An Index Number Formula Problem: the Aggregation of Broadly Comparable Items

Mick Silver*
International Monetary Fund

Presentation to the (Ottawa) International Working Group on Price Indices (May 27–29, 2009) Swiss Federal Statistical Office (FSO), Neuchâtel, Switzerland.

The problem with unit value indices is that we hear a lot about their use for trade price indices and little about their use elsewhere, when in fact we should be hearing much less about their use for trade price indices and a lot more elsewhere.

*The views expressed herein are those of the author and should not be attributed to the IMF, its Executive Board, or its management.

On formula and product heterogeneity

If data on matched prices and quantities are available:

☐ a superlative price index number formula such as Fisher is best to aggregate *heterogeneous* items, and a unit value index is biased, and

☐ a unit value index is best to aggregate *homogeneous* ones, and a superlative index is biased.

SNA 1993 and *2008*; *CPI* and *PPI Manuals*; Balk, Diewert and others.

☐ What about broadly comparable ones?

An illustration

		period 0				period 1		
	price	quantity	value		price	quantity	value	
A	10	6	60		10	8	80	
B	12	6	72		12	4	48	
total		12	132			12	128	
		Unit value				0.97		
		Laspeyres/Fisher/Paasche				1		

When to use unit value indexes and when Fisher? – potential areas of application

- ☐ CPIs: same item different outlet: Reinsdorf; Hausman and Leibtag (2008).
- ☐ Health treatments: Aizcorbe and Nestoriak (2008).
- ☐ PPI: same output (input) different customers (suppliers).
- ☐ XMPIs: same output (input) different country customers (suppliers).

SNA 2008 notes exception of institutionalized price discrimination to using unit values for homogeneous case. Asymmetry in theory for imports and exports.

Unit value and Fisher price index

$$P_{UV} = \left(\frac{\sum\limits_{m=1}^{M} p_m^t q_m^t}{\sum\limits_{m=1}^{M} q_m^t} \right) \Bigg/ \left(\frac{\sum\limits_{n=1}^{N} p_n^0 q_n^0}{\sum\limits_{n=1}^{N} q_n^0} \right)$$

$$P_F \equiv \sqrt{ \frac{\sum\limits_{i=1}^{n} p_i^t q_i^t}{\sum\limits_{i=1}^{n} p_i^0 q_i^t} \times \frac{\sum\limits_{i=1}^{n} p_i^t q_i^0}{\sum\limits_{i=1}^{n} p_i^0 q_i^0} }$$

Unit value indices for heterogeneous items:

- ❑ **The test approach** (Balk and Diewert)
  - ■ The unit value index fails the *Proportionality Test*: that is, if all prices are multiplied by the positive number λ, then the new price index is λ.
  - ■ The unit value index fails the *Invariance to Changes in the Units of Measurement (commensurability) Test*: that is, the price index does not change if the units of measurement for each product are changed.
  - ■ Inappropriate for XMPIs using customs data (Silver)
- ❑ Economic theoretic approach
  - ■ Bradley (2005) compares the bias that results from using unit values as "plug-ins" for prices for a COLI. Only in the uninteresting case of no price dispersion in either the current or reference period will the unit value (plug-in) index be unbiased against the COLI.

Unit value indices for homogeneous items:

- ☐ Passes the time aggregation problem.
- ☐ If unit value index is used to deflate a corresponding value change, the result is a change in total quantity which is intuitively appropriate,

$$\frac{\sum\limits_{m=1}^{M} p_m^t q_m^t}{\sum\limits_{n=1}^{N} p_n^0 q_n^0} \Bigg/ \left[\left(\frac{\sum\limits_{m=1}^{M} p_m^t q_m^t}{\sum\limits_{m=1}^{M} q_m^t} \right) \Bigg/ \left(\frac{\sum\limits_{m=1}^{M} p_m^0 q_m^0}{\sum\limits_{m=1}^{M} q_m^0} \right) \right] = \frac{\sum\limits_{m=1}^{M} q_m^t}{NM \sum\limits_{m=1}^{M} q_m^0}$$

- ☐ Tests not designed for homogeneous items.

When to use unit value indexes and when Fisher?

- ☐ Balk: if the splitting into homogeneous and heterogeneous items was not feasible, he advised a price index.

- ☐ Diewert: if detailed data on strictly homogeneous goods unavailable (specified item by outlet for CPI), then unit value over outlets.

- ☐ Dálen argued for quality-adjusted unit value indexes that remove the effect on prices of product heterogeneity and de Haan implemented this in a hedonic setting.

Numerical relationship: unit value and Fisher price index

- Balk (1998) and Parniczky (1974) seminal work: decomposition in terms of quantity-weighted covariances.

$$\frac{P_{UV}}{P_F} = \frac{P_{UV}}{P_L} \times \frac{P_L}{P_F} = \frac{P_{UV}}{P_L} \times \left[\frac{P_L}{P_P}\right]^{\frac{1}{2}}$$

$$\frac{P_F}{P_L} = \frac{P_P}{P_F} = \left[\frac{P_P}{P_L}\right]^{\frac{1}{2}} = \left[\rho_{x,y}^{s_0} cv^{s_0}(x) cv^{s_0}(y) + 1\right]^{\frac{1}{2}}$$

Numerical relationship: unit value and Fisher price index

$$\frac{P_{UV}}{P_P} = \frac{\overline{q}^0}{\overline{q}^t} \frac{\hat{\beta}^t}{\hat{\beta}^0}$$

$$\frac{P_{UV}}{P_L} = \left(\rho_{x,y}^{s_0} cv(x)cv(y) + 1 \right) \left(\frac{\overline{q}^0}{\overline{q}^t} \frac{\hat{\beta}^t}{\hat{\beta}^0} \right).$$

$$\frac{P_{UV}}{P_F} = \frac{P_{UV}}{P_L} \times \frac{P_L}{P_F} = = \left(\rho_{x,y}^{s_0} cv^{s_0}(x)cv^{s_0}(y) + 1 \right)^{\frac{1}{2}} \left(\frac{\overline{q}^0}{\overline{q}^t} \frac{\hat{\beta}^t}{\hat{\beta}^0} \right)$$

The difference between a Fisher price and a unit value index

☐ The first term is the **substitution effect.**

☐ The second term is the **levels effect**, that is the effect, for negatively sloping demand, of quantities shifting to prices at a lower level. The measure is based on the ratio of the slope coefficients from the regressions of $q^t{}_m$ on $p^0{}_m$ and of $q^0{}_m$ also on $p^0{}_m$. As the slope of the period t line say increases, above average prices have lower quantities and below average prices have higher quantities—the larger the increase, the greater the shift. The change in slopes capture a shift in levels.

Figure 1, Depiction of levels effect
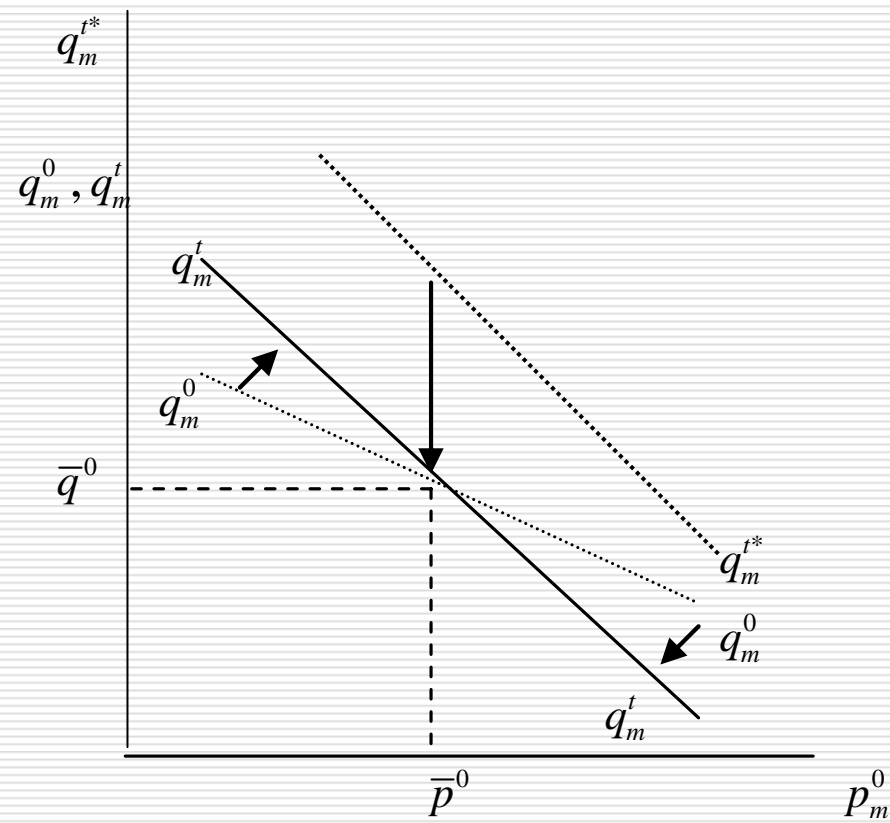
Decomposition helps

- ☐ Identifies role of substitution effect;
- ☐ the unit value bias will be equal to zero if:
  - ■ all base period price OR quantity changes are equal to each other OR there is no (weighted) correlation between the base period price and quantity changes;

  AND

  - ■ all base period prices OR base and current period quantities are equal to each other OR there is no (unweighted) correlation between the base period prices and base and current period quantities;
- ☐ relative position of UV, Las, Pas, and Fisher;
- ☐ identifies levels effect;
- ☐ product heterogeneity via CVs.

When to use unit value indexes and when Fisher?

- □ Balk: if the splitting into homogeneous and heterogeneous items was not feasible, he advised a price index.

- □ Diewert: if detailed data on strictly homogeneous goods unavailable (specified item by outlet for CPI), then unit value over outlets.

- □ Dálen argued for quality-adjusted unit value indexes that remove the effect on prices of product heterogeneity and de Haan implemented this in a hedonic setting.

If unit value indices are right for homogeneous items, what about broadly comparable goods and services?

- What about a quality-stripped unit value index? Dálen (2001); De Haan (2004 and 2007)?

- Consider a regression of price on *k* quality characteristics:

$$p_m^\tau = \beta_0^\tau + \sum_{k=1}^{K} \beta_k^\tau z_{km}^\tau + u_m^\tau$$

$$\hat{p}_m^\tau = p_m^\tau - \left[\sum_{k=1}^{K} \beta_k^\tau \left(z_{km}^\tau - \bar{z}_{km}^\tau \right) \right]$$

$$\tilde{p}_m^\tau = \beta_0^\tau + \sum_{k=1}^{K} \beta_k^\tau z_{km}^\tau$$

$$P_U^* = \left(\frac{\sum_{m=1}^{M} \hat{p}_m^t q_m^t}{\sum_{m=1}^{M} q_m^t} \right) \Bigg/ \left(\frac{\sum_{m=1}^{M} \hat{p}_n^0 q_n^0}{\sum_{m=1}^{M} q_n^0} \right)$$

But for broadly comparable goods some of the price change is a unit value shift in levels and some not…

- ☐ Weighted average of Fisher and quality-adjusted Fisher
- ☐ Weights based on:
  - ■ elasticity of substitution
  - ■ proportion of price variation due to heterogeneity: RSS

$$P_U^* \overline{w}_U + P_F^*(1-\overline{w}_U) = \frac{\sum_{m=1}^{M} \hat{p}_m^t q_m^t \Big/ \sum_{m=1}^{M} q_m^t}{\sum_{m=1}^{M} \hat{p}_m^0 q_m^0 \Big/ \sum_{m=1}^{M} q_m^0} \times \overline{w}_U + \sqrt{\frac{\sum_{m=1}^{M} \tilde{p}_m^t q_m^t}{\sum_{m=1}^{M} \tilde{p}_m^0 q_m^t} \times \frac{\sum_{m=1}^{M} \tilde{p}_m^t q_m^0}{\sum_{m=1}^{M} \tilde{p}_m^0 q_m^0}} \times (1-\overline{w}_U)$$

An appropriate index..

- □ *should have the property that if all price variation is explained by the hedonic regression, the index is a Fisher index; if none of the price variation is explained by the hedonic regression, the index is a unit value index; as the percentage of price variation explained by the hedonic regression increases, so too will the weight given to the Fisher component.*

Empirical work

- Useful work by de Haan on unit values and quality-adjusted ones: TVs, washing machines, refrigerators and PCs – for matched and unmatched, OLS and WLS. For all products, except PCs, QAUV was less than weighted TDI, and for PCs pretty close.

- This work in progress is on television sets for matched models – simpler – it is clear that hedonic is about product heterogeneity only. The decomposition is to explain why formulas differ p in terms of signs on rho and magnitude of CVs and to examine different formulas.

Empirical work

Figure 1, Unit value and price indices for 14in TVs

Empirical work

Figure 2, Quality adjusted unit value and Fisher price indices
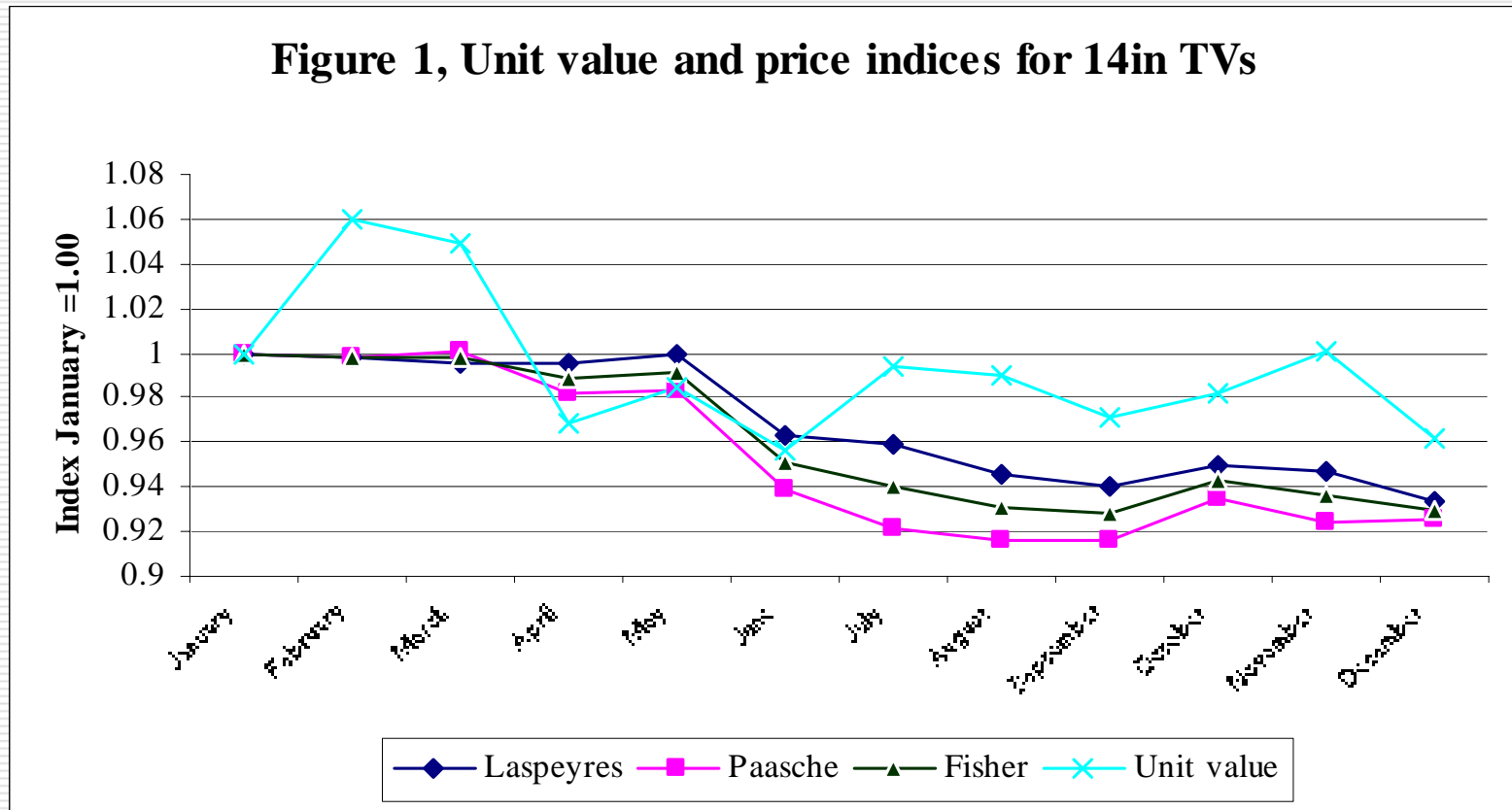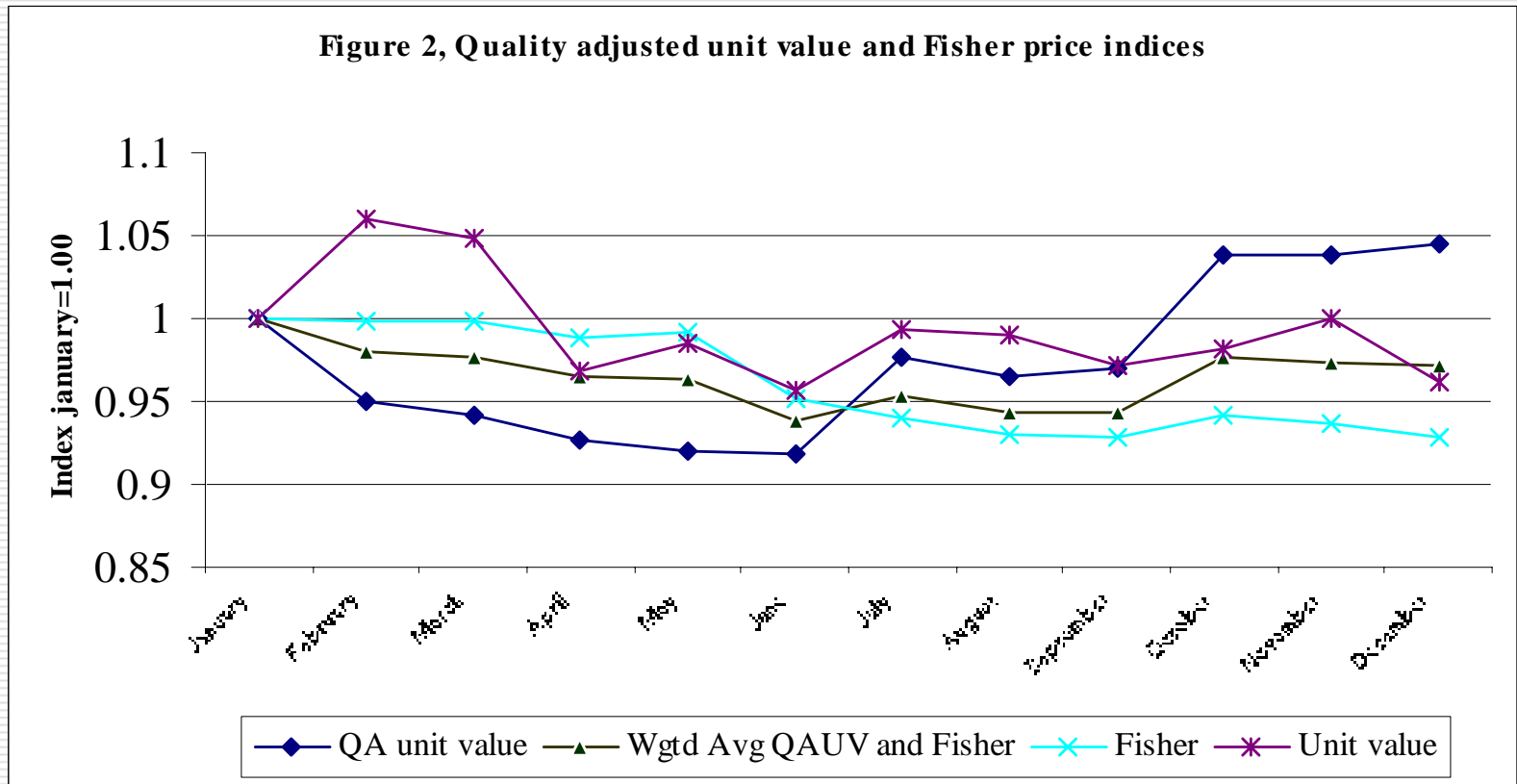
Summary

- For the aggregation of homogeneous items, the unit value index is the best index and superlative index numbers biased, and for the aggregation of heterogeneous items, superlative index numbers are best index and unit value index numbers biased.

- Institutionalized price discrimination is exception for buyer of homogeneous items, but not seller.

- The factors determining the difference between unit value indices and Laspeyres, Paasche and Fisher price indices were established.

- Quality adjustments to the prices to mitigate price dispersion due to the slight product heterogeneity would be appropriate for unit value indices.

- For broadly comparable goods and services, it is more complex.