Hedonics and the Consumer Price Index.

Ariel Pakes, Harvard University and the N.B.E.R. [1]

May 11, 2005

In this paper I focus on issues that arise in incorporating hedonic price indices into the Consumer Price Index ("CPI"). Since this paper is being published in a volume commemorating Zvi Griliches' contributions, I begin by commenting on the relationship between my work (both in this paper and in its companion piece (Pakes [2003]), and Zvi's contributions to and views of hedonic price indices.

Zvi's 1961 article (Griliches [1961]) revived Court's [1939] view that hedonics could be used to ameliorate the "new goods" problems in price indices. Both argued that since newer models often had a greater amount of desirable characteristics, the difference between prices of the newer and older models should not entirely attributed to inflation. Both Griliches and Court suggested first estimating a surface by least squares which related prices to characteristics and the passage of time, and then using this estimated surface to derive estimates of price changes for products with constant characteristics.

Both in his writings and in conversations, Zvi was ever uneasy and cautious concerning the theoretical foundations of the hedonic surface he advocated be estimated. While he thought highly of the work on characteristics-based demand systems initiated by his friend Kelvin Lancaster (Lancaster [1971]), he was wary of and less convinced by the literature that went from a characteristics-based demand system to an equation for price as a function of characteristics. In the terminology of industrial organization economics, Zvi was worried about pricing equations that did not provide an adequate role for markups over (marginal) costs. In the absence of a rigorous and reasonable theoretical integration of markups into price equations, Zvi settled for a practical interpretation of hedonic price indices. An hedonic price index was simply an empirical summary of the relationship between price and characteristics.

Within the last decade characteristics-based models have been utilized as tools for the empirical analysis of differentiated product oligopolies (see Berry, Levinsohn and Pakes, 1995, "BLP"). These models typicall employ a Nash in prices (or Bertrand) equilibrium as a conceptual framework. The markup on a given model ends up being a complex function of the characteristics and costs of all marketed goods, and the distribution of consumer attributes. The projection of these markups onto the good's own characteristics – the "hedonic function" – is simply a "reduced form" whose coefficients do not necessarily obey any of the restrictions one usully associates with utility or cost functions. Nonetheless, as I will explain below, the hedonic function does

1

provide a way to partially control for new goods problems in price indices, just as Zvi had thought.

In the first section of the paper I explain the rationale underlying hedonic price indices. This section provides a verbal exposition and a small extension of topics treated in a more formal way in Pakes (2003). Section II expands on this by pointing out some of the properties of hedonic price indices, and showing how problems can arise in indices which do not possess those properties. In the remainder of the paper I discuss some of the most frequently cited "problems" with using hedonics as components of the CPI construction.

More specifically, I will first explain why a number of issues underlying these problems have already been "essentially" solved. This includes the question of what characteristis and what functional form should be used to estimate the hedonic surface, as well as how to produce an hedonic index within the real time constraints faced by the Bureau of Labor Statistics ("BLS"). What I mean by the words "essentially solved" is that procedures are either already in place, or are being put in place at the BLS, that should resolve these issues[1].

Most others issues that I am aware of are common to both hedonic and matched model indices. In this context I first discuss the problems caused by sample attrition. It is the greater ability of hedonic indices to handle sample attrition that is the most important reason favoring hedonic indices. This section also considers the reasons why both indices may not produce very "tight" upper bounds on the compensating variation we are after. These include our inability to capture the inframarginal rents that go to consumers who would purchase a good at all observed prices, and substitution biases. We do have some suggestions for mitigating these problems, but it is clear that more empirical work is needed to clarify their effects and to explore alternative correction procedures. Even so, however, it seems fairly clear that these issues cause more problems for the matched model than for hedonic indices, and so should not be thought of as a cause for slowing down the introduction of hedonics.

Finally there are some technical issues about how to actually construct the hedonic indices we use. Since it may be both politically and administratively

---

[1]Throughout I will condition the discussion on using the BLS's current data gathering arrangements. This, not because I endorse them, but rather because revolutionary changes in data gathering procedures are unlikely to be implemented in the near future.

difficult to change the way we construct hedonics after we introduce them, these technical issues ought probably be investigated prior to, or together with, a broader role out of hedonics.

Still the basic result is clear enough. With the improved technologies currently available it is both feasible and desirable to move many of the commodity group indices in the CPI from a matched model to a hedonic basis.

1   The Goal and Its Relationship to Hedonics.

I will assume that we are interested in finding an upper bound to the compensating variation. Of course the lower the upper bound the better (and we would like a least upper bound). Konus (1939) has shown how to find a lower bound when utility is defined in "goods" space. Pakes (2003) presents an analogous argument for characteristic based models.

The Hedonic Function

In characteristic based models

- Utility functions are defined on the characteristics of products (instead of on products per se), and

- The distribution of utility and cost functions, together with an equilibrium assumption, determine the relationship between prices and characteristics.

The hedonic function is a reduced form empirical summary of the relationship between a product's price and its characteristics. Formally it is the mathematical expectation of price conditional on characteristics and it can be estimated by regressing prices on characteristics (or, if one is worried about nonlinearlities in this expectation, it can be estimated by regressing price on a sufficiently rich set of functions of those characteristics[2]). The hedonic function summarizes how much a consumer has to pay to obtain different bundles of characteristics.

----

[2]In a standard Nash in prices, or Bertrand, equilibrium a product's price depends on the characteristics of its competitors' products, as well as on its own characteristics; see Pakes, 2003, for a discussion. The hedonic function therefore depends upon the expectation of competing products' characteristics conditional on own product characteristics. As a result the link between the hedonic function and the primitives determining behavior depends

The Hedonic Price Adjustment.

The hedonic price adjustment is the "Laspeyres" adjustment in characteristic space. I.e. it is

the increment in money income needed to enable the consumer to buy the same tuple of characteristics that the consumer bought in the base period in the comparison period.

To obtain an empirical estimate of the hedonic price adjustment we estimate the price of the good in the comparison period by evaluating the comparison period's hedonic function at the characteristic tuple bought in the base period. The compensating variation is obtained as the difference between this estimate and either the base period price, or the price that the base period's hedonic function predicts for that good's characteristics. If all goods are available in both periods, analogous logic to that used by Konus (1939), shows that this is an unbiased estimate of an upper bound to the compensating variation.

2   Two Properties of Konus-Type Bounds

Both Konus' original and the characteristics based analogue of our bound

1. are nonparametric, i.e. they do not require any assumption on the form of the individual's utility function, and

2. allow for arbitrary differences in utility functions among different individuals.

Differences in tastes cause differences in base period purchases among individuals. As a result there is a different bound for each individual. The aggregate index is formed as a weighted average of the individual bounds.

Throughout the discussion I will assume that any index we want to use is a sum of "nonparametric" bounds to the compensating variations of "different" consuming units.

---

on the details of the equilibrium process generating new products. Though models that compute equilibria to the dynamic oligopolistic interactions that generate new products are available (see, for e.g. Pakes and McGuire, 1994), to my knowledge they have not yet been used to investigate the nature of the reduced form relationship between price and own-product characteristics.

Two Implications of Our Requirements.

I want to point out two implications of this view. First once we allow individuals to differ in their income, demographics, and other determinants of tastes, all other bounds that I know of require strong, and I would say unreasonable, assumptions.

In particular, bounds that depend on the second derivative of the demand function having a single sign do not allow for arbitrary differences in utility functions among individuals; indeed I would argue that they typically require totally unrealistic distributional assumptions. To see this it suffices to consider the simplest of "vertical" models.

Let individual's differ in their income. If an individual with income y purchases the good in question its utility is determined by the amount of income it has for purchasing other goods, or $(y - p)$ where p is the price of the good, and the utility it derives from consuming the good, or K. In Cobb-Dougals form the individual's utility if it does purchase the good is[3]

$$(y - p)^\alpha K.$$

It the individual chooses not to purchase the good its utility is

$$y^\alpha.$$

It follows that individuals' with income above p/θ will consume the good, where $\theta = \theta(K, \alpha)$. Thus if demand is given by $D(p)$ and $G_y(\cdot)$ provides the distribution of income, we have

$$D(p) = [1 - G_y(p/\theta)].$$

Consequently

$$\frac{\partial D(p)}{\partial p} = -g_y(p/\theta)\theta^{-1}, \quad \text{and} \quad \frac{\partial^2 D(p)}{(\partial p)^2} = -g'_y(p/\theta)\theta^{-2},$$

where $g(\cdot)$ is the density of the income distribution and $g'(\cdot)$ is its derivative.

That is the demand curve will be *concave* before every mode of the income distribution *and convex* thereafter. The "moral" of this example is that once

---

[3]Actually, since individual choices are invariant to monotone transforms of the utility function, all we require is that utility be an individual specific monotone transform of the functional form that follows.

we allow for heterogeneity in tastes the convexity of the demand surface will depend on the distribution of preferences. Moreover the convexity is likely to change signs if the distribution of the consumer attributes that determine purchase patterns (like income) has at least one mode. Thus bounding arguments that depend on the convexity of the demand function being of one sign are not very appealing.

The second point I want to make concerns how we aggregate over individuals to produce a single index for the population as a whole. Currently the way we aggregate over individuals to produce a single index is determined by the way goods are sampled. There are many possible improvements which could make the aggregation procedure either reflect current choices better, or make the index easier to produce (either from a cost or from a time standpoint). However the issue of how we aggregate is independent of the issue of what index we are aggregating over. In particular it arises whether we use hedonic or matched model price relatives for individual goods.

3   Issues In Using Hedonics.

Different authors have raised different questions with respect to the use of hedonic indices. I am going to divide these issues into three groups, each of which will be treated in a separate subsection. The first set of issues addresses problems which I think have been effectively "solved". These are issues we know how to resolve though for one of them the BLS is not quite set up to implement the solutions yet. The second set of issues addresses problems which are in fact common to both matched model and hedonic indices. I will argue that we should expect hedonic indices to compensate for these problems at least as well as the matched model indices do. On the other hand there is little doubt that these are among the, if not *the* most important problems with the index, and it is disappointing that there is so little research on them. Finally there are a set of technical issues to be resolved on which among alternative possible ways of constructing hedonic indices is likely to produce an index with the most desirable properties.

3.1   Issues That Are Effectively Solved.

This subsection considers three problems which have effective solutions. It is the fact that these solutions exist that makes it feasible to use hedonics in

the construction of official price indices.

3.1.1 Data

The data required to use hedonic techniques are already gathered by the BLS. This is because

> *the matched model comparisons currently in use condition on product characteristics, just as hedonic comparisons do.*

The current data gathering procedure actually generates two lists of characteristics. An initial list of characteristics determines which goods the gatherer can sample from when sampling a new entry level item at an outlet. The data gatherer then makes a more detailed list of characteristics of the sampled good. It is the second list which determines which goods can be sampled from when the data gatherer (or his or her colleague) returns to the outlet to obtain a comparison period price to compare to the base period price of the entry level item.

Thus the second list of characteristics is the list of characteristics the BLS implicit conditions its matched model comparisons on. These lists

- have been used for years,

- have never been a source of contention, and

- are generally rich enough to base a hedonic analysis on.

If we used this list of characteristics for our hedonic index then the conditioning sets of hedonic and matched model comparisons would be *identical* *except* that matched model comparisons implicitly also conditions on the precise outlet sampled. The data gatherers, however, do provide a detailed characterization of the type of outlet (when there are chain stores the characterization provides the precise name of the chain, otherwise it is in the form of a categorization of store types). Since the characteristics of the sale (in contrast to the characteristics of the product sold) are deteminants of utility, and the type of outlet is such a characteristic, the information on outlet type should be incorporated into the hedonic analysis[4].

---

[4]Note that by conditioning on the outlet sampled the matched model index does not condition on all the conditions of the sale that matter to the consumer. For example, the consumer might care about the day of the week, or the quantities and price of the other

Omitted Characteristics.

If there are characteristics that consumers care about and that neither index conditions on, the expected effect of those omitted characteristics on both the matched model and hedonic price relatives of the goods that do survive until the comparison period is *identical*(see Pakes, 2003). Note that this implies that provided both indexes condition on the same characteristics any difference in the expectation of the two indices is due to solely to the differences in their treatment of the price changes of the goods that drop out (a point we focus on below).

As noted above there is still a bit of room for characteristics that are omitted from the hedonic analysis but conditioned on in the matched model analysis. First the outlet characterization available to the hedonic analyst may not be sufficiently rich. Also it is possible that the same data gatherer obtained the observation on price in both the comparison and the base period, and that this data gatherer held in memory characteristics that the first period good had but which were not listed in the list of characteristics.

If there are characteristics that are omitted from the hedonic but conditioned on in the matched model, it is only their independent variance, that is their residual variance after we condition on all included characteristics, that will have any effect on the comparison between the hedonic and the matched model index. Often product characteristics are highly correlated. Indeed, as a referee pointed out, in practice perhaps the most frequent set of characteristics that the hedonic does not condition on and the matched model does are characteristics which are available to the hedonic analyst but not used in hedonic predictions because they were not "significant" in the hedonic regression. That is they are left out because hedonic predictions that do not use them are thought to be more accurate than predictions that do (see the discussion of the trade off between bias and variance of indices in Pakes,2003). Since such characteristics do not change the R^2 of the regression by much, they should not have much of an effect on the resulting price relative.

The impact of any residual omissions of characteristics from the hedonic function on the difference between the hedonic and matched model price relatives depends on the details of how the hedonic is constructed. For the

goods sold, at the time of sale. Both the matched model and the hedonic analysis will always face a tradeoff between the costs of gathering characteristics and getting as good an approximation as possible to the determinants of utility.

main argument in the paper it suffices to note that if their is any difference it is likely to make the hedonic generate a larger compensating differential than the matched model index for goods that are available in both the reference and the comparison period (which are the only goods that we can get matched model price comparisons for)[5]. So if the matched model price comparisons provide an upper bound to the compensating differential for these goods, so should the hedonic. That is the possibility of omitted characteristics should not reverse the argument that the hedonic index provides an upper bound to the compensating variation.

3.1.2 Timeliness.

For the BLS to produce a hedonic index it needs to

- transfer data from the data gatherers' to a computer,

- estimate the hedonic function, and

- compute the hedonic index,

all within its monthly time constraints.

In the past the difficulty with producing a hedonic index of the form discussed here was that it could not be produced within the BLS's time constraints. As a result where the the BLS did use "hedonic-like" corrections they used them only on part of the data, and used a correction methodology that leads to a number of quite severe problems (see the discussion of the "incomplete hybrid" in Pakes, 2003).

There has, however, been a major change in data gathering procedures at the BLS, and this should enable the BLS to produce a proper hedonic index within their time constraints. Data gatherer's now carry hand held computers and download the data they gather directly onto the BLS's main computer after every work day. Given the downloaded data, it is easy to

---

[5]To see this let the hedonic function in period τ be $h_\tau(x)$ and the price be $p_\tau = h_\tau + \epsilon_\tau$. Then the difference between the matched model and hedonic price relatives is

$$p_{t+1}/p_t - h_{t+1}/p_t = \epsilon_{t+1}/(h_t + \epsilon_t).$$

Assume $E[\epsilon_{t+1}|\epsilon_t] = \rho\epsilon_t$ for a positive ρ. Then the fact that ϵ_t is mean zero implies that the expectation of this difference is negative, or that the the hedonic price relatives will, on average, be larger then the matched model price relatives.

automate the procedure that constructs a hedonic index (see the discussion below), and the automated procedure could produce a hedonic index virtually instantaneously.

Of course there will be some costs of getting hedonics up and running as; (i) the hand held computers must be programmed appropriately, (ii) the data gatherers will have to be trained to use them, and (iii) we will want to experiment some with the automated procedures before we embed them in the index. These are largely, however, "start-up" costs of transferring to hedonics, and do not impact on their longer term possiblities for producing a hedonic within the BLS's time constraints.

3.1.3 Functional Forms.

There is still a question of which functional form to use in estimation. The hedonic function is just a regression of costs and markups onto characteristics. This implies

- Any sufficiently rich functional form will do, and all sufficiently rich functional forms will generate approximately the same result,

- There are no constraints that can be used to constrain the form of the function to be similar in different time periods,

- There are no constraints on the coefficients at a given point of time.

A couple of comments are in order. Pakes (2003) provides a discussion of why it is unreasonable to assume that hedonic coefficient are stable over time and shows just how strongly the data reject the stability assumption in his PC example. Prices are determined by marginal costs and markups. Though marginal costs might not differ much over time, markups will; especially in "high tech" industries. In those industries products which are innovative and desired by consumers will earn the markups that justify the investments in their development. These markups provide incentives for entry into the part of characteristic space the innovative product opened up, and the entry should cause markups to fall. For this reason the statement that coefficients in hedonic regressions should not be expected to be the same over time is widely accepted by now.

The statement that there are no constraints across coefficients at a given point in time, in particular no sign constraints, seems to be more contentious.

It is not terribly hard to construct theoretical examples where "vertical" characteristics (characteristics that each consumer prefers more of) can have a negative sign in a hedonic regression (see for e.g., Erickson, 2004), and there are numerous empirical results with signs that are, at lest apparently, "wrong". Many of the empirical instances seem, however, to be a result of a misunderstanding of what a "vertical" characteristic is; so I would like to focus the discussion on this issue and show how it plays out in a real example due to Cockburn and Anis (1998).

When we use a characteristic model to provide an explanation of behavior we often have to go to a more complete characteristic model than the model the analyst who estimates hedonic functions can use. The more complete model has consumers producing utility from combining product characteristics with their own, individual-specific, attributes (see Pollack,1989, and the literature cited their for a deeper discussion). The productivity of different product characteristics in this formulation would generally differ with the consumer's attributes and the characteristics that utility is defined over (the "utility" characteristics") would be different than the characteristics of the product. In particular a given vertical utility characteristic might be produced with different inputs by consumers with different attributes.

Often the best the hedonic analyst can do is find some summary measures of the values of the utility characteristic produced by the product, and use this as a "product characteristic" in the hedonic analysis of input prices. However the logic that leads one to expect a positive effect of a vertical "characteristic" on price requires the assumption that we can order the inputs by the amount of the vertical utility characteristic they produce, and that this ordering is the same for *all consumers*. As the discussion indicates, there are many cases when this assumption is *not* satisfied, and then there is no reason to expect that a product which, say on average, produces more of the vertical characteristic than a competitor, will have a higher price.

An example, taken from Cockburn and Anis' (1998) study of drugs for the treatment of Rheumatoid Arthritis, will clarify this point. Assume people that are sick with this disease care about their ability to grip objects and their overall health (we include the health index because the drugs that treat Rheumatoid Arthritis can have toxic side effects). The drugs actually marketed are defined by their content of various chemicals, and the transformation from chemicals to "grip ability" and "overall health" varies by individual. We simplify and assume that their are two types of drugs that treat Rheumatoid Arthritis (type A and type B). The National Institute of

11

Health performed a series of tests and the results showed that for the vast majority of patients drug type A is as effective as drug B and is far less toxic. In particular the "toxicity" rating of drug A (measured as fraction of people who the drug causes serious harm to) is essentially zero. On the other hand Drug B is effective on most of the, say, 5% of the population that drug A does not help, but has a toxicity rating of .7.

Once the results of these tests are made public drug companies rush to produce different versions of the type A drug. Indeed companies keep entering into that part of the market until the expected discounted value of profits from marketing such a drug falls below its development costs. The large number of producers of type A drugs generates competition in that drug's market and this forces down mark-ups on drugs of type A. Of course the large number of patients consuming these drugs implies that the firms that produced it are still able to cover their development costs (even given the small mark-ups). The market for drugs of type B is too small to support more than one firm, so that producer sells its product at a "monopoly" price.

The marginal cost of production is similar for both types of drugs, so the hedonic function will be largely determined by differences in mark-ups. Since mark-ups are higher for the more toxic drug, we should expect the hedonic regression of price against efficacy and toxicity to pick up a strong positive coefficient on toxicity. This is precisely what Cockburn and Anis (1998) find. There is nothing "wrong" with his result, indeed standard economic arguments lead us to expect it. In particular it does not mean that the market prefers more toxic drugs; it just means that profits were such that entry drove down mark-ups more on the less toxic then on the more toxic drugs – not an unreasonable finding at all.

Hedonic regressions have been used in research for some time and they are often found to have coefficients which are "unstable" either over time or across markets, and which clash with the naive intuition that characteristics which are generally thought to be desirable should have positive coefficients. This intuition was formalized in a series of early models whose equilibrium implied that the "marginal willingness to pay for a characteristic equaled its marginal cost of production". I hope this discussion has made it amply clear that these models can be *very misleading*. The derivatives of a hedonic price function should not be interpreted as either willingness to pay derivatives or cost derivatives; rather they are formed from a complex equilibrium process and are not interpretable. What can be interpreted is the value of the regression function for existing tuples of characteristics; this provides an

estimate of the average current price for goods which are defined by those characteristic values.

One final point on functional forms. Often a double log functional form seems to fit the data. This can be estimated either by using non-linear least squares on the linear version of the variables, that is by choosing coefficients to minimize

$$\sum_j \left(p_j - A\Pi_k x_{j,k}^{\alpha_k} \right)^2,$$

or by applying the log transform and estimating the linear model

$$log[p_j] = logA + \sum_k \alpha_k log[x_{j,k}] + \epsilon_j.$$

If we use non-linear least squares the hedonic estimate of the value of the characteristics of a product is simply

$$\tilde{h}(x) = \tilde{A}\Pi_k x_{j,k}^{\tilde{\alpha}_k},$$

where a tilde over a variable denotes its estimate when estimated by non-linear least squares. The hedonic estimate of the value of the characteristics of the product when the log-log form is used to estimate the hedonic function is slightly different. In particular it must correct for the fact that when we transform back to the linear form the expectation of the disturbance is no longer zero so

$$\hat{h}(x) = exp[\hat{log}A]\Pi_j x_j^{\hat{\alpha}_j} E[exp(\epsilon)|x],$$

where the hats over variables indicate the coefficients estimated in the linear in logs version. If we do not make a correction for $E[exp(\epsilon)|x]$, we will obtain a biased estimate of prices. This point is also made by Berndt (1991) and can be quantitatively important.

So why use the linear in logs regression instead of non-linear least squares? There are typically two reasons. First it allows us to use the least squares formula to express the coefficient estimates as an analytic function of the data (in non linear least squares we must employ a search routine to find the coefficient estimates). As a result using the linear in logs form allows us to fully automate the process of obtaining the hedonic coefficient estimates; a property which is extremely desirable given the time constraints and the concern for impartiality in producing the CPI.

The second reason for going to the log-log form is that the disturbances in that form often behave as if they are identically distributed across products

(in particular their variance does not depend on x). If the log transform does produce identically distributed disturbances then $E[exp(\epsilon)|x] = E[exp(\epsilon)]$ and can be consistently estimated by $J^{-1}\sum_j exp(e_j)$, where e_j is the residual for product j produced by the log log regression. On the other hand the residuals in the log-log form need not be identically distributed, and if they are not then the prediction equation obtained by ignoring the dependence of $E[exp(\epsilon)|x]$ on x is also biased. So if one does go to the log log form some examination of the residuals for possible heteroscedasticity is advised. To my knowledge there has not been much investigation of heteroskedasticity in this context, and even less examination of its impact on the hedonic price estimates.

3.2 Problems That Still Need Attention.

All of the problems that still need attention that I am aware of are common to the matched model and hedonic indices. Here I focus on two of them; both of which could be addressed by research using existing tools. I begin by introducing each problem and comparing their impacts on the two indices. It is the differential impacts of these problems on matched model and hedonic indices that I believe leads to a preference for hedonic indexes, at least for component indices in industries with a lot of technological change. Nevertheless the fact that neither index truly "solves" these problems points to the need for research to determine the magnitude of the residual issues and what we can do about them.

3.2.1 Sample Attrition.

Bundles of characteristics available in the base period are often not available in the comparison period. When this occurs there is no comparison period price for the matched model comparison. We can, of course, use the hedonic function to predict a comparison period price for the good even if it is not available. However if the good is not in the second period choice set, the Konus rational for the hedonic compensation providing a lower bound to the compensating variation is suspect. Consequently, when a good disappears *both* indices have a conceptual problem.

The matched model index basically averages the price-changes of the goods that were drawn in the base period from a sampled outlet and were

available in the comparison period in the same outlet[6].

Since the price changes of the goods that survive to the comparison period are not a random sample of all price changes, the matched model index incurs a selection bias.

In many types of products the goods that are not available in the comparison period are disproportionately goods that have been obsoleted by newly entering goods[7]. I.e. they are goods whose characteristics tuples could only be sold in the comparison period if the good's price was reduced so much that it would have made the good unprofitable to market. Thus by selecting from the price changes of the goods that do survive, the matched model index is selecting from the right tail of the distribution or price changes, and this gives it a positive selection bias.

The hedonic predicts the reference period price for the good even when it is no longer on the shelves of the particular outlet sampled in that period. However without additional conditions, the compensation that the hedonically predicted price gives to the consumer need not insure the consumer its base period utility, and this violates the conditions for a Konus-type bound given earlier.

There are two alternative conditions which insure that the hedonic compensation for a good which is not available in the reference period is larger than the compensating variation. There is also an informal argument that suggests that we should be able to handle the problems in hedonics when neither of these conditions are met.

We begin with the formal conditions. That is when either $C1$ or $C2$ below are satisfied then the hedonic compensation can be shown to be greater than the compensating variation even if the good is not on the sampled store's shelves in the reference period.

---

[6]For more details on how this is done see Triplett, 2003.

[7]There are other reasons why the good might not be available; stockouts being one of them. Moreover if stockouts were the major reason leading to attrition and they indicated excess demand which was followed by price increases, we would tend to omit goods whose price were rising, producing the opposite bias in the index. The BLS data gatherers code goods have special codes for goods that are thought to be temporarily unavailable (expected to return to the shelves in future periods). This should give us some indication of the extent to which attrition could be caused by stockouts. For further discussion of the implications of this distinction see Erickson and Pakes, forthcoming.

- $C1$ (A gradient argument from Pakes, 2003). Let $N(x_1)$ be a neighborhood of x_1 and C^j be the period "j" choice set. If there are a sufficient number of goods in $N(x_1) \cap C^1$ and $h(\cdot)$ is smooth then if

$$sup_{\ x \in N(x_1) \cap C^2} \left[\frac{\partial h^1(x)}{\partial x}|_{x_1} - \frac{\partial h^2(x)}{\partial x}|_{x_1} \right] [x - x_1] \geq 0,$$

the consumer will prefer the hedonic compensation and the period two choice set to period 1 utility.

This condition insures that there is a good which is available in period 2, say x_2, that satisfies the condition that the utility change from choosing it rather than x_1 is larger then the utility change due to changing costs from $h^2(x_2)$ to $h^1(x_1)$. The requirement that there be goods in the neighborhood of x^1 in the initial period (some or all of which may have exited by the reference period), is what allows us to use gradients.

- $C2$ (Vertical Characteristics). Alternatively if the important characteristics of the product are vertical, so that all consumers prefer more of them, or if the only non-vertical difference between the product that disappears and a product which *is* in C_2 is the outlet of sale and we could bound compensation for those differences, then we can modify the hedonic estimate for the second period to give individuals who bought the exiting good the cost of a good which *is* available in the second period and is at least as valuable to the consumer as the good that exited.

Condition $C2$ can often be applied for "high-tech" goods, like computers, whose major characteristics are constantly improving.

When C1 or C2 can be applied sample attrition does not cause a problem for the validity of the hedonic bound. If neither condition is satisfied we must look for other ways of compensating individuals for price changes that might satisfy the bound. We note that sometimes the cause of our inability to satisfy either C1 or C2 is that the underlying utility function is not additively separability in the goods from the different commodity groups. When this occurs we sometimes can find a bound for the compensating differential needed to compensate a consumer for attrition by forming an index for a combined commodity group[8].

---

[8]The fact that the utility function is not likely to be additively seperable in the goods

An example might help here. Consider the price index for software and assume that a base period software product can not be found in the reference period. All the software sold in the reference period that can perform the functions of the base period product require more memory than did the base period product. Since there are no goods in period two with similar memory requirements to those of the base year product we cannot use our condition C1, and since there are no products with less memory requirements (memory requirements are a "bad") we cannot use our condition C2. However if we define utility on the ability to perform "computer/software functions" we can obtain a compensating variation for couples of computer-software purchases.

Consider a person whose base period computer is available in the reference period. If the base period computer did not have the memory required to run the new software, then the compensation required to keep the consumer at base period utility from its software/computer purchase is obtained as the difference in price between the new software and last year's price for the old software *plus* the difference in price between the *new* computer and last years price for the old computer. If the base period computer did have the memory required to run the software, then all we require for the compensating variation is the difference in price between the new and the old software.

To my knowledge there has not been any research on the extent to which a combination of C1 or C2 are satisfied in cases of sample attrition. Given that the absence of C1 or C2 when there is attrition is the major remaining conceptual problem in the argument for hedonics, we would like to know more about its empirical relevance. On the other hand there is a more general argument which suggests that attrition problems are only likely to be important in very particular cases.

For a good which exits to cause a problem for our bound it must have gone from profitable to unprofitable despite still being preferred by a significant share of consumers (at price $h^1(x)$). At least if commodity groups are defined to minimize problems caused by non-additive utility functions, this is an unlikely event. However it could occur if the good; (i) has major "horizontal" characteristics or characteristics which some consuming units prefers more of

---

from the different commodity groups is a more general problem with all the price indices currently in use. Though non of us believe the utility function is additive, a discussion of the rather dramatic changes in both the data generation process and the index construction process required to adress general non-additivities is beyond the scope of this paper. I consider non-additivities here only because in some instances it can interact with sample attrition in a way that invites more detailed analysis.

while others prefer less (thus invalidating the use of our $C2$), *and* (ii)does not have close substitutes in one of the two periods (thus invalidating the use of $C1$), *and* either (iii)

- has large fixed costs of production (since then a non-negligible fraction of the population could prefer the old good to the new goods without making the old good profitable to market)[9], or

- has experienced a marked change in production costs over the two periods, or

- has experienced a change in ownership, so that one firm now produces two close substitutes and finds that it is profitable to take one of them off the market.

One could easily argue that the BLS analyst should be able to spot instances where this combination of conditions is likely to have occurred and make some adjustment for it.

3.2.2 Distance Between Both Indices and The Least Upper Bound (Substitution Bias and New Goods).

So far our discussion has been aimed at producing an index which insures that the consumer is not *worse* off in the second period given first period income and the compensating variation determined by the price index. There are at least two reasons, however, why the compensating variations the indices generate might lead the consumer to be *better* off than the compensation needed to keep the consumer at base period utility.

- As described here neither index makes an adjustment for the "substitution effect" that enables a consumer to substitute from the base period good to another good in the reference period as a result of the difference in relative prices in the two periods (and in principle we should be considering substitution between as well as within commodity groups).

- Neither the hedonic nor the matched model make an adjustment for the inframarginal rents which accrue when a new good is first introduced.

---

[9] An example might be an old automobile (say a model "T" Ford) that antique car buffs might be willing to purchase at its old price (adjusted for inflation), but that would not sell sufficient quantity at that price to cover the sunk costs of producing it.

There has been a great deal of experimentation with alternative ways of combining price relatives in a given commodity group to form a commodity group index which might be less subject to substitution bias then a simple average of the price relatives for that commodity group. Indeed for most commodities in the CPI the BLS uses a geometric mean to compute its area commodity price indices, and then an arithmetic average to average over the area indices to produce an overall index for the commodity. The use of the geometric instead of arithmetic mean in forming the area commodity group indices is to mitigate substitution biases. To my knowledge there has been no discussion of whether we should treat substitution differently were we to use a hedonic index, and some research on this question is needed (see also the discussion in section 3.3).

There are technical problems associated with incorporating the inframarginal rents from new goods into the index, and they can not be fully accomodated in either the matched model or hedonic indices. The logic of the Konus argument, whether applied in goods or characteristic space, never allows us to account for the inframarginal rents that go to consumers who preferred a new good to all existing goods at the introductory price (it only seeks to measure the change in price of goods that were marketed in the base period). An implication of this fact is that neither the hedonic nor the matched model index has any correction for increases in either the the variety, or for the "quality" of the goods marketed *per se*. The only increase that is recorded as a result of the appearance of new (and/or improved) goods is that caused by the fall in the price of these new varieties or higher quality goods over time.

An example will illustrate just how problematic this omission can be. Consider a new good that is launched at an "introductory" low price to induce consumers to experiment and spread information on the benefits of the new good. By period two the information has diffused throughout the population and the producer lets the new good's price rise to a level consistent with a full information Nash in prices (or Bertrand) equilibrium. If the good is sampled as an entry level item in period one, the only impact of the good on the CPI is the increase in price recorded in period 2. Assume that the rest of the choice set remains unchanged between period zero and two. Then the new good increases the computed CPI between periods 0 and 2 even though the utility of consumers who buy the good in period 2 has actually increased, and no consumer's utility has decreased.

19

The basic problem here is that there is no known way of obtaining estimates of compensating variation that capture the inframarginal rents to new goods that does not involve estimating the distribution of utility functions. Indeed we should note that the problem would not be entirely resolved even if we were able to base our indices on an estimated distribution of utility functions; at least not without imposing a priori restrictions on the form of those utility functions.

To see this assume we were able to estimate the distribution of utility functions nonparametrically (so we would not have to impose any *a priori* functional form restrictions). If the analysis is done in product space what we do is nonparametrically analyze the relationship between quantity demanded, price, and various household attributes (including income). Since there is simply no information in the data on the inframarginal rents of consumers who bought the good at *all observed prices*, there would be no estimates of utility gains to consumers who bought the good at the highest observed price. Of course putting enough *a priori* restrictions on functional forms would allow us to obtain estimates of the average inframarginal rents to consumers who bought the good at all observed prices, but then the estimates are determined by the imposed restrictions, not by the data.

On the other hand if we were to estimate demand directly, the "introductory pricing" problem explained in the last paragraph would no longer creep into the index; that is we would only not be able to estimate inframarginal rents above the highest observed price (which in this example is not the introductory price). Moreover even the problem caused by inframarginal rents at the highest price offered is attenuated somewhat when the demand analysis is done in characteristic, rather than in product, space (as in BLP, 1995). When demand is estimated as a function of product characteristics and household attributes, we can find the highest reservation price ever associated with a given tuple of characteristics, and we only can not estimate inframarginal rents for individuals who bought at that price.

Of course obtaining demand estimates is a difficult task that, at least with current technology and data, requires a number of questionable assumptions. As a result we should probably not expect any of the federal agencies to base their price index computations on demand estimates any time soon. Still there is good reason to push research on just how much of a problem our inability to measure the inframarginal benefits to initial purchasers of new goods is, and which markets those benefits are concentrated in.

There are at least two avenues to pursue here.

20

- Follow introductory pricing patterns in different industries, and attempt to adjust the index for initial increases in new goods prices in the industries where we find them.

- Attempt to construct experimental indices based on explicit estimation of demand curves, and then focus on adjustments for the product groups where the difference between the experimental and the actual indices are particularly large.

Though, for the reasons mentioned above, we are unlikely to be able to solve the "new goods" problem entirely, there should be a number of ways in which, with a bit more effort and care, we can ameliorate its effects on the index.

3.3   Practical Issues.

This section considers practical questions that will need to be resolved before we incorporate hedonics in our indices for different commodity groups. The discussion will be brief, as most of the issues can not really be resolved without more empirical research, and the preferred resolution may well differ for different commodity groups.

3.3.1   The Form of Hedonic Indices.

At least two related questions need to be answered.

- Should we construct individual hedonic price relatives by dividing the hedonic estimate of the reference period price by the actual base period price, or by the hedonic estimate of that price?

- Should we use an arithmetic or a geometric mean to aggregate the estimates of the hedonic price relatives into an index for the commodity group?

Recall that
$$p_{t,i} = h^t(x_i) + \epsilon_{t,i}.$$

Letting a "hat" over a variable indicate its estimated value, the first question is whether we should form hedonic price relatives as

$$I^h(x_i) = \hat{h}^{t+1}(x_i)/\hat{h}^t(x_i) \quad \text{or as} \quad I^p(x_i) = \hat{h}^{t+1}(x_i)/p_{t,i}.$$

21

There is a conceptual issue here. $I^h(x_i)$ tries to compare the average price of the listed characteristics in the two periods, while $I^p(x_i)$ tries to compare the price of the actual good sampled in the base period to the price of a randomly chosen good with the same listed characteristics in the reference period. One can argue either case conceptually and which of the arguments dominates depends to a large extent on; (i) the source of the disturbance in the hedonic function, and on (ii) whether one wants to measure the change in the average price at a given vector of characteristics or the average of the price changes for that vector of characteristics. What is true is that the expectation of $I^p(x_i)$ is greater then that of $I^h(x_i)$, so measurred inflation is likely to be higher when using the $I^p(x_i)$ index (this follows from Jensen's inequality, see Erickson and Pakes, 2005).

Practically the extent of the differences between the two indices for any given commodity group will depend on the variance of the ϵ in the hedonic function and the precision of the predictions made using that function. The BLS's current hedonic like procedures produce hedonic regressions with very good fits (R^2's are typically well above ninety percent) and quite accurate predictions (see below). So there is the possibility that the various indices do not produce results that are very different from one another. On the other hand, as discussed in section 3.1.2, the procedures for estimating hedonic functions are currently being modified to enable the production of timely hedonic indices, and the fit resulting from the new procedures may be different. In particular, as mentionned by a referee, hedonics are now more likely to be estimated from retail store data, and fits from this type of data are may well be worse. Either way, more research on the empirical effects of going with one or the other indices is needed before making a decision on which index seems appropriate for which commodity group.

As noted the issue of whether to use arithmetic or geometric averages has been debated extensively in the context of matched model indices. Geometric averages allow consumers to substitute in a particular parametric way in response to changes in relative prices before computing compensating variations. This does, however, violate the non-parametric spirit of Konus' bounds arguments (be they in product or characteristic space). Also it will change the way the variance in the hedonic's predictions affects the comparison between the index based on the I^p price relatives and that based on the I^h price relatives. This is another point at which more research is needed.

3.3.2 Variance and Hybrid Indices.

There are two sources of variance in the hedonic.

- Sampling variance or variance caused by the fact that the sample average of price changes does not equal the population average of those changes.

- Estimation variance or variance caused by the fact that the estimate of the hedonic function at any particular characteristic tuple will contain estimation error.

Sampling variance is present in both the matched model and in the hedonic index but is likely to be higher in the matched model index because sample attrition yields smaller samples for that index[10]. If we use the extreme assumptions that p_{t+1}/p_t are the price relatives we wish to measure, then there is no estimation error in the mathced model price relatives. Of course this "extreme" assumption implicitly assumes the data gatherer can condition on all relevant characteristics, an event which we noted earlier is highly unlikely, and that we are interested in the average of the price changes of goods with those characteritics rather than the change in price of the average good with those characteristics. Still even under less extreme assumptions it may well be the case that the mathced model price relatives have lower estimation variance than the hedonically estimated price relatives.

If there is less variance in the matched model price relatives then we might want to use them for the goods that do survive, and the hedonic price relatives for those goods that do not survive. This "hybrid" index uses the variance reduction properties of the matched model price relatives when they are available, and uses the the hedonic price relatives to correct for the selection biases in the availability of the matched model relatives. This is the index Pakes (2003) labels the complete hybrid index.

Though I think we should experiment with the complete hybrid, it should be noted that it does not get rid of all the selection bias. To the extent that ϵ does in fact correspond to omitted characteristics, and selection is based on those characteristics, then the complete hybrid contains a term equal to $E[\epsilon_{t+1}|p_t, survival]$ times the probability of survival (and no compensating

---

[10]The fact that one sample is selected and the other is not, implies that the variance per draw is likely to be different in the matched model and the hedonic samples, but their is no obvious ordering on which one will be larger.

term for the goods that did not survive). So when experimenting with the complete hybrid we should consider both likely biases and variances in the estimate.

There are at least two other procedures which might reduce the variance of the hedonic predictions. First we could instruct data gatherers to gather price information on a larger number of items at each outlet sampled then the number of items that will actually be used to form price relatives. The extra items, if they are found, could be used to increase the precision of the estimates of the hedonic surface, but, to maintain the integrity of the original sampling scheme, would not be used to form actual price relatives. We note that a major cost in increasing the BLS's sample size are the "fixed costs" of getting permission to obtain price quotes at a store and transporting the data gatherer to it. As a result this way of increasing sample size should be relatively cost effective.

The second procedure that might be used to reduce the variance of the hedonic predictions is to augment the sampled data with data downloaded from the internet (or from other sources) and use the augmented data to estimate the hedonic function. The BLS currently uses data downloaded from the internet to estimate the hedonic function used in the "hedonic-like" indices that are currently in use. The problem with their current procedure is that they are having difficulty in downloading the data and estimating the hedonic function within their monthly time constraints. The suggestions in this paper and in Pakes (2003) make it much easier to estimate the ehdonic, so it might also be time to revisit this idea.

4   Conclusion.

The timing of a broader role out of hedonics at the BLS should probably only depend on the speed at which the BLS can train their data gatherer's and analysts to use the procedures that are required to implement hedonics. In the meantime two types of research projects would seem worthwhile. The first is to take particular commodity groups for which the BLS has stored historical data on characteristics and/or can obtain such data in some other way, and use it to experiment with hedonic techniques. The experiments should evaluate different forms of hedonic indices (as discussed in section 3.3), and check for the extent to which the sample attrition adjustment implicit in hedonics is valid (as discussed in 3.2.1). Second, there should be some

attempt to look across commodity groups to see which of them might benefit most from a more detailed analysis then hedonics provides. This should include studies of introductory pricing patterns, as well as studies that form indices from estimated demand systems (as discussed in 3.2.2).

References.

Berry, S., J. Levinsohn, and A.Pakes (1995):"Automobile Prices in Market Equilibrium" *Econometrica*, Vol. 63, No. 4, pp.841-890.

Berndt, E (1991); "The Practice of Econometrics". Reading, Mass: Addison-Wesley.

Cockburn I. and A. Anis (1998); "Hedonic Analysis of Arthritis Drugs", N.B.E.R. Working paper, 6574, Cambridge, Mass.

Court, Andrew. "Hedonic Price Indexes with Automotive Examples", in *The dynamics of automobile demand*, General Motors Corporation, 1939, pp. 99-117.

Erickson, T. (2004); "On Incorrect Signs in Hedonic Regressions", *mimeo* Bureau of Labor Statistics.

Erickson, T. and A. Pakes (2005); " An Empirical Comparison of Quality Adjusted Price Indices for TV's" *mimeo* Bureau of Labor Statistics.

Griliches, Zvi. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change", in *The price statistics of the federal government*, New York, National Bureau of Economic Research, 1961, pp. 173-196.

Konus, A. (1924); "The Problem of the True Cost of Living Index" translated in *Econometrica* 7, (1939), January, pp.10-29.

Lancaster,K.(1971), *Consumer Demand: A New Approach*, Columbia University Press, New York, New York.

Pakes, A. and P. McGuire (1994), "Computing Markov Perfect Equilibria; Numerical Implications of a Dynamic Differentiated Product Model" *RAND Journal of Economics*, vol. 25, no.4, pp 555-589.

Pakes (2003); "A Reconsideration of Hedonic Price Indices With An Application to PC's" *American Economic Review*, December, pp 1578-1596. .

Pollak, R. (1989): *The Theory of the Cost-of-Living Index*, Oxford University Press.

Triplett, Jack. "The OECD Hedonic Handbook." Mimeo, *The Brookings Institution*, forthcoming.