Quality Adjustment and Hedonics: An Attempt at a Unified Approach

by Erwin Diewert
University of British Columbia and University of New South Wales

EMG 2018 Workshop

University of New South Wales

Centre for Economic Applied Research,

UNSW Business School and

The School of Economics

November 29-30, 2018

Introduction

- **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**

(i) the use of inflation adjusted carry forward and carry backward prices,

(ii)the use of hedonic regressions and

(iii) the estimation of Hicksian reservation prices.

- **But unfortunately, the paper is not finished and so I will not be able to cover all of the above topics! Sorry about this!**
- **However, last year at the EMG, I did cover the last topic!**

The Basic Consumer Theory Framework

- **Notation: Let $p^t \equiv [p_{t1},\ldots,p_{tN}]$ and $q^t \equiv [q_{t1},\ldots,q_{tN}]$ 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.**
- **Initially, we assume that all prices, quantities and values are positive; in subsequent sections, this assumption will be relaxed.**
- **I have in mind a scanner data context for an elementary category.**

The Basic Consumer Theory Framework (cont)

- **Let $q \equiv [q_1,\ldots,q_N]$ 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.**

Properties of $Q(q^t)$

- **Q(q) has the following properties:**

  (i) $Q(q) > 0$ if $q >> 0_N$;

  (ii) Q(q) is nondecreasing in its components;

  (iii) $Q(\lambda q) = \lambda Q(q)$ for $q \geq 0_N$ and $\lambda \geq 0$; (linear homogeneity);

  (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, $P(p,q) \equiv p \cdot q / Q(q)$ that corresponds to Q(q) does not depend on the scale of q.**
- **Property (iv) will ensure that the first order necessary conditions for the budget constrained maximization of Q(q) are also sufficient.**

The Aggregate Price Level Defined

- Let $p \equiv [p_1,...,p_N] > 0_N$ and $q \equiv [q_1,...,q_N] > 0_N$ be generic price and quantity vectors 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, $p \cdot q$. (Product Test for Levels).

More Introduction

- 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) $Q^t \equiv Q(q^t)$; $t = 1,\ldots,T$.

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

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

- 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*.

The Assumption of Maximizing Behavior is Introduced

- **Two additional assumptions are made:**

(v) Q(q) is once differentiable with respect to the components of q;

(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:

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

- **The first order conditions for solving (4) for period t are the following conditions:**

(5) $\nabla_q Q(q^t) = \lambda_t p^t$; $t = 1,\ldots,T$;

(6) $p^t \cdot q^t = v^t$; $t = 1,\ldots,T$.

- **This theory dates back to Konüs and Byushgens (1926), Shephard (1953) (in the context of a cost minimization framework), Samuelson and Swamy (1974) and Diewert (1976).**

Some Implications of Maximizing Behavior

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

- 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) $\lambda_t = Q(q^t)/p^t \cdot q^t$ $\quad t = 1,\ldots,T$

$= 1/P^t$

using definitions (3).

- Thus the Lagrange multipliers for the utility maximization problems are equal to the reciprocals of the aggregate price levels.

Additional Implications of Maximizing Behavior

- **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) $P^tQ^t = p^t \cdot q^t = v_t$; t = 1,...,T. (the product test for levels).**
- **Substitute equations (8) into equations (5) and after a bit of rearrangement, the following *fundamental equations* are obtained:**
- **(10) $p^t = P^t \nabla_q Q(q^t)$; t = 1,...,T. (Note the appearance of P^t here).**

The Path Forward

- **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.**

A Nonstochastic Method for Quality Adjustment: A Simple Model

- **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).**
- **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:**

The Simple 3 Product, 2 Period Model

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

- **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.**

The Simple 3 Product, 2 Period Model (cont)

- **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) $P^1 = 1$.

- **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*.**

Reservation Prices for the Missing Prices

- **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*; i.e., they are the lowest possible prices that are just high enough to deter purchasers from purchasing the products during periods if the unavailable products hypothetically became available.**
- **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.**

Reservation Prices for the Missing Prices (cont)

- **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.**
- **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).**
- **The simple model explained above provides a consumer theory justification for the use of these imputed prices.**

Two Methods for Computing Price and Quantity Levels

- **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, $P^1 = 1$ 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 $q^2 \equiv [q_{21}, 0, q_{23}]$.**

Two Methods for Computing Price and Quantity Levels (cont)

- **The period t aggregate quantity level Q^t can be calculated directly using only information on q^t and the vector of 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) $Q^t \equiv \alpha \cdot q^t$; or $Q^t \equiv p^t \cdot q^t / P^t$; $t = 1,2$.

- **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.**

A More Complicated Model

- **A problem with the above simple model is that there is only one product that is present in both periods. We need to generalize the simple model to allow for multiple overlapping products.**
- **In order to generalize the very simple model for dealing with new and disappearing products that was presented in the previous section, we develop another application of the fundamental equations (10).**
- **Define the aggregator function Q(q) as follows:**
- **(17) $Q_{KBF}(q^*) \equiv [q^* \cdot A q^*]^{1/2}$ (Note that I used this function last year)**
- **where q^* is defined as the N dimensional quantity vector $[q_1^*,\ldots,q_N^*]$ and $A \equiv [a_{ij}]$ is an N by N symmetric matrix of parameters which satisfies certain regularity conditions.**
- **Suppose further that that the observed price and quantity vectors for periods 1 and 2 are the positive price and quantity vectors, $p^{t*} \equiv [p_{t1}^*,\ldots,p_{tN}^*]$ and $q^{t*} \equiv [q_{t1}^*,\ldots,q_{tN}^*]$ for t = 1,2.**
- **Why the stars? You will see why in due course!**

The KBF Model

- **The model that we are about to develop is due to Konüs and Byushgens (1926) who showed the relationship of the KB functional form defined by (17) to the Fisher (1922) ideal index.**
- **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 $P^{t*} \equiv p^{t*}\cdot q^{t*}/Q^{t*}$ for t = 1,2.**
- **In the following slide, we define the *Fisher* (1922) *ideal quantity index*:**

The KBF Model (cont)

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

The KBF Model (cont)

- **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:**

A More General New and Disappearing Product Model

(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.

A More General New and Disappearing Product Model (cont)

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

- **where $Q_{KBF}(q^{*})$ is the KBF aggregator function defined by (17) and α_1 is set equal to 1.**
- **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.**

Time Product Dummy Regressions: The Case of No Missing Observations and Equal Weighting

- **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, we assume 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:**

Time Product Dummy Regressions (cont)

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

- where we have redefined the period t price levels P^t in equations (10) as the parameters π_t for $t = 1,\ldots,T$.
- Note that equations (28) form the basis for the ***time dummy hedonic regression model*** which is due to Court (1939). 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) 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.

Time Product Dummy Regressions (cont)

- **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).**
- **Here are the two key questions that we need to address:**

  (i) How exactly are the error terms to be introduced into the exact equations (28)?

  (ii) Should we weight equations (28) according to their economic importance and if so, what weights should be used?
- **Our approach to answering both questions will be a pragmatic one. We will experiment with different ways of introducing error terms and weights into equations (28) and reject specifications which give rise to indexes which have awkward axiomatic or economic properties.**

Time Product Dummy Regressions (cont)

- **We will postpone the weighting problem for a while and look at different ways of introducing the error terms into equations (28).**
- **Our approach will not be very rigorous from an econometric point of view; we will simply generate different indexes as solutions to various least squares minimization problems.**
- **Our first approach is to simply add error terms, e_{tn}, to the right hand sides of equations (28). The unknown parameters, $\pi \equiv [\pi_1,\ldots,\pi_T]$ and $\alpha \equiv [\alpha_1,\ldots,\alpha_N]$, will be estimated as solutions to the following (nonlinear) least squares minimization problem:**

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

- **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.**

Time Product Dummy Regressions: Approach 1

- **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) $\alpha_n = \Sigma_{t=1}^{T} \pi_t p_{tn} / \Sigma_{t=1}^{T} \pi_t^2$; $\quad$ n = 1,...,N

$= \Sigma_{t=1}^{T} \pi_t^2 (p_{tn}/\pi_t) / \Sigma_{t=1}^{T} \pi_t^2$;

(31) $\pi_t = \Sigma_{n=1}^{N} \alpha_n p_{tn} / \Sigma_{t=1}^{T} \alpha_n^2$; $\quad$ 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.**
- **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 price levels that is generated by the least squares minimization problem defined by (29).**

Time Product Dummy Regressions: Approach 1 (cont)

- 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 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.

Time Product Dummy Regressions: Approach 1 Rejected

- **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.**
- **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.**

Time Product Dummy Regressions: Approach 2

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

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

Time Product Dummy Regressions: Approach 2 (cont)

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

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

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

- **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) $\pi_t = \Pi_{n=1}^{N} (p_{tn}/\alpha_n)^{1/N}$; $\quad t = 1,\ldots,T;$

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

- **Thus the period t aggregate price level, π_t, is equal to the geometric average of the N quality adjusted prices for period t, $p_{t1}/\alpha_1, \ldots, p_{tN}/\alpha_N$, 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, $p_{1n}/\pi_1, \ldots, p_{Tn}/\pi_T$.**
- **These estimators look very reasonable (if quantity weights are not available).**

Time Product Dummy Regressions: Approach 2 (cont)

- If $\pi^* \equiv [\pi_1^*,\ldots,\pi_T^*]$ and $\alpha^* \equiv [\alpha_1^*,\ldots,\alpha_N^*]$ 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 $1,\pi_2^*,\ldots,\pi_T^*$ is the sequence of price levels that is generated by the least squares minimization problem defined by (33).
- Once we have the unique solution $1,\pi_2^*,\ldots,\pi_T^*$ 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:

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

$$= \Pi_{n=1}^N (p_{tn}/p_{sn})^{1/N}.$$

- This is simply the Jevons index for period t relative to period s.

Time Product Dummy Regressions: Approach 2 (conc)

- **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}).**
- **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 given the same weight in the Jevons measure of inflation between the two periods as is given to high expenditure products.**
- **This result indicates the importance of weighting.**
- **In the next section, I look at this second approach when there are missing observations.**

6. Time Product Dummy Regressions: The Case of Missing Observations with no Weighting by Economic Importance

- 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:

(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.

6. Time Product Dummy Regressions with Missing Observations

- **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.**
- **The first order necessary conditions for $\rho_1,\ldots,\rho_T$ and $\beta_1,\ldots,\beta_N$ to solve (42) are the following counterparts to (34) and (35):**

(43) $\Sigma_{n\in S(t)} [\rho_t + \beta_n] = \Sigma_{n\in S(t)} \ln p_{tn}$; $\quad t = 1,\ldots,T;$

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

- **Let $\rho_t \equiv \ln\pi_t$ for $t = 1,\ldots,T$ and let $\beta_n \equiv \ln\alpha_n$ for $n = 1,\ldots,N$.**
- **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:**

6. Time Product Dummy Regressions with Missing Observations

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

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

- 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 $t = 2,3,...,T$.
- We have the following expressions for π_t^*/π_r^* and α_n^*/α_m^*:

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

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

- Now the α_n^*enter into the indexes π_t^*/π_r^* defined by (47).

6. Time Product Dummy Regressions with Missing Observations

- **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^*.**
- **We turn to hedonic models which involve weighting.**

7. Weighted Time Product Dummy Regressions: The Bilateral Case: Approach 1

- **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 $\rho_1 = 0$.**
- **But this weighted model generates the following solution for ρ_2**

(54) $\rho_2^* \equiv \Sigma_{n=1}^{N} q_{1n}q_{2n}(q_{1n} + q_{2n})^{-1} \ln(p_{2n}/p_{1n})/\Sigma_{i=1}^{N} q_{1i}q_{2i}(q_{1i} + q_{2i})^{-1}$.

- **The resulting $\pi_2^* = \exp[\rho_2^*]$ is not invariant to changes in the units of measurement; this index is not satisfactory!**
- **Now replace the quantity weights in (51) with value weights which leads to Approach 2:**

7. Weighted Time Product Dummy Regressions: The Bilateral Case: Approach 2

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

7. Weighted Time Product Dummy Regressions: The Bilateral Case: Approach 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.**
- **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.**
- **Weighting Approach 2 also fails.**

7. Weighted Time Product Dummy Regressions: The Bilateral Case: Approach 3

- **The above deficiencies with Approach 2 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:**
- **(61) $\min_{\rho, \beta} \Sigma_{n=1}^{N} s_{1n}[\ln p_{1n} - \beta_n]^2 + \Sigma_{n=1}^{N} s_{2n}[\ln p_{2n} - \rho_2 - \beta_n]^2$.**
- **The solution to (61) is given by:**

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

- **where $\pi_2^* \equiv \exp[\rho_2^*]$.**
- **This index has satisfactory properties. It will tend to give a somewhat lower index value π_2^* than the superlative Törnqvist-Theil index which it approximates to the second order.**

7. Weighted Time Product Dummy Regressions: The Bilateral Case: Approach 3 Concluded

- **From equation (38), 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.***

7. Weighted Time Product Dummy Regressions: The Bilateral Case: Approach 4

- **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:**

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

- **The solution to (65) simplifies to the following solution:**

$$(66)\ \rho_2^* \equiv \Sigma_{n=1}^{N} (1/2)(s_{1n}+s_{2n})\ln(p_{2n}/p_{1n});$$

$$(67)\ \beta_n^* \equiv (1/2)\ln(p_{1n}) + (1/2)\ln(p_{2n}/\pi_2^*)\ ; \qquad n = 1,\ldots,N$$

- **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 $\pi_2^*/\pi_1^* = P_T(p^1,p^2,q^1,q^2)$, which is the Törnqvist (1936) Theil (1967; 137-138) index.**

7. Weighted Time Product Dummy Regressions: The Bilateral Case: Approach 4

- **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).**
- **Note that all of the models in this section 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. But the indexes that result from alternative forms of weighting can be very different.**
- **This result reinforces the case for using appropriately weighted versions of the basic time product dummy hedonic regression model.**
- **Similar comments apply to more general hedonic regressions that use information on characteristics: in general, it is preferable to use share weights in these regressions.**

Additional Materials

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