Estimating the Benefits of New Products

W. Erwin Diewert
University of British Columbia and NBER

Robert C. Feenstra
University of California, Davis, and NBER

Conference on Research on Income and Wealth, *Big Data for 21st Century Economic Statistics*, March 16, 2019

How to incorporate new products into price indexes, to consistently reflect the gains from their appearance?

General answer due to Hicks (1940): Before products are available, use their reservation prices

Hausman (1999): Gains from appearance of cell phones

Suggested taking a *linear approximation to the demand curve* at the point of consumption, and computing the consumer surplus gain to a new product under this linear demand curve. Provided that the demand curve is convex, then this linear approximation will be a *lower bound* to the consumer surplus gain.

Question: is a method like this feasible when there are many new (and disappearing) products, as typically occurs with barcode data?

Answer: It is necessary to adopt a utility function consistent with an underlying consumer

We will explore *two utility functions* here:

1) Constant Elasticity of Substitution (CES)
2) Quadratic Utility Function, also used by Konüs and Byushgens (1926) and Fisher (1922), and so we will also call it the *KBF functional form*

Each for these function forms has their limitations:

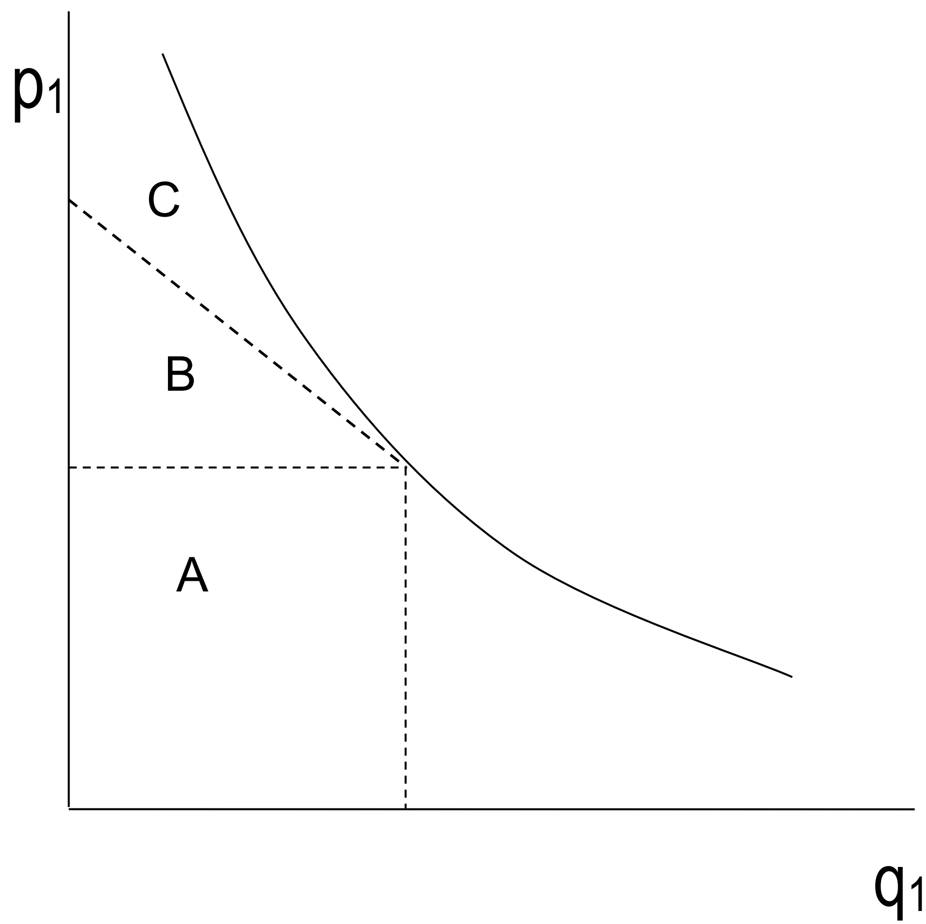
1) For CES, the reservation price is infinite, which may seem too high. But the gains from new products are still finite when $\sigma > 1$ (Feenstra, 1994)

2) For the KBF, there are too many parameters to estimate, so we restrict the number of parameters by extending the approach of the Normalized Quadratic functional form (Diewert and Wales, 1987, 1988).

We will construct exact price indexes for these two functional forms using a sample of frozen juice products sold in one store, and compare their results.

Constant-Elasticity Demand & Linear Approximation

Constant-elasticity demand $q_1 = kp_1^{-\sigma}$, with $k > 0$, $\sigma > 1$



Consumer surplus relative to expenditure,

$$\frac{B+C}{E_t} = \frac{1}{E_t}\int_{p_{1t}}^{\infty} kp^{-\sigma}dp = \frac{s_{1t}}{(\sigma-1)}, \quad \sigma > 1,$$

Where s_{1t} is the share of spending on good 1. Hausman's recommendation is to the consumer surplus area B,

$$\frac{B}{E_t} = \frac{s_{1t}}{2\sigma},$$

Comparing these equations, the ratio of the consumer surplus from the linear approximation to that from the constant-elasticity demand curve is *less than one-half,*

$$B/(B+C) = (\sigma - 1)/2\sigma < 1/2.$$

Recommendation: *Under weak conditions, taking* ***one-half*** *of the CES gain from new goods will be above but reasonably close to the Hausman linear approximation method, provided that the elasticity of demand is high.*

Consumer Gains with Share= 0.1 (% Expenditure)

σ	B/E_t	$(B+C)/2E_t$	$G_{CES}/2$	$G_H\vert_U$	G_H
2	2.50	5.00	5.56	2.78	2.70
3	1.67	2.50	2.70	1.85	1.82
4	1.25	1.67	1.79	1.39	1.37
5	1.00	1.25	1.33	1.11	1.10
6	0.83	1.00	1.06	0.93	0.92
10	0.50	0.56	0.59	0.56	0.55

CES Utility Function

$$U(q_t, I_t) = \left[\sum_{i \in I_t} a_i q_{it}^{(\sigma-1)/\sigma} \right]^{\sigma/(\sigma-1)}, \quad \sigma > 1,$$

The unit-expenditure function is:

$$e(p_t, I_t) = \left[\sum_{i \in I_t} b_i p_{it}^{1-\sigma} \right]^{1/(1-\sigma)}, \quad \sigma > 1, \; b_i \equiv a_i^{\sigma},$$

Hicksian demand q_{it} for that good,

$$q_{it}(p_t, U_t) = U_t \left[\sum\nolimits_{i \in I_t} b_i p_{it}^{1-\sigma} \right]^{\frac{\sigma}{1-\sigma}} b_i p_{it}^{-\sigma}, \qquad i \in I_t.$$

The (positive) Hicksian own-price elasticity

$$\eta_{it}\Big|_U \equiv -\frac{\partial \ln q_{it}}{\partial \ln p_{it}}\Bigg|_U = \sigma(1 - s_{it}).$$

Gains from new products (Feenstra, 1994):

$$\frac{e(p_t, I_t)}{e(p_{t-1}, I_{t-1})} = P_{SV}(I)\left(\frac{\lambda_t}{\lambda_{t-1}}\right)^{1/(\sigma-1)} .$$

For $I \equiv I_t \cap I_{t-1} \neq \varnothing$, "common" or maximum overlap items

$$P_{SV}(I) \equiv \prod_{i \in I}\left(\frac{p_{it}}{p_{it-1}}\right)^{w_i(I)},$$

$$w_i(I) \equiv \left(\frac{s_{it}(I) - s_{it-1}(I)}{\ln s_{it}(I) - \ln s_{it-1}(I)}\right) \Bigg/ \sum_{n \in I}\left(\frac{s_{nt}(I) - s_{nt-1}(I)}{\ln s_{nt}(I) - \ln s_{nt-1}(I)}\right).$$

And, $$\lambda_\tau \equiv \frac{\sum_{n\in I} p_{n\tau} q_{n\tau}}{\sum_{n\in I_\tau} p_{n\tau} q_{n\tau}} \leq 1,$$

$\lambda_\tau(I)$ is the *period τ expenditure on the goods in the common set I, relative to the period τ total expenditure.*

Alternatively, $\lambda_t(I)$ is interpreted as *one minus the period t expenditure on new or disappearing (not in the set I), relative to the period t total expenditure,*

When there is a greater expenditure share on new goods in period t than on disappearing goods in period t-1, then the ratio $\lambda_t(I)/\lambda_{t-1}(I) \leq 1$, a *fall* in the exact price index

With a single new product (a compensating variation):

$$G_{CES} = \left(1 - s_{1t}\right)^{-1/(\sigma-1)} - 1 \geq \frac{s_{1t}}{(\sigma - 1)}$$

Consumer Gains with Share= 0.1 (% Expenditure)

σ	B/E_t	$(B+C)/2E_t$	$G_{CES}/2$	$G_H\|_U$	G_H
2	2.50	5.00	5.56	2.78	2.70
3	1.67	2.50	2.70	1.85	1.82
4	1.25	1.67	1.79	1.39	1.37
5	1.00	1.25	1.33	1.11	1.10
6	0.83	1.00	1.06	0.93	0.92
10	0.50	0.56	0.59	0.56	0.55

Konüs-Byushgens-Fisher (KBF) Utility Function

$$U = (q^T A q)^{1/2},$$

where $A^T = A$ has $N(N+1)/2$ unknown a_{ik} elements.

Unit-expenditure function,

$$e(p_t) = (p_t^T A^* p_t)^{1/2},$$

provided $A^* = A^{-1}$ if A is of full rank. Hicksian demand is

$$q_{it}(p_t, U_t) = U_t \left[\frac{\sum_{n=1}^{N} a_{in}^* p_{nt}}{(p_t^T A^* p_t)^{1/2}} \right],$$

The (positive) Hicksian own-price elasticity,

$$\eta_{it}\Big|_U \equiv -\frac{\partial \ln q_{it}}{\partial \ln p_{it}}\Bigg|_U = \frac{-a^*_{ii}p_{it}}{\sum_{n=1}^{N} a^*_{in}p_{nt}} + s_{it},$$

Result: KBF demand curves have finite reservation price; they are convex; lie *in-between* the constant-elasticity CES demand curves and the straight-line Hausman approximation (all with same slope at consumption point)

Problem: What if A is not of full rank?

Then we can still work with the inverse demand curves:

$$p_t = E_t A q_t / (q_t^T A q_t) \ .$$

Furthermore, we can still derive a ***constant-utility Hicksian inverse demand curve*****,** and we can derive the Hausman approximation using the linear approximation to the linear approximation to this inverse demand curve.

IF the KBF and the CES demand curves have the same estimated elasticity at the consumption point, then the CES gain should be *less than twice as large* as the KBF.

(But we will see that the CES gain are actually 5x larger, meaning that the elasticities are not the same.)

Estimation on Frozen Juice Products at Barcode level

Weekly data aggregated to four-week periods that we call "months". There are 39 months and 19 products, or 741 obs. in total. There are 20 observations for products 2, 4 and 12 for the months when these products are missing.

Estimation of the CES Demand with Error in Prices

Measurement error in prices that is *also* reflected in the expenditure shares

$$s_{it} \equiv \frac{p_{it}q_{it}}{E_t} = \frac{a_i q_{it}^{(\sigma-1)/\sigma}}{\sum_{n \in I_t} a_n q_{nt}^{(\sigma-1)/\sigma}}$$

Take logs and add error terms u_{it} reflect the measurement error in shares,

$$\ln s_{it} = \ln a_i + \frac{(\sigma-1)}{\sigma}\ln q_{it} - \sum_{n=1}^{N} a_n q_{nt}^{(\sigma-1)/\sigma} + u_{it},$$

Feenstra *double-differenced variables* are defined as:

$$\Delta \ln s_{it} \equiv \ln(s_{it}) - \ln(s_{it-1}) \ ,$$

Now pick product N as the numeraire product and difference the $\Delta \ln s_{it}$ with respect to product N:

$$\mathrm{D}\ln s_{it} \equiv \Delta \ln s_{it} - \Delta \ln s_{Nt} \ ,$$

$$= \ln(s_{nt}) - \ln(s_{nt-1}) - \ln(s_{Nt}) - \ln(s_{Nt-1}).$$

Likewise for the *double-differenced log quantity and error term variables*:

$$\mathrm{Dln}q_{it} \equiv \Delta \ln q_{it} - \Delta \ln q_{Nt} \; ;$$

$$= \ln(q_{nit}) - \ln(q_{it\text{-}1}) - \ln(q_{Nt}) - \ln(q_{Nt\text{-}1}).$$

$$\mathrm{D}u_{it} \equiv u_{it} - u_{it\text{-}1} - u_{Nt} + u_{Nt\text{-}1} \; ,$$

Then the **CES demand equation** simplifies as

$$\mathrm{Dln}s_{it} = \frac{(\sigma-1)}{\sigma}\mathrm{Dln}q_{it} + \mathrm{D}u_{it} \; ,$$

The OLS estimated σ is equal to 7.40.

Can extend to errors in prices and (symmetrically) in quantities, along with (independent) measurement error in shares due to random taste parameters, resulting in $\sigma = 8.0$ for weekly, and $\sigma = 6.0$ from monthly data.

CES Gains with $\sigma = 7.40$

Month	Availability	$(\lambda_t / \lambda_{t-1})^{1/(\sigma-1)}$	G_{CES}
9	2 and 4 new	0.9928	1.0073
10	12 disappears	1.0036	0.9964
11	12 reappears	0.9957	1.0043
20	12 disappears	1.0039	0.9962
23	12 reappears	0.9969	1.0031
Cumulative Gain		0.9928	1.0073

In month 9, products 2 and 4 make their appearance, lowering the price level and increasing utility for month 9 by 0.73 percentage points. After that, product 12 disappears and re-appears several times (cancels out)

Estimation of the KBF Utility Function

The quadratic or KBF utility function

$$U = (q^T A q)^{1/2},$$

A is symmetric so there are N(N-1)/2 free parameters. That is too many to estimate, so we rewrite A matrix as

$$A = bb^T + B;\ b >> 0_N\ ;\ B = -CC^T$$

where C is a lower triangular matrix, so that B is negative semidefinite; $Bq^* = 0_N$,

Then we estimate the parameters of A in stages.

$$\hat{A} \equiv \hat{b}\hat{b}^T - \hat{c}^1\hat{c}^{1T} - \hat{c}^2\hat{c}^{2T} - \hat{c}^3\hat{c}^{3T} - \hat{c}^4\hat{c}^{4T} - \hat{c}^5\hat{c}^{5T} - \hat{c}^6\hat{c}^{6T}$$

This is the same type of procedure that Diewert and Wales (1988) used to estimate normalized quadratic preferences, resulting in a *semiflexible* functional form.

Estimation of KBF System Using Share Equations

$$s_{it} \equiv \frac{p_{it}q_{it}}{p_t \cdot q_t} = \frac{q_{it}\sum_{n=1}^{N} a_{in}q_{nt}}{q_t^T A q_t} + \varepsilon_{it}$$

We drop the months when the products are no available. The *predicted price* for product i in month t is:

$$p_t^* = E_t \hat{A} q_t / (q_t^T \hat{A} q_t)$$

This approach does not work very well: even though the predicted expenditure shares are quite close to the actual expenditure shares, *the predicted prices are not close to the actual prices* (so we do not think that the reservation prices for missing products are accurate)

Estimation of KBF Preferences Using Price Equations

We next estimate using *prices* as the dependent variables,

$$p_t = E_t A q_t/(q_t^T A q_t) + \varepsilon_t$$

Again take predicted price:

$$p_t^* = E_t \hat{A} q_t/(q_t^T \hat{A} q_t)$$

The predicted prices fit much better, so we believe that the reservation prices are also more accurate.

Gains for the KBF system, using Fisher Ideal Indexes:

CES: $$\frac{e(p_t, I_t)}{e(p_{t-1}, I_{t-1})} = \underbrace{P_{SV}(I)}_{\text{Maximum overlap index}} \underbrace{\left(\lambda_t / \lambda_{t-1}\right)^{1/(\sigma-1)}}_{\text{Inverse Gain}}.$$

KBF: $$\underbrace{P_{FH}(t-1,t)}_{\text{Using reservation prices}} = \underbrace{P_{FM}(t-1,t)}_{\text{Maximum overlap index}} / \underbrace{G_{KBF}}_{\text{Gain}}.$$

Or KBF*: $$\underbrace{P^{*}_{FH}(t-1,t)}_{\text{Using \textbf{all} reservation prices}} = \underbrace{P_{FM}(t-1,t)}_{\text{Max. overlap index}} / \underbrace{G^{*}_{KBF}}_{\text{Gain}}.$$

Reservation Prices for Unavailable Products or Using Predicted Prices for All Products

Month	**Availability**	$\mathbf{G_{KBF}}$	$\mathbf{G^{*}_{KBF}}$
9	2 and 4 new	1.0004	1.0016
10	12 disappears	0.9965	0.9988
11	12 reappears	1.0025	1.0015
20	12 disappears	0.9998	0.9971
23	12 reappears	0.9991	1.0001
Cumulative Gain		0.9983	0.9991

Does not work, as it leads to negative gains in total,
Problem seems to be with the *maximum overlap* index.

Second Approach: Combine λ_t term (like CES) with calculations for the estimated utility function, to measure the welfare gain/loss from new/disappearing goods

Month	**Availability**	$G_{A,KBF}$ $L_{A,KBF}$	G_{CES}
9	2 and 4 new	1.0013	1.0073
10	12 disappears	0.9975	0.9964
11	12 reappears	1.0030	1.0043
20	12 disappears	0.9988	0.9962
23	12 reappears	1.0008	1.0031
Cumulative Gain		1.0014	1.0073

CES gains are *five times larger* that the KBF gains.

Conclusions

Taking one-half of the gains from CES preferences will be close that from a linear approximation

KBF preferences have demand curves that lie in-between the CES and the linear approximation *if the elasticities of demand are the same,* so the gains from new products will greater than one-half the CES

Since the CES gains are 5x larger, we conclude that KBF has *more elastic* demand for new products 2 and 4

This illustrates the tradeoff: between CES (easy to estimate, but constrains the elasticity) and the KBF (harder to estimate, but it allows the new product to have a different and possibly lower elasticity than CES).