln p_{2n} - \rho_2^*]$

where we have denoted the implicit prices p_{tn}^* in (103) by p_{tn} where necessary. Finally substitute (103) into (99) and using $\Sigma_{n=1}^N s_{tn} = 1$ for $t = 1,2$, we obtain the following expression for ρ_2^*:

(104) $\rho_2^* = \frac{1}{2}\Sigma_{n=1}^N (s_{1n}+s_{2n})\ln(p_{2n}/p_{1n})$.

Thus the ρ_2^* defined by (91) is also equal to the ρ_2^* defined by (104) and thus the Haan index $P_{WTPDH}(p^{1*},p^{2*},q^{1*},q^{2*})$ defined by (92) over the set of products that are available in both periods is also equal to the Törnqvist Theil index over the complete data, $P_T(p^1,p^2,q^1,q^2)$, where the missing prices are defined by (94) and (95).

In the following section, we define weighted time dummy regression models for the general case of T periods.

9. Weighted Time Product Dummy Regressions: The T period Case

10. Time Dummy Hedonic Regression Models with Characteristics Information

11. Hedonics and the Problem of Taste Change: Hedonic Imputation Indexes

12. Estimating Reservation Prices: The Case of CES Preferences

13. Estimating Reservation Prices: The Case of KBF Preferences

14. Conclusion

References

Aizcorbe, A., C. Corrado and M. Doms (2000), "Constructing Price and Quantity Indexes for High Technology Goods", Industrial Output Section, Division of Research and Statistics, Board of Governors of the Federal Reserve System, Washington DC.

Court, A.T. (1939), "Hedonic Price Indexes with Automotive Examples", pp. 99-117 in *The Dynamics of Automobile Demand*, New York: General Motors Corporation.

Diewert, W.E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics* 4, 114-145.

Diewert, W.E. (1978), "Superlative Index Numbers and Consistency in Aggregation", *Econometrica* 46, 883-900.

Diewert, W.E. (1987), "Index Numbers", pp. 767-780 in J. Eatwell, M. Milgate and P. Newman, (eds.), The New Pa1grave: A Dictionary of Economics, London: The Macmillan Press.

Diewert, W.E. (1992), "Fisher Ideal Output, Input and Productivity Indexes Revisited", *Journal of Productivity Analysis* 3, 211-248.

Diewert, W.E. (1993), "The Early History of Price Index Research", pp. 33-65 in *Essays in Index Number Theory*, Volume 1, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North-Holland.

Diewert, W.E. (2002), "Weighted Country Product Dummy Variable Regressions and Index Number Formulae", Department of Economics, Discussion Paper 02-15, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z1.

Diewert, W.E. (2003a), "Hedonic Regressions: A Consumer Theory Approach", in Scanner Data and Price Indexes, Studies in Income and Wealth (Vol. 61), eds.

R.C. Feenstra and M.D. Shapiro, Chicago: University of Chicago Press, pp. 317-348.

Diewert, W.E. (2003b), “Hedonic Regressions: A Review of Some Unresolved Issues”, Paper presented at the Seventh Meeting of the Ottawa Group, Paris, 27–29 May.

Diewert, W.E. (2004), “On the Stochastic Approach to Linking the Regions in the ICP”, Discussion Paper no. 04-16, Department of Economics, The University of British Columbia, Vancouver, Canada.

Diewert, W.E. (2005a), “Weighted Country Product Dummy Variable Regressions and Index Number Formulae”, *Review of Income and Wealth* 51, 561-570.

Diewert, W.E. (2005b), “Adjacent Period Dummy Variable Hedonic Regressions and Bilateral Index Number Theory”, *Annales D’Économie et de Statistique*, No. 79/80, 759-786.

Diewert, W.E., S. Heravi and M. Silver (2009), “Hedonic Imputation versus Time Dummy Hedonic Indexes”, pp. 87-116 in W.E. Diewert, J. Greenlees and C. Hulten (eds.), *Price Index Concepts and Measurement*, Studies in Income and Wealth, Vol. 70. Chicago: University of Chicago Press.

Diewert, W.E. (2018), “Scanner Data, Elementary Price Indexes and the Chain Drift Problem”, Discussion Paper 18-06, Vancouver School of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

Diewert, W.E. and R. Feenstra (2017), “Estimating the Benefits and Costs of New and Disappearing Products”, Vancouver School of Economics, Discussion Paper 17-10, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

Diewert, W.E. and K.J. Fox (2017), “Substitution Bias in Multilateral Methods for CPI Construction using Scanner Data”, Discussion Paper 17-02, Vancouver School of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

Diewert, W.E., K.J. Fox and P. Schreyer (2017), “The Digital Economy, New Products and Consumer Welfare”, Vancouver School of Economics Discussion Paper 17-09, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

Diewert, W.E. and R.J. Hill (2010), “Alternative Approaches to Index Number Theory”, pp. 263-278 in *Price and Productivity Measurement*, W.E. Diewert, Bert M. Balk, Dennis Fixler, Kevin J. Fox and Alice O. Nakamura (eds.), Victoria Canada: Trafford Press.

Feenstra, R.C. (1994), “New Product Varieties and the Measurement of International Prices”, *American Economic Review* 84:1, 157-177.

Fisher, I. (1922), *The Making of Index Numbers*, Houghton-Mifflin, Boston.

Griliches, Z. (1971b), "Introduction: Hedonic Price Indexes Revisited", pp. 3-15 in *Price Indexes and Quality Change*, Z. Griliches (ed.), Cambridge MA: Harvard University Press.

Griliches, Z. (1971), "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change", in Price Indexes and Quality Change, Z. Griliches (ed.), Harvard University Press, Cambridge (Mass.), 55-87.

Gorajek, A. (2018), "Econometric Perspectives on Economic Measurement", Research Discussion Paper 2018-08, Reserve Bank of Australia, 65 Martin Pl, Sydney NSW 2000.

de Haan, J. (2004), "The Time Dummy Index as a Special Case of the Imputation Törnqvist Index," paper presented at The Eighth Meeting of the International Working Group on Price Indices, Helsinki, Finland.

de Haan, J. (2009). Comment on "Hedonic Imputation Versus Time Dummy Hedonic Indexes". In E. Diewert, J. Greenlees, and C. Hulten (eds.), *Price Index Concepts and Measurements*, pp. 196-200. University of Chicago Press.

de Haan, J. and F. Krsinich (2012), "The Treatment of Unmatched Items in Rolling Year GEKS Price Indexes: Evidence from New Zealand Scanner Data", paper presented at the Economic Measurement Group Workshop 2012, Australian School of Business, University of New South Wales, November 23.

de Haan, J. and F. Krsinich (2014), "Scanner Data and the Treatment of Quality Change in Nonrevisable Price Indexes", *Journal of Business and Economic Statistics* 32:3, 341-358.

Hicks, J.R. (1940), "The Valuation of the Social Income", *Economica* 7, 105–124.

Hofsten, E. von (1952), *Price Indexes and Quality Change*, London: George Allen and Unwin.

*Hulten, C. (2003), "Price Hedonics: A Critical Review", *Economic Policy Review* 9:3, 5-15.

ILO, Eurostat, IMF, OECD, UNECE and the World Bank (2004), *Consumer Price Index Manual: Theory and Practice*, Peter Hill (ed.), Geneva: International Labour Office.

Keynes, J.M. (1930), *Treatise on Money*, Vol. 1. London: Macmillan.

Keynes, J.M. (1983), "The Method of Index Numbers with Special Reference to the Measurement of General Exchange Value", pp. 49-156 in *The Collected Writings* *of John Maynard Keynes*, Volume 11, D. Moggridge (ed.), Cambridge: Cambridge University Press.

Konüs, A.A. (1924), “The Problem of the True Index of the Cost of Living”, translated in *Econometrica* 7, (1939), 10-29.

Konüs, A.A. and S.S. Byushgens (1926), “K Probleme Pokupatelnoi Cili Deneg”, *Voprosi Konyunkturi* 2, 151-172.

Marshall, A. (1887), “Remedies for Fluctuations of General Prices’, *Contemporary Review* 51, 355-375.

Muellbauer, J. (1974), “Household Production Theory, Quality and the .Hedonic Technique”, *American Economic Review* 64:6, 977-994.

Pakes, A. (2001), “A Reconsideration of Hedonic Price Indices with and Application to PCs”, NBER Working Paper 8715, Cambridge MA..

*Pakes, A. (2003), “A Reconsideration of Hedonic Price Indexes with an Application to PC’s”, *American Economic Review*, 93, 1578-1596.

Rao, D.S. Prasada (1995), “On the Equivalence of the Generalized Country-Product-Dummy (CPD) Method and the Rao-System for Multilateral Comparisons”, Working Paper No. 5, Centre for International Comparisons, University of Pennsylvania, Philadelphia.

Rao, D.S. Prasada (2004), “The Country-Product-Dummy Method: A Stochastic Approach to the Computation of Purchasing Power parities in the ICP”, paper presented at the SSHRC Conference on Index Numbers and Productivity Measurement, June 30-July 3, 2004, Vancouver, Canada.

Rao, D.S. Prasada (2005), “On the Equivalence of the Weighted Country Product Dummy (CPD) Method and the Rao System for Multilateral Price Comparisons”, Review of Income and Wealth 51:4, 571-580.

Rao, D.S. Prasada and G. Hajargasht (2016), “Stochastic Approach to Computation of Purchasing Power Parities in the International Comparison Program”, *Journal of Econometrics* 191:2, 414-425.

Summers, R. (1973), “International Price Comparisons Based Upon Incomplete Data”, Review of Income and Wealth 19, 1-16.

Theil, H. (1967), *Economics and Information Theory*, Amsterdam: North-Holland.

Törnqvist, L. (1936), “The Bank of Finland's Consumption Price Index”, *Bank of Finland Monthly Bulletin* 10, 1-8.

Törnqvist, L. and E. Törnqvist (1937), Vilket är förhällandet mellan finska markens och svenska kronans köpkraft?", *Ekonomiska Samfundets Tidskrift* 39, 1-39 reprinted as pp. 121-160 in *Collected Scientific Papers of Leo Törnqvist*, Helsinki: The Research Institute of the Finnish Economy, 1981.

Triplett, J. (1987), "Hedonic Functions and Hedonic Indexes", pp. 630-634 in: John Eatwell, Murray Milgate, and Peter Newman (eds.), *The New Palgrave: A Dictionary of Economics*, Volume 2. New York, NY: Stockton Press.

Triplett, J. (2004), *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*, Directorate for Science, Technology and Industry, DSTI/DOC(2004)9, Paris: OECD.

Triplett, J. E. and R. J. McDonald (1977), "Assessing the Quality Error in Output Measures: The Case of Refrigerators", *The Review of Income and Wealth* 23:2, 137-156.

Walsh, C.M. (1901), *The Measurement of General Exchange Value*, New York: Macmillan and Co.

Walsh, C.M. (1921), *The Problem of Estimation*, London: P.S. King & Son.

Whittaker, E. T. and G. Robinson (1940), *The Calculus of Observations*, Third Edition, London: Blackie & Sons.