WHEN DO HEDONIC AND MATCHED MODEL INDEXES GIVE DIFFERENT RESULTS? AND WHY?

Jack E. Triplett
Brookings Institution
Washington, D.C., USA

SSHRC Conference on Index Number Theory and the Measurement of Prices and Productivity

Vancouver, B.C.

June 30 - July 3, 2004

Note: This paper constitutes chapter IV in the author's forthcoming OECD publication:

Triplett, Jack E. Forthcoming. Handbook on hedonic indexes and quality adjustments in price indexes. Directorate for Science, Technology and Industry. Paris: Organization for Economic Co-operation and Development.

Revised: Jan 31, 03

Additional:Feb13,Mar18,correctMay21

Final: feb12, 04

CHAPTER IV

WHEN DO HEDONIC AND MATCHED MODEL INDEXES GIVE DIFFERENT RESULTS? AND WHY?

Many economists believe that hedonic indexes generally rise more slowly, or fall more rapidly, than matched model indexes. One also frequently hears the opposite, particularly from within statistical agencies: Hedonic indexes should give the same result as matched model indexes.

The question "Do they differ?" is too simple. Empirical studies show that matched model and hedonic indexes seldom coincide, they usually differ. I review the empirical work in section D of this chapter. More crucial are the "when?" and the "why?" parts of the questions in the chapter title. One needs to understand the circumstances under which matched model indexes and hedonic indexes will give the same result and the circumstances under which they will differ.

A portion of this topic—but only a portion—was addressed in Chapters II and III. The content of Chapters II and III was kept narrowly focussed on the *forced replacement* problem inside price index samples: Some items that were in the index sample disappear and their disappearance forces selection of a replacement. Forced replacements require the statistical agency to make suitable quality adjustments for the quality changes between replacement items and the old items they replace in the sample. Hedonic quality adjustments for forced replacements may yield different indexes from those produced with conventional quality adjustments, but this will depend (among other things) on which conventional quality adjustment method is used for the forced replacement. See the conclusions of Chapters II and III.

This forced replacement focus was preserved partly for expositional reasons, in order to highlight differences in the way conventional and hedonic indexes handle quality change. The framework used in Chapters II and III explicitly and purposefully set the rest of the index number context the same: the index databases (whether small or large sample or universe), sampling procedures, calculation methods, and collection strategies were held fixed. By abstracting from other issues that may also arise in constructing accurate price indexes for high tech products, the "other things equal" exposition in Chapters II and III directed attention to one set of issues on which there has been much confusion—the difference between hedonic and conventional indexes, particularly conventional and hedonic quality adjustments—without introducing elements that, though important, complicate the discussion.

However, a complete treatment goes beyond forced replacements in fixed samples (no matter how large). This chapter compares hedonic and matched model indexes in a broader context, in which the effects of entries and exits in product markets—not just exits and replacements in the sample—are considered.

A. INSIDE-THE-SAMPLE FORCED REPLACEMENTS AND OUTSIDE-THE-SAMPLE QUALITY CHANGE

The forced replacement problem addressed in Chapters II and III is perfectly general, it does not apply solely to price indexes that are based on small samples. It is indeed true that most statistical agency samples of high tech products normally include only a small part of the universe of those products. However, even if the price index were based on the initial period's universe of transactions for some product, the forced-replacement quality adjustment problem arises as items in the initial period's universe exit and are replaced by new product varieties in subsequent periods.

Numerous studies of personal computers have documented an extraordinary rate of model turnover: Pakes (2003) records annual sample attrition of 80 per cent in IDC data for U.S. PCs, and Van Mulligen (2002) reports that in a near universe of Dutch computer models sample attrition rates average nearly 20 percent *per month*. Koskim≅ki and Vartia (2001) and McKenzie (2002) present comparable sample attrition rates from two widely separated markets (Finland and Australia). Rapid model turnover is not confined to computers. Silver and Heravi (2002) examined the scanner data transactions universe for U.K. appliances; they show that exits reduce the original coverage by 20 percent in twelve months.[1] Moreover, when entrants were also ignored, the fixed sample's coverage declined to only 50 percent of expenditures by the end of the year.

Because exits and entries are the essence of the forced replacement problem, the analysis of Chapters II and III still applies to large samples. A number of computer price indexes have now been produced which use matched model methods on a large sample or near universe (see the review in section D of this chapter). Regardless of the size of the sample, one can still ask: Would the index have been different had hedonic quality adjustments been employed for forced replacements instead of matched model methods?

However, the impacts of quality change on price indexes go beyond adjusting the fixed sample for forced replacements. Forced replacements are caused by exits of old product varieties from the sample. Equally important are the price impacts of entering new product varieties, new varieties that may not be in the sample at all. What matters is not only how the agency (or the researcher) adjusts for inside-the-sample quality changes, but also whether the fixed-sample design systematically misses price change from rapid turnover of product varieties in high tech products. Size of the sample may be an issue. Of more importance are frequency of sample updating, weighting of the sample, principles for selecting the item samples, item replacement rules when an item exits from the sample, and—especially—the impact of new product entries on the pricing structure. I refer to this complex of measurement problems as "outside-the-sample" quality change problems, because they cannot be ameliorated in any way by improving the adjustment methods for quality changes encountered inside the sample, nor can they be understood in the context of the inside-the-sample (forced replacements) analysis.

In turn, one can think of the outside-the-sample problem as having two parts: First, do the relatively small fixed samples normally drawn for price indexes remain representative for multiple periods? The very rapid rates of sample deterioration already cited suggest they are not representative for high tech products.

Second, do large samples, or even universe samples that are replenished frequently, account for all price change when used in a matched model price index calculation? This second question is considerably more complicated. A matched model index that is based on a broad sample, or a universe, and that is updated frequently is sometimes called the "frequently resample and reweight" method

[1] To avoid misinterpretation, one should note that a 20 percent (or 80 percent) attrition rate does not mean that the statistical sample at the end of the period will be 20 percent (or 80 percent) smaller than in the initiation period, because agencies will normally replace exits.

(hereafter, FR&R).[2] Can an agency get equivalent results from an FR&R reform of the traditional matched model index methodology, without adopting hedonic methods? If it can, are there advantages in cost and data requirements of expanded, FR&R matched model indexes over hedonic indexes for high tech products? Conversely, if hedonic and FR&R matched model indexes differ, why do they and under what circumstances?

The answers to all these questions depend on the operation of markets—the way price changes come about, the way new products are introduced and priced, and the marketing strategies pursued by producers of high tech (and much low tech) equipment. The following sections develop these ideas.

B. FIXED SAMPLES AND PRICE CHANGES OUTSIDE THE SAMPLE

Statistical agencies typically draw product samples at some period and hold them fixed, or attempt to hold them fixed, over some interval. If probability samples, fixed samples may be drawn conceptually from what Dalϑn (2001) has called the "fixed universe" or they may have been drawn with the intention of representing Dalen's "dynamic universe" but without the statistical methodology to do it adequately. Even if originally representative, the longer the interval over which the sample remains fixed, the more likely that it will become unrepresentative of the universe of transactions.

In some countries' price indexes, product or item samples of high-tech products are judgemental. Items may be selected with an eye to reducing the incidence of future forced replacements—the rule is to find products that are likely to remain for sale, and less likely to disappear. When quality change takes the form of new product varieties, samples of varieties that do not change may be very unrepresentative right from the start. Trying to minimise the incidence of forced replacements generates another "quality problem" that is equally, and perhaps more, serious.

Though we speak of the "fixed sample," the sample never really remains fixed; replacement items are brought in as old varieties exit, and the index is linked over in some manner, or adjusted, as explained in Chapter II. The sample is "fixed" in the sense that a new sample is not drawn for some interval (which varies across countries and sometimes across products), and in the sense that entering computers have no probability of entering the sample, except as replacements for computers that exit the sample. The sample is not fixed in the sense that the products in it never change; but changes are forced, they are not changes by design.

Accordingly, replacement rules matter. Whether the sample was originally drawn on probability or judgemental principles, when forced replacements occur the agency may select a replacement item that is as close as possible to the one that exited the sample. The logic of this "nearest product variety" rule is to minimize the quality changes for which adjustments are required. This replacement rule has been called (by Walter Lane, of the U.S. BLS): "find the next most obsolete product." "Nearest product variety" replacement rules assure that the sample will increasingly be unrepresentative of high-tech product varieties that are for sale. Again, attempting to minimise one kind of quality error risks introducing another.

Whether fixed samples are drawn on a probability or judgemental basis, the price behaviour of the fixed sample will not adequately represent the price behaviour of the total market in the face of rapid technological change and the introduction of new product varieties. Replacing products that exit from the sample is not the same thing as bringing new products into the sample on a timely basis. Though much thinking about these topics has proceeded recently in conjunction with the HICP program in Europe, it is

[2] I have taken this term from European usage, there seems no equivalent in English language usage elsewhere. I am told that Eurostat has introduced a new term recently,

important not to overlook antecedents, for the lessons that can be learned from them are not the less valuable for being old lessons.

1. Research Studies

One of the first studies on the fixed sample quality change problem was Berndt, Griliches, and Rosett (1993). Working on a product completely unrelated to IT, they found that prices for newer pharmaceuticals fell, relative to older ones. Indeed, this study produced the arresting finding that drug manufacturers typically raised the prices of older branded prescription pharmaceuticals when new generic competition followed expiration of a patent.

At that time, the U.S. BLS selected pharmaceuticals for the PPI on a probability basis, but once they were selected, they stayed in the sample for five years, until another probability sample was drawn.[3] Using an exhaustive database of pharmaceutical prices obtained from manufacturers, Berndt, Griliches, and Rosett were able to mimic the U.S. PPI for pharmaceuticals. They then constructed a price index that brought new pharmaceuticals into the index as soon as they were introduced. Because prices of newer pharmaceuticals declined relative to older ones, the price index that mimicked the PPI sample rose considerably more than the universe of drug prices. Berndt, Griliches, and Rosett (1993) showed that holding the sample fixed in the face of technological innovation and new products resulted in a substantially upward-biased price index. Theirs was a very influential study, for it caused changes in PPI sampling procedures for pharmaceuticals.

A second pair of studies concerned semiconductors. Dulberger (1993) and Flamm (1993) documented the effect of delay in bringing new semiconductor chip products into the price index. In the PPI for semiconductors at that time, the BLS chose a probability sample of chips and held the sample fixed for approximately five years, except for forced replacements. The prices of newer types of semiconductors were declining more rapidly than older ones, similar to the case of pharmaceutical prices. Both Dulberger and Flamm presented semiconductor price indexes that differed substantially from the published U.S. PPI indexes for semiconductors. Table 4.1 summarises Dulberger's findings. When new semiconductor chips were brought rapidly into the price index, standard index number formulas exhibited price declines of 20-35 per cent per year (see the first four entries in Table 4.1). In contrast, when Dulberger mimicked the PPI procedures on the same data, so that new chips entered the index with a delay, the price index fell only about 8.5 per cent, which was closer to the PPI's 4.5 per cent decline (last two lines of Table 4.1).

None of these studies had anything to do with hedonic indexes—they were all computed with matched model methods. Essentially, they were all FR&R indexes, though sometimes equally weighted. They provided great insight into the question of constructing price indexes for products that were experiencing rapid technological change: They found that rapid quality change made fixed samples obsolete because prices of new product varieties, *after their introduction*, declined more rapidly than did prices of older products. The problems they uncovered did not originate in the treatment of inside the sample, forced replacements.

Silver and Heravi present similar results in a series of papers, though they use hedonic indexes to quantify their conclusions. As Silver and Heravi (2002) put it, two conditions must be met for price index bias from fixed samples: First, the samples must lose representativeness. Second, the price changes of the varieties that remain in the fixed sample must be systematically different from those that do not (the new and the exiting products). They show that the fixed sample for appliances in the U.K. deteriorates very rapidly; thus, the first condition is met. The second condition is also met, because price

[3] Forced replacements are not frequent in drug price index samples.

changes of entrants and exits from the sample are different from the price changes for products that continue. Substantial bias arose from holding appliance samples fixed for even a few months.

Moulton, Lafleur, and Moses (1999) compared matched model and hedonic indexes using the database collected for the CPI. They estimated two hedonic indexes for colour TV's: (a) a characteristics price index for TV's, using hedonic functions estimated from the CPI database, and (b) an index in which hedonic quality adjustments (using the same hedonic function) were made for forced replacements in the actual U.S. CPI TV index.[4]

The index with hedonic quality adjustments for forced replacements differed only insignificantly from the published CPI TV index (table 4.2). This implies that hedonic quality adjustments for forced replacements—about 15 percent of the price quotes—equalled in magnitude the actual CPI quality adjustments, which were mostly by the "class mean" variant of the IP-IQ method (explained in Chapter II).

The characteristics price index, however, differed substantially from the actual CPI, even though both were based on the same CPI database (table 4.2). The U.S. CPI sample is periodically replenished; in new samples, new TV models are selected with a probability proportionate to their sales. Moulton, Lafleur, and Moses reasoned that (in the language used in this handbook) the characteristics price index allowed for the price impacts of new introductions into the CPI sample, where CPI quality adjustments (hedonic or not) were made only for forced replacements for sample exits. When allowance was made for the entry of new product varieties, and not solely for forced replacements, the price index showed a steeper rate of decline.[5]

As these studies suggest, new varieties may exhibit price changes *after their introduction* that differ from the price changes of older varieties. In the usual economists' notion of the pricing cycle, prices of newer varieties fall more than the prices of continuing products, and for this reason the newer ones earn larger market shares.

New products might also be initially introduced at low (quality adjusted) prices, to induce consumers to try them. In this case, their subsequent prices rise, relative to those of continuing products, especially if they are successful introductions. Though this also implies an error from use of the fixed sample of established varieties, the direction of price index error in this case is not always obvious.[6]

Most price indexes are constructed around some notion of a representative sample of sales in the initiation period. A sample of initiation period sales may be adequate when little change occurs in the range of products that are for sale, when yesterday's products are pretty much the same as today's. Or the fixed sample may work fairly well when the prices of any new products that are introduced move more or less consistently with those of established products. In semiconductors, pharmaceuticals, and computers in the US, and in appliances in the UK (and possibly in the US), a price index that records only price

[4] See Chapter III for definitions of the characteristics price index and of hedonic quality adjustment.

[5] In the Moulton, Lafleur, and Moses (1999) study, it is not clear whether prices of new entrants into the CPI sample—which are not, after all, necessarily new entrants into the market—were declining relative to established TV models after the new models entered the sample, or had prices that were lower (quality adjusted) at entry. For the latter problem, see the next section of this chapter. Most other studies have shown that matched model indexes overstate the rate of decline for TVs; see section D of this chapter.

[6] Pakes (2003) contains an example where the price of the new product rises after introduction, yet, because the new products are better on a quality-adjusted basis than those that they displace, the price index should fall as their market share rises. Though economists have speculated about which behavior of new product prices—falling or rising prices after introduction—is the dominant one, few empirical studies exist to confirm which speculation is born out in actual markets.

movements in established products misses much of the price change that occurs. These studies all show that price movements from new introductions, *after they are introduced*, can be different from price movements measured by the continuing varieties (and item replacements) that are contained in a fixed sample.

2. Hedonic and FR&R Indexes

The new introductions bias discussed in the previous section can be ameliorated by FR&R methods, in which new products are brought promptly into the sample. Hedonic indexes are not strictly necessary. It is true, nonetheless, that hedonic index samples often approach universes of product varieties, and their samples are normally replenished for each period covered in the investigation. Accordingly, hedonic indexes share sample size, sample replenishment, and near universe coverage with FR&R index methods; perhaps these are the features of hedonic indexes that ameliorate the fixed sample problem. One might expect, at any rate, that when new models have price changes *after introduction* that differ from price changes for continuing models, FR&R and hedonic indexes will coincide because both incorporate these new models promptly.

FR&R methods, however, cannot confront another potential problem posed by new product introductions. If price change takes place *at the introduction* of new product varieties, or if price change is associated with product exits, price changes can be missed even with rapid replenishment of samples. This class of problems is addressed in the following section.

C. PRICE CHANGES OUTSIDE FR&R SAMPLES

New varieties may experience different price movements *after their introductions* from those of older varieties, as explained in the previous section. However, there is a second price effect from new varieties. Price changes may *accompany the introduction* of new products, or the exit of old ones. A "residual diagram," similar to the ones used in Chapter III, illustrates (Figure 4.1).

Suppose a new computer (model n) is introduced that is faster than the one it replaces (model m), but is cheaper, because it embodies a new technology. Statistical agencies frequently observe that replacements appear that are higher in quality but lower in price than the products they replace, which is the situation depicted in Figure 4.1. Because the old hedonic function (h_{old}) reflects the old technology, computer n's introductory price is also lower than the old hedonic function.

Disregarding for the moment what happens to the prices of continuing computers (the matched models), the replacement of computer m by computer n amounts to a price reduction—the spectrum of (quality-adjusted) prices is lower than it was before computer n's introduction (refer to the discussion of this point in Chapter III, section C.3). The hedonic function "residual" in Figure 4.1 shows the price difference between the actual price of the new computer (model n) and an old one with the same specification—the value "est ln P_n" in Figure 4.1 is the estimated price of a computer with computer n's specification, estimated from the old hedonic function. The hedonic function residual therefore indicates the downward pressure on the prices of competitive computers implied by the introduction of the new computer.

The residual also measures—in this particular case—the price decline of the new computer (model n), compared with the one that it replaced (model m). The actual price difference between these two computers is labelled V in Figure 4.1, but V under-estimates the true price decline because computer n is faster than computer m. The quality-adjusted price difference is the actual price difference (V) plus (in this case) the value of the quality difference, which is labelled A(h) in Figure 4.1. The sum of these two terms equals the residual, so the residual measures the quality adjusted price difference that should go into the price index.

In this chapter, we do not need the more complicated notation that was introduced in Chapter III. I thus revert to the simpler notation of Chapter II, and designate the "old" or previous period's hedonic function as period t, the "new" or current period's hedonic function as period t+1.

1. Case One

Suppose initially that no other price changed in response to the new computer. Any hedonic method would give a declining price index. Matched model indexes generally will not, whether fixed or FR&R samples.

Consider first the hedonic dummy variable method. Even if prices of other computers do not change in response, computer n's introductory price (below the old hedonic surface) will pull the hedonic function downward, by an amount that depends on the size of the residual in Figure 4.1 and computer n's market share. The new hedonic function might, for example, look like the hedonic function h (t+1) shown in Figure 4.2. The coefficient of the dummy variable in an adjacent-period regression measures the price change (see Chapter III); in Figure 4.2, the dummy variable index is shown as the downward shift of the hedonic function, from h (t) to h (t+1), marked Δh, in Figure 4.2.[7]

Now consider the hedonic imputation method. The hedonic index also would fall. The imputed price decline for the entering computer is measured by the residual in Figure 4.1, as noted above. This decline is shown as ΔP_n in Figure 4.2 (ΔP_n is the same as the residual in Figure 4.1). The index declines by ΔP_n, weighted by computer n's index weight.

One could also impute a price change for the exiting computer (which would be good practice). This yields a price decline equal to the amount ΔP_m, which is also shown in Figure 4.2. It might not seem intuitively obvious that the imputed price for the exiting computer falls, when no continuing computer experiences price change. In this example, computer m's price in period t lay on the old hedonic line, h(t). The decline occurs because the hedonic function for the new period (t+1) is used to impute the price of exits (see Chapter III). The hedonic function h(t+1) is below h(t), so the imputed price of exiting computer m in period t+1 is below its actual price in period t, yielding ΔP_m as the price decline for computer m.[8]

What about the FR&R method? Will it record a price decline? Generally it will not, though the complete answer depends on which method of quality adjustment is used in the FR&R index.

Most FR&R indexes will probably shift over to the new computer when it is encountered and drop the old one, linking together the index from the old sample to the index that uses the replenished one. This implies the deletion (IP-IQ) quality adjustment method (Chapter II), in which the price change for the exiting/entering computer is implicitly imputed from actual price changes of continuing ones.

In our example, none of the continuing computers changes in price. Hence, imputed price change for exiting or entering computers is zero. The FR&R method clearly misses price change, and it

[7] Entry of the new computer might also change the slope of the hedonic function, but this is neglected for simplicity in the diagram (see the parallel point in Chapter III, section YY). One needs to worry about the weight of computer n in the regression (in the usual equally weighted regression, computer n might have a weight that is too high or too low), but that matter is discussed in Chapter VI.

[8] A different period t price for computer m would obviously change this example. Consider the implication if P_{mt} were below the (new) regression line; in this case the exit of computer m implies a price increase.

differs from the hedonic index. The result for the FR&R sample is the same as from the fixed sample index, which was discussed in Chapters II and III. The FR&R sample does not surmount the problems of the fixed sample when prices of continuing computers do not respond to new entrants.

Some alternative quality adjustment method might be combined with the FR&R method. For example, the price index agency might use option prices to adjust for the quality difference between computer n and computer m, or it might apply a production cost adjustment, or an option price adjustment, or even use a hedonic quality adjustment. In these cases, some or all the price decline might be recorded in the FR&R index. But the claimed advantages of the FR&R method evaporate, because FR&R has been proposed as a way to estimate accurate price indexes for high tech products without going to the expense of hedonic methods.

2. Case Two

The first case (no price response of continuing computers to the entering computer) is a limiting or extreme one. Consider the case at the other extreme: Suppose prices of the continuing computers fall fully and instantaneously to respond to the new computer. This yields a new hedonic function, such as h(t+1) in Figure 4.3, in which computer n is on the hedonic surface (for simplicity, I continue to suppose that the price of computer m lay on the old hedonic surface).

In the case of instantaneous price response, the FR&R index will pick up all the price decline accompanying the introduction of computer n, because the market effect of the new computer is reflected in the price changes of continuing computers that are inside the FR&R sample. On the other hand, the fixed sample index will also suffice, because with instantaneous market adjustment, the index will record the full price decline even if computer n is not included in the sample. In case two, then, the FR&R index equals the hedonic index, but the FR&R index is not necessary, the fixed sample will also pick up all the price change.

3. Case Three

The most realistic case lies between the two extremes: The introduction of the new technology brings about some price response, but the market does not fully adjust instantaneously. Some of the price decline accompanying the new computer's introduction will be picked up by the matched model FR&R index, but not all of it. The reasoning is the same as for case one: If the FR&R index is implemented with the deletion (IP-IQ) quality adjustment method, as would normally be the case, the imputed price change for the exiting/entering computer will be the price change for the continuing computers; by specification in the example, prices of continuing computers do not decline as much as the price of the entering computer. The FR&R index misses some price change measured by the hedonic index, but not necessarily all of it. The hedonic index measures all of the price decline, whether estimated by the dummy variable method, the hedonic imputation method, or another hedonic method.

Ultimately, as more new computers similar to computer n are introduced, the new technology will shift the hedonic function down to a new level, but this may take some time. When that happens, all computers made with the old technology may disappear, because it is not possible to produce computers on the new price/quality frontier with the old technology. For example, 386 chip technology for PC's replaced 286, Pentium replaced 386 and 486, and was in turn replaced by Pentium II, III and IV.

One might contend therefore that the FR&R index eventually follows the hedonic index, so the error in case three would be small, or that it would just amount to a lag. This is an empirical proposition that might be true, but it is hazardous to count on it. Suppose all of the matched models initially remain in the sample (that is, computer m does not exit, so it does not trigger a forced replacement). Computer n enters at a lower price/quality level and is followed by its technology mates. No prices of m-type computers fall. The pricing agency updates its sample frequently (a FR&R index), so it gradually replaces m-type computers with n-type computers. Then at some point all the m-type computers, having gradually been replaced in the sample, exit from the market. No price decline will be recorded in any matched models comparison. Yet, the quality-adjusted prices are lower in the end period than at the beginning.[9]

Summary. Whether a FR&R matched model index incorporates all the price change recorded in a hedonic index depends on several factors:

(a) *The speed with which the FR&R matched model index incorporates new varieties*. An annual resampled and reweighted index might be considerably less effective than a monthly one, because too much model turnover takes place before the annual sample is replenished. The hedonic index is better because it contains an explicit price change measure for exits and entries, which may be large over a year's time. Conversely, the higher the frequency of the FR&R, the more its measurement should approach that of the hedonic index.

(b) *The amount of price change that occurs at the point of introduction of new varieties, rather than after their introduction*. If new varieties are introduced with (quality-adjusted) prices that are lower—or higher—than the ones they replace, these price effects will be missed by the FR&R index. If the price/quality ratios of entering computers are similar to old ones, little price change is implied by entries, and little is lost by omitting them from the FR&R index. Similar statements apply to price changes implied by exits.

(c) *The effect of factor (b) is combined with a third factor—how rapidly do prices of continuing varieties respond to the new introductions?* If markets respond instantaneously to the competition of new and improved varieties, then the FR&R price index, which necessarily imputes the price changes of entering and exiting varieties from the price changes of continuing ones, nevertheless will pick up the price impact of the entering varieties. If market responses are slow, the FR&R index will miss price change associated with entries and exits, or will incorporate them with a lag.

(d) *The weight of the entering and exiting varieties*. If the weight of entering and exiting varieties is small, one expects that their market impact will also be small, whether or not the entering varieties have lower—or higher—quality-adjusted prices than the varieties they replace (factor b), and whether the market adjusts instantaneously or does not (factor c). One would expect, though, that the weight of entering varieties will be greater when they offer price reductions and when the market response is slow.

These four factors imply the need for empirical work to quantify them in different markets and for different ICT products, for the purpose of guiding practical strategies for constructing accurate ICT price indexes. Available information is reviewed in section D.

Before turning to the existing empirical studies, it is worth noting that the assertion that hedonic indexes and conventional indexes should give the same answer is quite old. Jaszi (1964) gave an example of coffee-chicory mixes in different proportions, and asserted that the price should reflect the costs of the

[9] This speculation appears broadly consistent with the findings of Moulton, Lefleur, and Moses (1999). It is explicit in Dulberger (1989). See also Von Hosten (1952) for a parallel discussion involving new car models. The point is quite old, and the potential problem is quite well known and discussed in the price index literature, yet it is still neglected.

different mixes, which would also be recovered in a production cost quality adjustment method (see Chapter II), or in the ratios of their prices in the overlapping link method.[10]

Jaszi's example, as does other similar reasoning on the issue, implicitly assumes market equilibrium prices for both periods in a price index comparison. If newer computers had higher quality relative to their price, this logic goes, that should push down the prices of existing models—in factor (c), above, response is instantaneous. Both the hedonic index and the matched model index should reflect the effects of quality improvements, whether the quality change occurred on the models in the sample or outside it. A properly constructed matched model index, on this line of reasoning, should equal a hedonic index.

The interesting cases are those where market equilibrium might not prevail. Moreover, what might seem reasonable for coffee-chicory mixes (which anyone can mix) might not seem so reasonable for differentiated high-tech products. With differentiated products, sellers look for market niches that are not served by the existing variety spectrum. If they locate such a niche, innovation in creating a new variety implies a return, and so implies at least short-run market power, which can hardly be conceived for a new coffee-chicory mix. Reasoning that markets should respond quickly to innovation in new product varieties is not confirmation that markets do in fact respond quickly.

D. EMPIRICAL STUDIES

Many comparisons between hedonic indexes and matched model indexes have been carried out, on many products.

1. Early Studies: Research Hedonic Indexes and Statistical Agency Matched Model Indexes

The early studies that compared hedonic indexes with statistical agencies matched model indexes are reviewed in Triplett (1975). Gordon (1990) contains similar comparisons for many investment goods.

In many of these studies, researchers found that hedonic indexes rose less rapidly, or fell, compared with published government matched model indexes, and this result has found its way into the folklore of economics, despite the fact that contrary examples are numerous. These contrary findings—when hedonic indexes rose more rapidly than the price indexes with which they were compared—should have received more attention. The analysis in Chapter II shows that the direction of bias from the application of conventional methods is sometimes more nearly a function of the direction (rising or falling) of true price change than the direction (improving or deteriorating) of quality change. Thus, if hedonic indexes are more accurate than conventional ones, they might show less price increase or more price increase, depending on the direction of quality change bias from conventional methods; they might also agree, if conventional quality adjustments happened to give the same result as hedonic methods.

In any case, these early studies were not true evaluations of alternative quality adjustment methodologies. Most researchers computed hedonic indexes from databases consisting of a series of large, extensive annual samples or near universes of models, in which the prices were usually list prices, not transactions prices. They compared them with statistical agency matched model price indexes for the same or a similar product. The statistical agency indexes were derived from small samples, and their prices were transactions prices, as nearly as the agencies could obtain them.

---

[10] The example is interestingly dated. Chicory root was once used to "cut" coffee in some places in the U.S. to make a less expensive beverage.

Thus, these early hedonic-matched model comparisons confounded differences in databases, fixed sample-universe sample differences, and so forth, with differences in quality adjustment methods. To an undetermined extent, reported differences between research hedonic indexes and published government agency matched model indexes reflected something more than the difference between matched model and hedonic *methodologies*. Missing from the early studies were hedonic index-matched model index comparisons that covered the same database.

2. Same Database Studies: Hedonic Indexes and Matched Model Indexes

We need studies that estimate, solely, differences associated with matched model and hedonic methodologies. I am accordingly very selective in the studies included in this review. I include only studies that (1) estimate matched model and hedonic indexes using the same database, and (2) conform to best practice for both methodologies. That means FR&R indexes for matched model method and one of the best practice examples of hedonic indexes. Even so, no doubt some studies that could have been included in this review were not, but any exclusions were not purposeful.

Dulberger (1989) was the first computer study that compared hedonic and matched model indexes computed from the same database. Dulberger's data consisted of a universe of IBM and "plug compatible" mainframe computers, a selection that was done to assure comparability in the speed measure in her hedonic function. The data were, then, not a full universe of computers, but also not a sample in the usual sense, and consisted of many more computers than would be included in normal statistical agency samples. In addition to hedonic price indexes constructed by several different methods, already discussed in chapter III, Dulberger also computed a matched model index from the same data. The results are summarised in table 4.3.

Over the entire interval of her study, Dulberger's hedonic indexes declined at similar rates, 17-19 percent per year for the three hedonic indexes. But her matched model index declined only half as much—8 ½ percent per year. These calculations involved the whole dataset, not a relatively small statistical sample. All indexes were based on annual data.

McKenzie (2002) used IDC data that covered desktop and laptop computers sold in Australia. McKenzie's IDC database was a large sample, though coverage of the Australian market was by no means complete, and it contained "high frequency" data, in his case bi-monthly observations. However, market share weights were only available at the firm, not the model, level. Separate indexes for desktops and notebook computers were estimated; in Table 4.4, these are weighted together (desktop and notebook price movements were not greatly different in Australia over this period).

McKenzie computed four indexes: (a) small sample and large sample matched model indexes, both of which used ordinary linking methods (IP-IQ, see Chapter II) for changing sample composition, and (b) two forms of hedonic index (dummy variable using the large sample, hedonic imputation using the small sample). The four indexes resulted in two different hedonic and matched model index comparisons, a large sample comparison and a small sample comparison. In both comparisons, sample size and composition were held constant. This permitted a more complete and thorough analysis of the issues than in some other studies.

The large-sample matched model index declined rapidly—more than 30 per cent in less than two years. The small-sample matched model index was designed to mimic the sample that a statistical agency might adopt, except that the models were chosen so that an overlap always existed for entries and exits from the sample (in practice, this would not be likely). This "overlap sample" index declined about the same rate as the full sample, matched model index, actually a bit more (see Table 4.4). This somewhat surprising result may be an artifact of choosing, retroactively, an overlap sample.

McKenzie then computed two hedonic indexes. One applied hedonic quality adjustments for sample forced replacements to the small sample matched model index. This is methodologically similar to the method used by the U.S. BLS (see Chapter III). The second hedonic index used the classic dummy variable method (Chapter III), applied to the large sample. As Table 4.4 shows, these two hedonic indexes are not identical.[11] However, both hedonic indexes declined substantially faster (10 to 20 percentage points faster, depending on the comparison) than either the large sample or the small sample matched model indexes.

It is intriguing that McKenzie's two large-sample indexes differed the most—the matched model, all models index fell by 32 percent, the hedonic pooled dummy variable index by 52 percent, a difference of 20 index points. Moreover, the FR&R index declined more slowly than the fixed-sample, hedonically adjusted index (32 percent, compared with 45 percent). In Australian data, a large-sample, FR&R matched model index does not give the same answer as a hedonic index calculated on the same data, nor is the large sample index demonstrably better than the small sample index with hedonic quality adjustments.

Okamoto and Sato (2001) compared FR&R matched model indexes with hedonic indexes covering several products, all computed from monthly scanner data for Japan. The authors computed alternative hedonic indexes (dummy variable method and characteristics price index methods) as well as alternative matched model indexes (different formulas). Table 4.5 shows an extract of results for PC's, for colour TV's, and for digital cameras. All the indexes show rapid rates of decline. PC indexes decline at the fastest rate (over 40 percent per year) The price indexes for colour TV's show similar price declines as those estimated for other countries.

Two result from this study stand out. First, among hedonic indexes, on the one hand, and matched model indexes, on the other, alternative implementation methods are not important, quantitatively. Dummy variable and characteristics price methods give exactly or nearly the same hedonic indexes in each product—over 45 per cent per year in the case of PC's, approximately 22 percent per year for digital cameras, 10.4 percent per year for TV's. Similarly, matched model indexes are not sensitive to alternative index number formulas, when the formulas are restricted to those with good index number properties.[12]

Second, hedonic indexes differ from FR&R matched model indexes. One might judge the difference to be small in the case of PC's, where the hedonic index declines at 45 percent annually and the matched model index at 42 percent. At around 3 percentage points, this difference is still several times larger than the difference between alternative implementations of each method, so the difference is in some not quite precise sense statistically significant (see the summary section, below). The difference between hedonic indexes and matched model FR&R indexes is larger for TV's, where hedonic indexes fall substantially more slowly (10 ½ percent annually) than matched model FR&R indexes (nearly 19 percent annually); it is nearly as large (nearly 6 percentage points, over a shorter interval) for digital cameras. Okamoto and Sato present a chart for digital camera indexes (reproduced as Figure 4.4) showing that these correspondences do not just apply to the end periods: For every month, alternative hedonic indexes agree and alternative matched model indexes agree, but hedonic indexes differ from matched model indexes.

[11] This comparison has already been discussed in chapter III.

[12] The authors also present chained Laspeyres versions of the matched model index and the characteristics price index; the Laspeyres versions predictably drop less than either the geometric mean or the superlative index number formulas.

Evans (2002) constructed hedonic indexes with the hedonic imputation method and matched model indexes, using IDC data for France. Subjects were three ICT products (desktop PC's, laptops, and servers). Indexes were produced at quarterly frequency, where the hedonic function was re-estimated quarterly, and then aggregated into a price index for computers. Table 4.6 presents aggregated indexes. As Evans notes: "Over the six quarter period, the hedonic computer producer price index fell 42.1 per cent, compared to a decline of 13.7 per cent recorded for a matched model index derived from the same database."

Van Mulligen (2002) computed FR&R matched model and hedonic indexes for desktops (PC's), laptops and servers, using a near universe of "market intelligence" data (from the GfKcompany) for the Netherlands. Average annual rates of change for these indexes are presented in Table 4.7.

For each of the three products in Van Mulligen's study, the FR&R matched model index declined by 21-22 percent per year over the three-year period. Hedonic indexes for all three products recorded more price decline than the corresponding matched model index, but the matched model-hedonic difference varied with hedonic computation method and with the product: The smallest (about 3 percentage points per year) applied to the comparison using the hedonic imputation indexes for notebooks and servers, the largest (nearly 11 percentage points per year) emerged for the dummy variable index for PC's.

Even though a difference of 3 percentage points per year might seem small, it still suggests that the matched model index understates price decline by roughly 14-15 percent (3 percentage points over 22 percent decline). The largest difference, 10.7 percentage points for PCs (21.9 percent per year for the matched model, compared with 32.6 percent per year for the dummy variable index) indicates an understatement of nearly half, measured using the matched model index as the base. As with Okamoto and Sato's results, what may appear small differences (by some metric) still suggest that methodology creates significant differences.

Note that in every case, Van Mulligen's dummy variable indexes recorded the greatest decline. At this writing, it is not entirely clear why his hedonic imputation indexes decline so much less than his dummy variable indexes.[13]

Silver and Heravi (2001, 2003) compared matched-model and hedonic indexes for U.K. television sets and washing machines. Although the latter do not qualify as ICT products, they complement the relatively small number of available ICT studies, so are included in the present discussion. Table 4.8 presents the results.

Like Okamoto and Sato, Silver and Heravi computed hedonic indexes according to several methods and matched model indexes by several formulas. They find that matched model indexes uniformly fall more rapidly than hedonic indexes, which implies that new appliance models enter at prices that are above the quality-adjusted average for continuing models (contrary to all the results for computers). This finding is also repeated in the TV study by Okamoto and Sato (2001). A particularly valuable comparison involves the two Fisher-formula indexes. The Fisher matched model index weights by sales of models; the Fisher characteristics price index weights by quantities of characteristics. Thus, the weighting patterns are not the same. Nevertheless, both are superlative indexes, so comparisons involving them are cleansed of poor index number properties, and might therefore be said to be the best examples of their respective types. In each case, the matched model index records more than two percentage points more price decline than the hedonic index.

[13] See the discussion in Chapter III, and also Moulton, Lafleur and Moses's (1999) somewhat similar findings.

Summary. These are a small number of studies. The differences between hedonic indexes and matched model indexes are also small in some cases, though not small in others. Nevertheless, the estimates that exist provide little support for the idea that matched model and hedonic indexes generally give the same result when FR&R methods are used.

We can sharpen the conclusions drawn in this section by utilising the old statistical distinction of variability within groups, compared with variability between groups. An extensive price index literature on differences in formula exists, and it is well known that different index number formulas can produce different results. This is within-group variability: Differences in matched model index number outcomes across index number formulas record variability *within* the matched-model method (one group).

It is also established that hedonic indexes may differ according to the method of application, though as discussed in Chapter III, the divergence among alternative hedonic indexes is not so great as has often been supposed. Differences in index number outcomes among alternative hedonic methods constitute within-group variability for the second group (hedonic methods).

Consider the size of the variability within index type (that is, within the *group* of hedonic indexes and within the *group* of matched model indexes). One can ask whether systematic differences *between* the two groups is larger than the differences *within* the groups. In the absence of a good measure of central tendency for this problem, I use the range.

Silver and Heravi, and also Okamoto and Sato, estimate different matched model indexes and different hedonic indexes on the same data. In Table 4.9, I show the ranges of matched model indexes and of hedonic indexes calculated by these two sets of authors. In every case but one, the within-group range is greatly exceeded by the between-group range.

For example, for PC's Okamoto and Sato (2001) find that the range of matched model indexes amounts to 0.5 index points, and the range of hedonic indexes is 0.6 index points; these represent about ¼ to 1/5 of the between-group range (variously, 2.4 to 3.0 percentage points). Thus, in a not quite statistical sense, we can say that hedonic indexes and matched model indexes for PCs differ statistically. Differences are larger for the other two products.

Similarly, Silver and Heravi (2002) find ranges within matched model and hedonic indexes of 1.0 and 0.8, respectively, for TV indexes. The range of the between-group difference for this product (2.2 to 2.5 percentage points) is more than twice as large. The main exception in table 4.8 is the case of washing machines in Silver and Heravi, but this is ambiguous. Even though the between-group range is little larger than the within-group range for this product, every matched model index declines more than any hedonic index, so there is not much question that the two alternative methodologies give different answers.

Actually, there is no overlap in the ranges of hedonic and matched model indexes for any product studied by either pair of authors. All hedonic indexes lie outside the range of all the matched model indexes.

I have restricted the within-group and between-group comparisons to indexes with "good" properties—a rough and ready Baysian restriction. Okamoto and Sato, for example, also calculate a Laspeyres index, which predictably diverges from their two superlative indexes. Had I included that, the range for matched model indexes would have enlarged considerably. However, if a Laspeyres index differs from a superlative index, we know from index number theory that the superlative is better. There is little sense in including index numbers that are not best practice in the comparison in Table 4.9. Similarly, I do not report in Table 4.9 results from studies where the matched model index reported is not

FR&R. The same point applies to hedonic indexes that are not best practice. The appropriate test for whether hedonic methodology and matched model methodology matters rests on comparing best practice versions of each type.

3. Analysis

As the preceding discussion and tables 4.1 to 4.7 show, in most empirical comparisons matched model and hedonic indexes produce different indexes, even when the matched model indexes are constructed on FR&R methods. In some studies, the differences are small. One can analyse these studies in terms of the four factors discussed at the end of section C, namely:

- *Frequency of resampling and reweighting*.
- *The amount of price change that occurs at new product introduction.*
- *The speed with which prices of older products adjust.*
- *The weight of entering and exiting products.*

We expect that the difference between matched model and hedonic indexes will be smaller the greater are the first and third factors (e.g., if older products adjust rapidly to prices of new entrants, the difference between matched model FR&R indexes and hedonic indexes should be small). The difference will be greater when the second and fourth factors are larger (e.g., the larger the price change at introduction, the greater will be the difference between FR&R indexes, which miss such price changes, and hedonic indexes).

Information on these four factors is summarised in Table 1.10. A few points stand out.

Dulberger had weights, but hers was not a high frequency sample (it was annual, so R&R, but not "F"). Her matched model indexes might therefore not conform to the results of a full FR&R index. On the other hand, all the other studies produced full FR&R matched model indexes. Most of them reported that hedonic indexes differed from matched model indexes on the same data. *Frequent resampling and reweighting does not, by itself, assure that matched model indexes will coincide with hedonic indexes.*

Only two studies produced information on the second factor (price change at product introduction). Dulberger (1989) contended that new computer models were introduced at lower quality-adjusted prices than were offered by older computers. She produced estimates in support of her position: Entering computers implied substantial price reductions. Matched model indexes miss these introductory price changes, because they only track price changes after a new model's introduction. Silver and Heravi (2002) find that both entrants and exits produce price changes that differ from those of continuing products, so the matched model index misses both kinds of price change, in their case. No other studies have produced direct estimates of this effect.

The importance of evaluating price behaviour of entering and exiting computers cannot be stressed too strongly. Most matched model indexes—and nearly all FR&R matched model indexes—are constructed so that price changes for entering and exiting computers are implicitly imputed from the continuing models (the IP-IQ-deletion method, discussed in Chapter III). They build in, that is, the assumption that entrants have no independent impact on price changes at the point of their introductions, nor do exits imply any price changes by the fact of their exiting. Some researchers have speculated that entry and exit effects are not important, and perhaps they are not in some cases. But the only researchers who have explicitly evaluated the matter empirically (Dulberger and Silver and Heravi) have concluded that entries and exits do matter. Future comparisons of matched model and hedonic indexes would benefit from following the good examples of these studies.

Speeds of adjustment are the third factor. They are difficult to estimate. Dulberger and Silver and Heravi reasoned somewhat indirectly: Prices of continuing models did not adjust quickly (Dulberger) or did not decline as fast as those of new entrants (Silver and Heravi). No other study considered this matter.

With respect to the final factor, Aizcorbe, Corrado and Doms (2000) emphasise that their frequently-reweighted system means that entries and exits from their sample get low weights. Price changes that are associated with entries and exits, they contend, will therefore have a small impact on their indexes, even when the FR&R procedures miss these price changes. At some degree of frequency, their contention must be correct. On the other hand, Van Mulligen (also FR&R) found that entries and exits corresponded to around 20 percent of the expenditure weight, monthly. Silver and Heravi also emphasize that weights of entrants and exits are not low. Too little information has been presented on this matter; it is certainly an important factor that influences whether matched model and hedonic indexes differ.

E. MARKET EXITS

Market exits deserve more attention in empirical studies. A product that exits from the market might imply price change, either for all buyers or for buyers occupying some market niche. Pakes (2003) emphasises market exits.

Suppose the old computer, m, disappears from the matched model sample and that its price was higher than what would have been predicted from its characteristics (that is, it lies above the hedonic surface, as shown in Figure 3.5, in Chapter III). Its exit implies a price reduction in the sense that the range of (quality adjusted) prices is lower than it was previously. Pakes (2003) speculates that product varieties disappear because they decline in price more than surviving varieties (and hence are withdrawn). One might contend that the exits of overpriced computers cannot have a major impact because buyers could always have selected a model that was not overpriced, so their disappearance should be of concern to no buyer.

On the other hand, suppose the exiting computer was a "bargain," its price lay below the hedonic line, as in Figure 3.9 (in Chapter III). Then its exit from the sample implies a price increase, in the sense that the range of (quality adjusted) prices is higher than it was previously.

One normally expects the overpriced computer models disappear from the market, not the bargains. Many product exits no doubt occur because they do not offer good value.

However, the whole analysis of quality change requires that the buyers have different preferences. If not, as Rosen (1974) showed, little variety will exist, because everyone who wants a €2000 computer will buy exactly the same configuration. Producers will make a single computer model at each price. In fact, all consumers do not buy the same product variety; instead, they occupy distinct market niches.[14] If a product niche disappears, those consumers will be affected, even if other consumers are made better off by the products that replace them.

[14] Berry, Levinsohn, and Pakes (1995) analyse this market niche problem in a market for differentiated products.

Are exits predominantly of over-priced products, or do some of them represent disappearances of market niches? The latter implies a price increase for consumers who bought niche products that are no longer available. There is little research on this matter. The effect of product exits has been too little studied in the price index literature, and the topic has generally been treated far too cavalierly.

In any case, however, the matched model index has no way to take account of these market exits. In most matched model applications, the price behaviour of exits is imputed from price changes for continuing models.

F. SUMMARY AND CONCLUSION

When does the matched model methodology give the same result as hedonic methodology? The answer depends largely on the nature of competition in the market for computers and high tech products, and secondarily on the form of the matched model method that is used.

1. Three Price Effects

The discussion in this chapter suggests that three price index effects arise from the introduction of new product varieties outside the price index sample. These effects apply to universe samples with FR&R, they are not restricted to small, fixed samples.

First, the introduction of new varieties at a more favourable price/quality ratio than established varieties will put pressure on the prices of the older varieties in the price index sample. If prices of older computers adjust rapidly to the competitive threat of the new computers, then the matched model methodology may adequately record the price change, especially when the deletion (IP-IQ) form of quality adjustment is used.[15] Matched model indexes—even fixed sample matched model indexes—and hedonic indexes should not differ greatly. This case is the one that motivates most contentions that matched model and hedonic indexes ought to record the same measure of price change.

The second price effect arises out of the stereotypical “product pricing cycle.” Prices of new product varieties are introduced at relatively high prices, but the prices of these new varieties subsequently fall, relative to the existing ones. Their market shares, initially quite small because of their initially high prices, expand rapidly as their prices fall, and the shares of the old varieties fall as they are displaced by newer, cheaper (on a quality-adjusted basis) varieties.

For this problem—price change for new introductions that differs from the price change for older computer models—it is important to distinguish between two kinds of samples. This product pricing cycle problem creates error in fixed sample indexes. It can be handled effectively by changing the sample frequently, and chaining the indexes (FR&R). As Griliches (1993) remarked: “The right way to compute such indexes is to weight these changes, and change the weights as one goes along. If the price declines are occurring during a period when very few of these models are being bought, then they will have very little weight in the total. And that is all there is to it. There is no special mystery about that.”

The rapidly replenished (FR&R) samples will typically show more price decline than the fixed sample in the case of high technology goods. Because hedonic indexes are typically run on extensive cross sections of computers and are updated in each period, they also bring new introductions into the price index quickly.[16] If samples are large and very frequently replenished, product cycle pricing effects

[15] The link-to-show-no-change method (still used in some price index programs—see Chapter II) will miss these changes entirely.

[16] If they are run on annual data, say, then they may not bring the new introductions in quickly, but they make a quality adjustment for them, which accomplishes much the same thing.

should be measured consistently with FR&R matched model methodology and with hedonic methodology, provided the "frequently" in FR&R is "frequent enough."

The third price effect is price change that occurs, not *after* the introduction of new product varieties, but *contemporaneously with* new introductions, and that is not matched by price changes in continuing models. This price effect occurs when the price/quality ratio of new machines differs from old machines and the prices of old machines do not adjust instantaneously. When these entry and exit effects are large, FR&R matched model indexes will differ from hedonic indexes, unless some non-traditional method of quality adjustment is incorporated into the FR&R methodology. On the other hand, when entry and exit price effects are relatively small, either because the price changes themselves are small, or the weight of the entry/exits is small, then keeping the sample up to date is effective in dealing with the new product variety problems. Matched model indexes in these cases may give price indexes that are close to hedonic indexes.

Some economists and statisticians have supposed that the third effect cannot be large, that it cannot dominate other price changes that can be measured with FR&R sampling methods. The empirical work cited in this chapter indicates that such faith in traditional methods is misplaced, at least for high tech products with a great amount of product turnover.

2. Price Measurement Implications: FR&R And Hedonic Indexes

The empirical review in section D indicates that hedonic indexes differ from FR&R matched model indexes in every study of computers where both types of indexes have been estimated. In some cases the differences may be small, or they may be small enough that statistical agencies feel they can be ignored. If difference are small, they might ask, why go to the expense of estimating hedonic indexes? Matched model indexes will do—or considering both costs and benefits, hedonic indexes, they may say, may not be cost effective.

Cost effectiveness is the right question to ask. One can ask the cost-effective question of FR&R methods, not just of hedonic methods. FR&R sampling is expensive. There is a general perception that fixed sample methods with conventional quality adjustments are cheapest, but least effective. A less generally held perception has hedonic indexes as the most effective approach, but they are judged the most costly, with FR&R somewhere in between. However, this simple ranking is oversimplified, both on the cost side and on the effectiveness side.

a. Effectiveness

Existing research suggests that for technologically dynamic products FR&R methods record some price change that fixed sample methods miss, and that hedonic indexes record price change that FR&R methods miss. Owing to the small number of studies that have been carried out on the same database, one cannot be that precise about the magnitudes of the differences: Sometimes they will be small, but not always. Small or not, where comparisons have been made on the same data, differences are statistically significant, judged by the within group, between group analysis in section D.

One should also underscore that hedonic indexes are not always lower than matched model indexes, despite the widespread perception that they will be. TV price indexes are a good example.

The effectiveness of FR&R over fixed sample methodology, and of hedonic over FR&R methodology, depends on the industrial organisation of particular markets and on the way that new products are introduced and priced. These considerations imply that the margins of effectiveness among the three approaches will vary market by market, that is, product by product and country by country. For computers, it seems clear that hedonic indexes are most effective. However, the results for computers and for some appliances that are summarised here may not apply to price indexes for other products, where market conditions differ. This represents a strong caveat to the findings of existing research.

This "it depends" conclusion is in a real sense discouraging, because it implies a continuing research program to find out how much gain in effectiveness more advanced price index methods will yield in particular markets. The most that can be said is: Reliance on simple market equilibrium notions to justify fixed sample methods with conventional quantity adjustments is holding on to a treacherous and unreliable standard.

Before turning to cost considerations, it is worth noting that hedonic indexes can be used to evaluate the potential error when agencies are not able to execute FR&R sampling strategies, because hedonic indexes facilitate constructing a price index when the observations in the beginning and ending periods are quite different. To avoid the expense of FR&R methods, price index agencies often try to measure technologically dynamic industries with a sample of old products, or make do with fixed samples where continuing products, hopefully, represent the price changes of all products that are not in the sample. If universes of prices and characteristics are available from some source, perhaps on an annual basis, the results of a hedonic index can be compared with the results from a normal statistical agency fixed sample. Evaluation studies using hedonic indexes can be valuable in tracking situations where the effectiveness of alternative methods is in doubt.

b. Cost

On the cost side, it is usually presumed that FR&R methods and hedonic methods are more expensive than fixed samples, mostly because FR&R and hedonic indexes require much more data. It is less clear that either is more expensive than a smaller fixed sample that adequately allows for quality change and is set up to reach out for new products and bring them rapidly into the index: For example, Lane (2001) discusses "directed" and "targeted" rotations of CPI samples to assure that new goods are rotated more rapidly into the index sample.

It is also less clear that hedonic methods are more expensive than FR&R methods, *when both are done with comparable data quality concerns*. It is sometimes contended that FR&R methods are cheaper than hedonic methods because FR&R indexes require only a large-scale sample of prices, and do not require data on characteristics. There is something to this. However, to do matched model indexes right one needs to control the matches for characteristics of the products, as discussed in Chapter II. For high-tech products, the characteristics incorporated into an statistical agency's pricing specification are often the same or nearly the same as the characteristics that go into a hedonic function. Thus, *FR&R methods need data on characteristics, even for a matched model calculation, and the characteristics will be the same as those needed to estimate hedonic indexes*.

It is sometimes contended that manufacturers' model numbers are sufficient for matching in a FR&R sample, that characteristics are not needed. Model numbers are certainly useful, but statistical agency commodity analysts generally are not satisfied with model numbers for assuring a match in conventional price index collections. There is little reason to suppose that model numbers are any more adequate when much larger samples are collected for FR&R indexes. One needs the characteristics to be sure that matching by model numbers creates real matches. From this, FR&R indexes need nearly the same body of characteristics data that are necessary for estimating a hedonic index.

It is often said: If one has the data for an FR&R index, why do the hedonic index? It might be better to turn this question around. If one has good data for an FR&R matched model index, including the characteristics necessary to assure an accurate match, one normally has the data for computing a hedonic index. So why not do the hedonic index? Research results indicate that the hedonic index is more effective.

Sometimes in-between cases are posited: The FR&R sample may have data on some characteristics, which can accordingly be used for matching, but perhaps too many variables are missing to estimate a reliable hedonic function. But this contention is also flawed: If variables are missing for the hedonic function, they are also missing for the matched model index. Undetected changes in these missing variables will bias the matched model index, just as they bias the hedonic index.

The FR&R matched model index that uses some of the characteristics might be cheaper than the full hedonic index that uses all of the relevant ones. But these are not like comparisons. The FR&R matched model index that makes matches on an inadequate set of variables is not as good as a hedonic index constructed on all of the characteristics. The "FR&R is cheaper" view does not, in this case, apply to equally effective methods.

In all these comparisons, one must ask: What is the cost of producing alternative quality adjustments? There is a great shortage of cost estimates that reflect actual conditions inside agencies and that are therefore relevant to the choices that must be made.

In work leading up to the adoption of hedonic indexes for computers in the U.K. (Ball, et al., 2002), the Office for National Statistics (ONS) compared costs for estimating a hedonic index for computer equipment with the costs of its previous method (matched model indexes, with quality adjustments by the option price method). Obtaining option prices for sample changes caused by forced replacements is not without cost. Making hedonic quality adjustments for changes in the sample is actually cheaper than the option cost method, case by case; but this cost saving was offset by the need to do the research and estimate the hedonic function. In all, whether the hedonic index was more expensive depended on how frequently the hedonic function needed to be re-estimated. But whatever the cost comparison, ONS concluded that hedonic quality adjustments were better than option price quality adjustments, in part because hedonic adjustments were less likely to require arbitrary judgements, and so were more objective than option price adjustments.

Decisions on cost and effectiveness must be made by individual countries' statistical agencies. Situations differ sufficiently across countries that different judgements may well emerge.

Revised: Dec 24, 03

APPENDIX TO CHAPTER IV

A MATCHED MODEL INDEX AND A NON-HEDONIC REGRESSION INDEX

A widely-noticed, unpublished paper by Aizcorbe, Corrado, and Doms (2000, hereafter, ACD) has been interpreted as saying that a matched model index that is computed from a frequently-replenished, frequently re-weighted sample would coincide with a hedonic index. A second paper on the same lines is Turvey (2000).

ACD computed a matched model index from cells that were defined by manufacturers' model nomenclatures. They replenished their sample frequently, it was a very large sample, and they had weights for each of the cells. Their matched model index declined at a rate that is not far from the rates recorded in the U.S. PPI personal computer index, which incorporates hedonic quality adjustments. Not unreasonably, ACD conclude that a matched model index constructed on the FR&R principle will probably decline more rapidly than one where the sample is not so large, where it is not replenished frequently, and where no weights are applied to the individual cells. This conclusion is consistent with previous work on high technology products, particularly the pioneering work by Dulberger (1993) and Berndt, Griliches and Rosset (1993).

ACD further state that the difference between their matched model index and a hedonic index will be smaller the smaller is the weight of entries and exits, which is correct (the weight of entries and exits is one of the four conditions discussed at the end of section IV.C). They contend that their frequent replenishment and re-weighting must surely work better than the typical statistical agency fixed sample, and again this contention is certainly correct and, as noted above, consistent with earlier work that has perhaps not been given sufficient attention. The fairly close correspondence between their index and the PPI computer index is suggestive: Carefully constructed matched model indexes, computed on a large frequently replenished database which contains current information on weights will show rapid declines in computer prices, and might well approximate hedonic indexes if conditions are right.

Their work is saluatory in another dimension: Some economists and statisticians still distrust hedonic indexes because the indexes fall so fast and they, like Denison (1989), see the hedonic index as a 'black box." For those persons, the size of the price declines in ACD's matched model indexes should be reassuring.

It might also be true that such a matched-model index will decline as fast as a hedonic index *estimated from the same data*. ACD have also been interpreted as showing that, and hence have been interpreted as providing a counter-example to the studies that are reviewed in the text of Chapter IV (all of which find that matched model and hedonic indexes computed on the same data give different results). However, ACD do not actually do what the other studies do, for they estimate no hedonic function. The regression index that they compare with their matched model index differs from a normal hedonic index in ways that assure that it will be close to a matched model index, and, actually, it will have the same deficiencies as a matched model index.

For both their indexes, ACD divide up the computer commodity spectrum into cells, which are defined by manufacturer's model numbers. They count on the model number to assure that within a cell no quality variation is permitted. They collect the price and sales quantity for each cell, but they have no information on the characteristics of the computer in each cell. Their matched model index uses the price and quantity associated with each cell in a traditional index number formula (they compute several).

ACD contend (communications with the authors) that their regression index methodology is like a hedonic function with "fixed effects" for each computer model. They assign a unique dummy variable to each computer *that is available in both of two periods* (this is the fixed effect). The fixed effects control for computer quality in a way that is similar to the use of explicit controls for characteristics in the usual hedonic function. The regression covering two periods (t and t+1) contains the prices of the computers, in the usual way, and the fixed effects variables plus the normal time dummy variable. Adapting the equation for the dummy variable method (equation 3.1 from Chapter III), this gives (for an adjacent period regression):

$$\ln P_{it} = d_o + \sum d_i f_i + b_1(D_{t+1}) + \varepsilon_{it},$$

where f_i designates the fixed effect variable for computer i and d_i is its coefficient. The index is estimated from the coefficient b_1, as explained in Chapter III.

The fixed effects idea has potential in estimating hedonic functions. As noted in Chapter V, researchers who estimate the usual hedonic functions impose smooth contours on the function. But computer models are unique. They are only available for some combinations of characteristics, not for every combination. The usual smooth hedonic functions that are estimated with OLS regressions imply that the product space is filled densely without any gaps, but there are gaps. Thus, it seems reasonable to suspect that the true hedonic function has "kinks and bumps" in it, rather than the smooth functions that everyone estimates. Estimating a function that has fixed effects for each computer model permits the function to take on shapes that are not smooth. In that sense, the function estimated by ACD is a substitute for the usual hedonic function, and one that might be justified empirically, if the objective were only to determine the shape of the hedonic function.

For estimating hedonic price indexes, though, the fixed effects model has exactly the same deficiency as the conventional matched model quality adjustment methods discussed in Chapter II. Suppose new computer models were introduced in period t + 1, models that were not available in period t. Those new computers cannot be assigned a fixed effect in period t + 1, because the fixed effects would be exactly collinear with the time dummy variable for period t + 1. For this reason, the new computers must either be left out of the regression, or they are treated with one of the three matched model quality methods discussed in Chapter II—though ACD do not discuss this point, leaving them out of the regression implies something equivalent to the deletion (IP-IQ) method described in Chapter II.

We know there is potential for price change when a new product enters the sample or when an old one exits. As explained in Chapters III and IV, the hedonic index measures these price changes, because it estimates a quality-corrected price change for the entries and the exits. The matched model index does not, whether or not it is frequently resampled and re-weighted. Fixed-effects regression indexes also contain no estimates for entering and exiting computers, they can only measure price change on the continuing part of the sample, as is also true of most conventional matched model indexes. The empirical magnitudes of entry and exit price effects are presented and discussed in Chapter IV.

Comparing a fixed effects regression index with a matched model index on the same data does not compare alternative quality adjustment procedures (as do the studies presented in Chapter IV), for the quality adjustment procedures are the same. Instead, it compares alternative index number formulas. The formula for the fixed effects regression time dummy coefficient (d_1 in the above equation) is generally a geometric mean formula (see Chapter III, section D); the matched model index uses a normal index number formula (like the Tornqvist index or the Fisher index). Finding that the fixed effects regression index approximates the normal matched model calculation is just showing that these index numbers are close together.

If it were useful to do so, one could also refer to the typical statistical agency matched model procedure as a fixed effects model. The agencies define the cell so that only "small" quality deviations are permitted from one pricing period to the next (Chapter II). This avoids having too many empty cells in the second period, which will be the case if the product (as is the case for computers) is changing rapidly. As discussed in Chapter II, quality change problems arise (a) when unrecorded changes arise within those cells, and (b) when new cells appear and old ones disappear and these appearances and disappearances result in unmeasured price changes. Bias infects the index when the procedures used to account for quality changes are inadequate, and when price changes are missed.

With respect to the first effect (a), ADC's cells are defined on manufacturer's nomenclature. Nomenclature frequently hides changes in specifications, so what are in effect inadvertent direct comparisons of unlike computers occur. This is why agencies normally collect the characteristics, rather than relying on nomenclature (as discussed in Chapter II). With respect to the second factor (b), price changes from exits and entries are also missed in the fixed effects index, a point made above.

In summary, the ACD study is valuable is showing that frequent resampling and reweighting produces an index that declines in magnitude similar to a hedonic index, such as the PPI index for PC's (the interpretation of one of the authors, with which I do not disagree). What is still not resolved, however, is how much a FR&F index with conventional procedures differs from an index with hedonic quality adjustments *carried out on the same data used by ACD*. On the datasets of other authors who have examined this issue (the studies are reviewed in the body of Chapter IV), these indexes differ, they do not coincide.

Table 4.1

Alternative Semiconductor Price Indexes, 1982-88

	Average Annual Percentage Change
Chain Laspeyres	-21.9
Chain Paasche	-34.9
Chain Fisher	-28.7
Chain Tornqvist	-27.9
Mimic of PPI (delayed introduction)	- 8.5
Actual PPI	- 4.4

Source: Dulberger (1993), Table 3.7

Table 4.2

Price Indexes for Televisions
(August 1993=100)

	August 1997
Published CPI	86.8
Simulated CPI, with hedonic quality adjustment	86.4
Hedonic characteristic price index	79.6

Source: Moulton, LaFleur, and Moses (1999), tables 5 and 6

Table 4.3

Price Indexes for Computer Processors
(Average Annual Rates of Change, 1972-1984)

Matched model	-8.5%
Dummy variable	-19.2%
Characteristic price index	-17.3%
Hedonic Imputation	-19.5%

Source: Dulberger (1989), table 2.6

Table 4.4

Matched Model and Hedonic PC Computer Price Indexes Australia,
(Total Price Decline, April 2000-December 2001)

Matched model index, all models	-32%
Matched model index, overlap sample	-35%
Matched model index, with hedonic quality adjustments	-45%
Hedonic dummy variable index	-52%

Source: McKenzie (2002), Figure 1

Table 4.5

Hedonic and Matched Model Price Indexes for Japanese PC's, TV's, and Digital Cameras
(Average Annual Rates of Change, 1995 I to 1999 I-- except cameras: January 2000-December 2001)

Hedonic Indexes[a]	PC's	TV's	Cameras
Adjacent-month dummy variable	-45.1%	-10.4%	-22.0%
Price index for characteristics (monthly, weighted geometric mean formula)	-45.7%	-10.4%	-21.9%
FR&R Matched Model Indexes (monthly samples)[a]			
Chained Fisher	-42.7%	-18.8%	-27.8%
Geometric mean formula	-42.7%	-18.4%	NA[b]

Source:
Calculated from Okamoto and Sato (2001), charts 2 and 5, supplemented with additional information on the digital camera index from Masato Okamoto (January 26, 2003).

Notes:
a. Other hedonic indexes and other matched model indexes are also presented in these charts; not all agree with the indexes cited above, but the ones cited are either the preferred measures (based on index number theory and hedonic index principles), or do not differ from other, equally preferred measures.
b. Geometric mean not available. Chained Tornquist = - 27.8%

Table 4.6

Comparison of Matched Model and Hedonic Price Indexes for Computers in France
(2001-I to 2002-II, total change)

Matched model index	- 13.7 %
Hedonic imputation index	- 42.1%

Source: Evans (2002), table 9.

Table 4.7
Matched Model and Hedonic Indexes for Computers in the Netherlands
(Average Annual Rates of Change, January 1999 to January 2002)

	PCs	Notebooks	Servers
Matched model (FR&R) index	- 21.9 %	- 20.5 %	- 22.1 %
Hedonic single imputed index (weighted)[a]	- 26.2 %	- 23.3 %	- 25.7 %
Hedonic dummy variable index (unweighted)	- 32.5 %	- 25.5 %	- 27.3 %

Source: Van Mulligen (2003, tables A.8, A.9 and A.10)

/a The author also presents another imputation index, which he calls "hedonic imputation (using dummy variable index)." These indexes are closer to the matched model indexes, as discussed in Chapter III. Average annual rates of change for this index are, respectively for the three products listed above, - 24.3 %, - 21.2 %, and – 24.8%. As explained in Chapter III, using the dummy variable price index to impute price changes for entering and exiting computers understates their price change whenever the true price change for exits and entrants is greater than the price change for continuing models.

Table 4.8

**Alternative Hedonic Indexes and Matched Model Indexes.
UK Television Sets and Washing Machines**

	TV's, Jan-Dec 1998	Washing machines, Jan-Dec 1998
Matched-model index (Fisher formula)	- 12.7%	- 9.3%
Hedonic dummy variable index (equal-weight geometric mean formula)	- 10.5%	- 7.4%[a]
Hedonic characteristics price index (Fisher formula)	- 10.1	- 7.6

Notes: a) Dummy variable index computed from chaining together dummy variable indexes from adjacent-period regressions. Personal communication from Saeed Heravi, January 24, 2003. The dummy variable from a 12-month pooled regression (which was the estimate published in Silver and Heravi, 2003) was – 6.0%.

Sources:
TV's—computed from Silver and Heravi (2001, table 4, indexes labeled, respectively, Matched Fisher, Dummy variable semilog and Hedonic (SEHI) Fisher.
Washing machines—computed from Silver and Heravi (1999, table 1), supplemented by email communication with Saeed Heravi.

Table 4.9

**Within Group and Between Group Variability,
FR&R Matched Model and Hedonic Indexes**

Authors and product	range, matched model	range, hedonic	between group differences (max-min)
Okamato and Sato (2001)			
PCs	0.5	0.6	2.4 - 3.0
TVs	0.4	0.0	8.0 - 8.4
Cameras	na	0.1	5.8 - 5.9
Silver and Heravi (2001)			
TVs	a)	.8	2.2 – 2.6
Washing Machines	1.3[b]	1.7	1.7 – 1.9

a) Only Fisher index calculated.
b) Includes Fisher, Tornqvist, and geometric mean indexes.

Table 4.10

Analysis of FR&R Matched Model Index and Hedonic Index Comparisons

Study	Diff. Index?	FR&R?	est P diff?	est P adjust?	Weight, ent/exit?
Dulberger	yes	partial (annual)	yes; differs	yes (informal)	no
ACD	/a	yes (quarterly)	no	no	yes, no number
McKenzie	yes	yes (monthly)	no	no	no
Okamoto and Sato	yes	yes (monthly)	yes; differs	no	no
Evans	yes	yes (monthly)	no	no	no
Van Mulligen	yes	yes (monthly)	no	no	yes, 20% monthly
Silver and Heravi	yes	yes (monthly)	yes; differs	yes?	4 % monthly[b]

Key:
Diff. Index? = whether matched model index differs from hedonic index
FR&R? = whether used full FR&R structure, and frequency
Est P diff? = study estimated whether entering/exiting models had different price/quality from continuing
Est P adjust? = study estimated whether continuing models adjusted quickly to price/quality of entrants
Weight, ent/exit? = study estimated the weight of entering and exiting models

Notes:
a. Not computed explicitly.
b. Sales proportion of unmatched observations (Silver and Heravi, 2002, page F399).

Figure 4.1

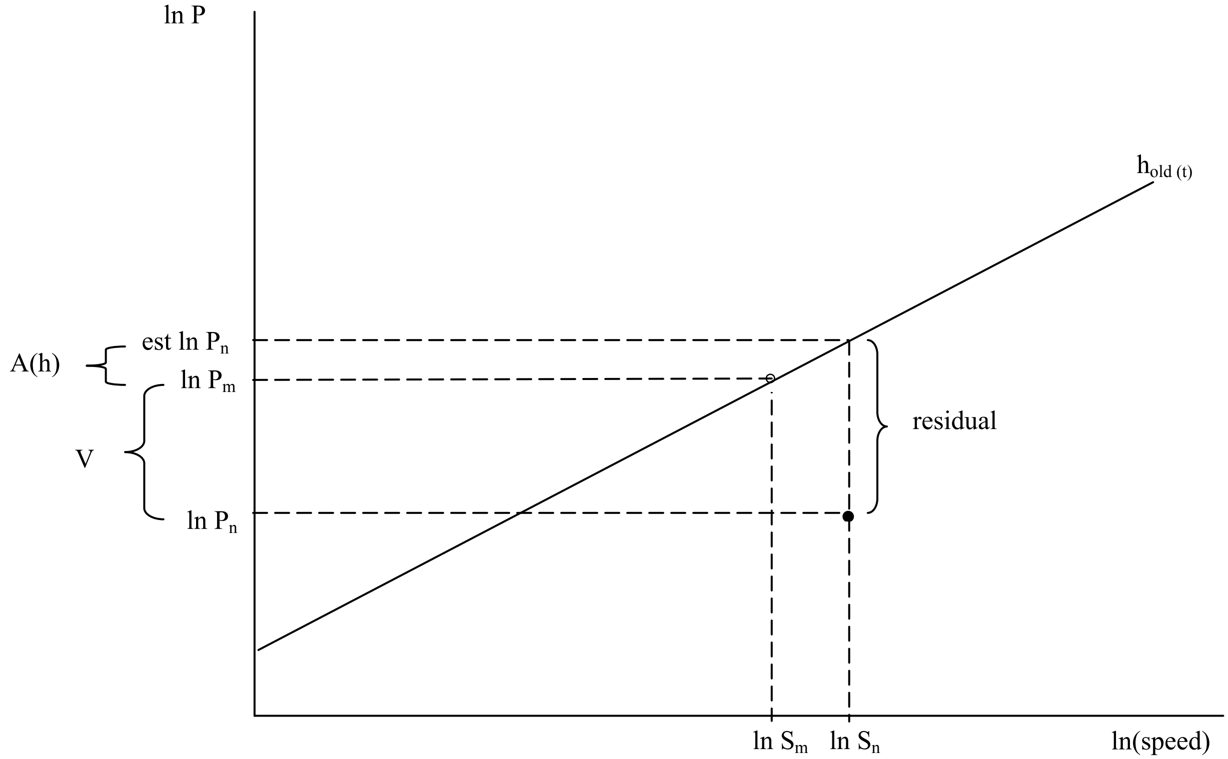


Figure 4.2

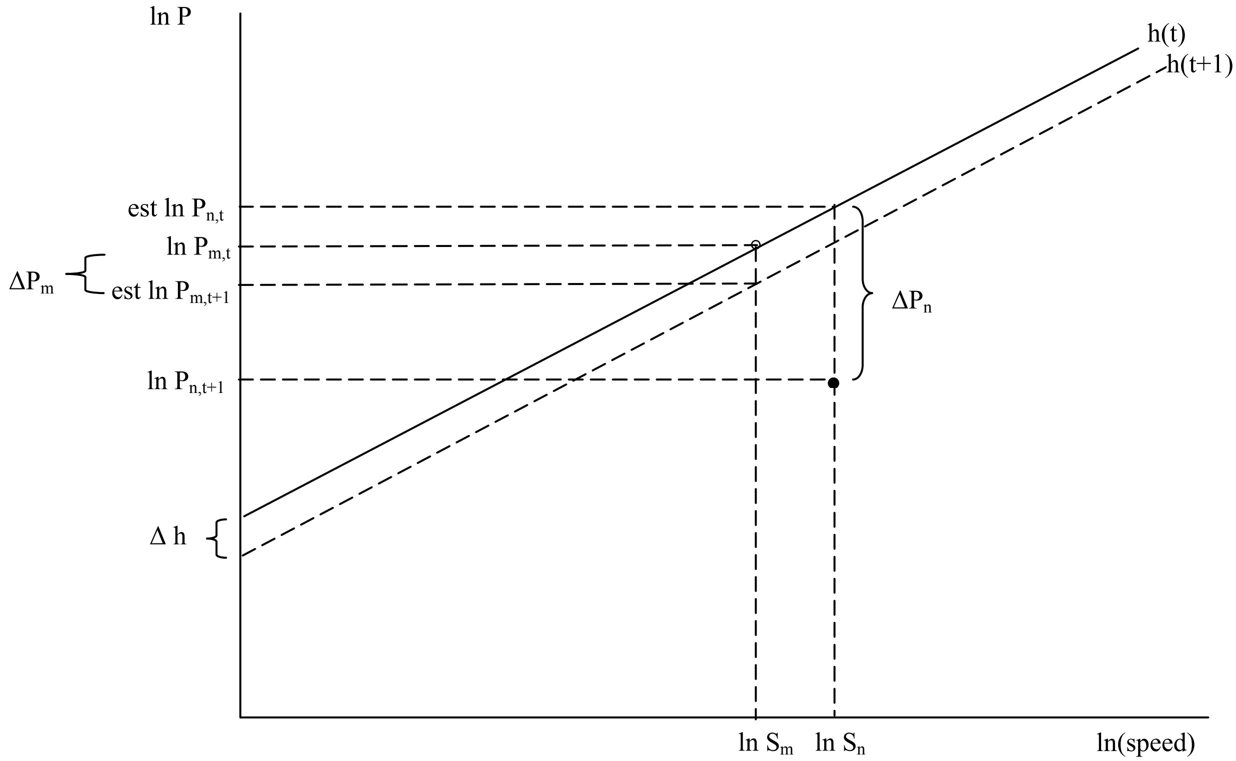

ln P
h(t)
h(t+1)
est ln $P_{n,t}$
ln $P_{m,t}$
ΔP_m
est ln $P_{m,t+1}$
ΔP_n
ln $P_{n,t+1}$
Δ h
ln S_m
ln S_n
ln(speed)

Figure 4.3

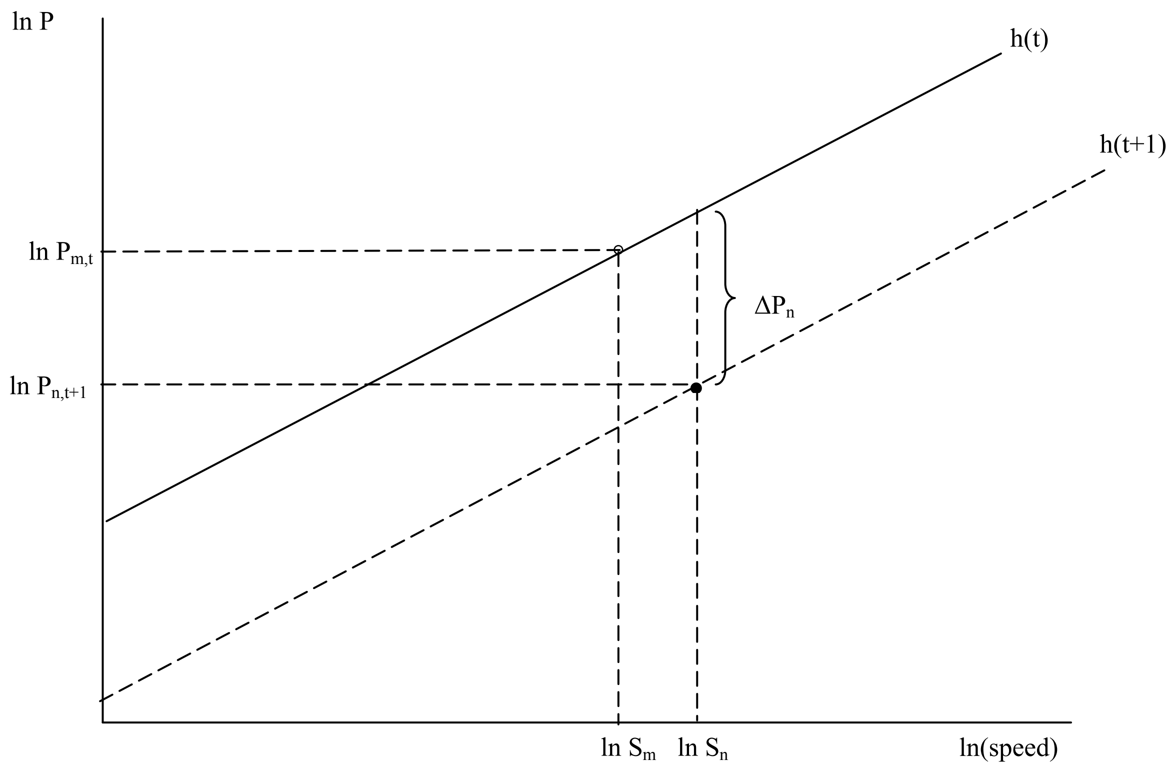
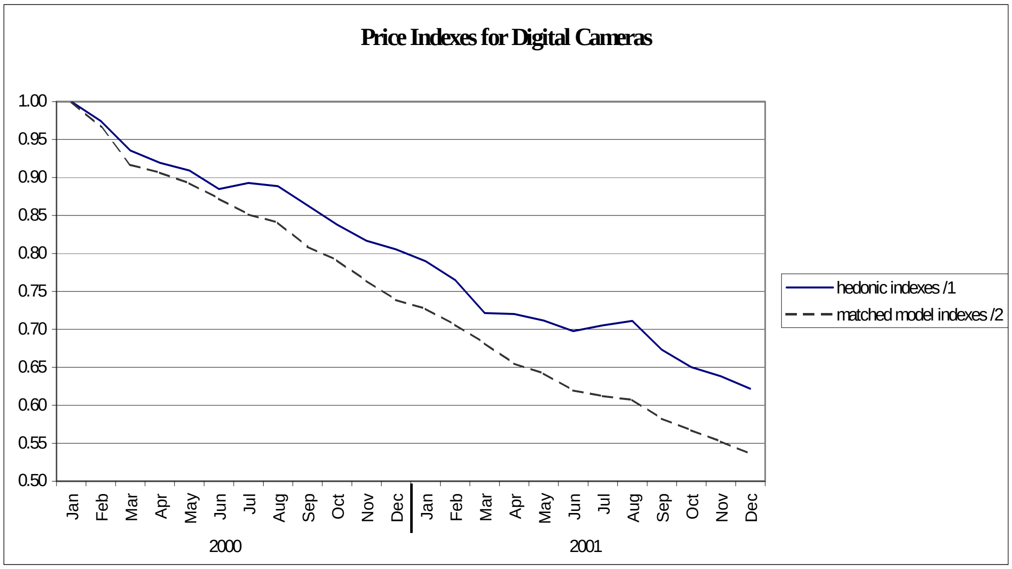


Figure 4.4



Source: Okamoto and Sato (2001), supplementary information provided by Masato Okamoto, January 26, 2003.
Notes:
/1 Graphs of two hedonic indexes (dummy variable index and characteristics price index)--indexes coincide for every month.
/2 Graphs of two matched model indexes (chained Fisher and chained Tornqvist)--indexes coincide for every month.