Price Indexes for Drugs:

A Review of the Issues*

Ana Aizcorbe**

Nicole Nestoriak***

* We thank Sarah Pack and Danit Kanal for research assistance and member of our internal drug price index reading group for helpful comments and discussion: Marissa Crawford, Harvey Davis, Matthew Donahoe, Bonnie Retus, Carol Robbins, Adam Shapiro and Shelly Smith. The views expressed in this paper are solely those of the authors and not necessarily those of their respective agencies.
** Ana.Aizcorbe@bea.gov; Office of the Chief Economist, Bureau of Economic Analysis.
*** Nestoriak.Nicole@bls.gov; Compensation Research and Program Development Group, Bureau of Labor Statistics.

1. Introduction

Price indexes provide a way to summarize changes in prices of individual goods and services using an aggregate statistic. An important use of these indexes is to decompose changes in spending into price and quantity components. This is the role that price indexes play in the National Income and Product Accounts to obtain measures of real output and productivity. Price indexes are also used in the National Health Expenditure Accounts to provide information on the drivers of spending growth in the nation's health care sector. More broadly, health economists have used similar decompositions to inform policy debates about which levers may be used to contain cost growth (Merlis 2000).

Numerically, most price indexes can be expressed as functions of weighted averages of price change; many of the measurement issues discussed in the literature relate to which particular formula and weights is most appropriate in constructing the index.[1] There has been significant research into these issues and, indeed, the Bureau of Labor Statistics' official statistics have undergone substantial improvements in recent decades owing to research that pointed to deficiencies in existing indexes and provided new methods to improve those measures.

Beyond weighting issues, there is the problem that some price changes are accompanied by changes in the quality of goods. Ideally, one would like a "constant-quality price index" that would allocate improvements in goods as an increase in quantity consumed, not price paid. Because goods are typically multifaceted and have many attributes that are valued by consumers, the measurement of "quality" is typically summarized using consumers' valuations of the attributes of the goods. For many goods (like computers), these valuations are inferred from the prices that consumers pay for them. For medical care, it is widely understood that the prices patients pay for goods and services will typically not reflect how much they value them. This presents a major obstacle in measuring changes in the quality of medical care using techniques that have been applied successfully in other industries.

Related to these difficulties in valuing quality change is the broader problem that price indexes for medical care do not have a clear link to patients' welfare. This is because those inferences usually rely on a cost of living index (COLI) interpretation for

[1] See Schultze and Mackie (2002) for a recent review of the issues

the indexes. The nature of health-related decisions is such that this interpretation is strained, at best, for those transactions: Doctors play a key role in prescribing medical care—raising potential principal agent issues—and most patients use health insurance to cover at least part of their payments for medical services—raising moral hazard issues. Finally, the simple static utility maximization paradigm underlying COLI theory does not align well with how decisions for medical care are actually made.[2]

We therefore focus on the measurement issues, how the indexes are constructed, and how they may be used to decompose the growth in spending into price and quantity components. We do this for one of the major inputs in the provision of medical care: prescription drugs. While one ultimately wants to measure the output of the health sector as the marginal improvement to health status from all goods and services (Abraham and Mackie 2005), there are nonetheless important uses of price indexes for the individual inputs, such as measuring productivity growth for the drug industry and parsing out the drivers of growth in spending.

The chapter is organized as follows. We begin with a discussion of two definitional issues that turn out to be numerically important. Section 3 discusses the relative merits of different aggregation methods as they relate to the drug industry. A review of what is known about the issue of quality change is given in Section 4 and Section 5 concludes with a summary of the issues.

2. Definitional Issues

We begin with a discussion of how the "product" provided by the drug industry should be defined and how quantity and price should be measured. The particular price that one pays for a drug depends importantly on the attributes of the drug: for example, active ingredient (sometimes called the "molecule"), strength (e.g. 25mg), and dosage form (e.g. tablet). Which of these attributes define the "product?" Another issue that turns out to matter is how one defines the unit of quantity and, hence, the price (for example, should it be price per day or price per prescription?).

Defining the product: Drug vs. Molecure

[2] See Cutler et. al. (1998) for an alternative paradigm that better takes into account the dynamic nature of medical care and the role of insurance payments.

How one defines the "product" or "output" provided by the drug industry has important implications for how changes in spending are split out into price versus quantity components.

An important issue in this regard is how price indexes should handle the entry of generic drugs: should branded and generic versions of the same drug be considered the same or separate drugs? One landmark contribution of this literature was the demonstration that this distinction is numerically important for several prominent medications: Berndt, Cockburn and Griliches (1996) for antidepressants and Griliches and Cockburn (1994) for antibiotics.

For price index purposes, one wants to define a homogeneous product, so that tracking its price over time does not reflect any changes in the good's attributes. The issue is very similar to the problem of defining the market in antitrust cases: should the market be defined as aluminum foil or as all wrapping materials—foil, wax paper, saran wrap, etc. The key issue is how substitutable consumers find these goods.

There are two commonly used product definitions for drugs. One approach is to use the NDC code (National Drug Code, a 10-digit code that gives information on the attributes of the drug). An example of the level of granularity is a prescription for Lipitor, 20mg, 30 tablets (NDC 0071-0156-23). An alternate definition treats drugs with the same active ingredient as identical goods: this uses the GPI code (Generic Product Identifier, a 14-digit code that contains 7 pairs of digits) to identify drugs that use the same active ingredient and strength. For Lipitor, atorvastatin is the active ingredient and GPI 39400010100320 is the GPI code for the 20mg version of this drug. Once generic versions enter the market, the 20mg versions of atorvastatin will share this GPI code with the branded version, Lipitor. By contrast, a generic drug would not share Lipitor's NDC code, but would have a unique code of its own.

The important distinction between the NDC and GPI definitions lies in their treatment of generic goods. Why does it matter? Consider a simple example where a branded antidepressant sells for $1 a day and its price stays constant. At some point, the branded drug loses patent protection and a generic version is introduced at 70 cents per day, with its price also remaining constant. As patients switch to the less-expensive generic drug, overall revenues received by the drug industry fall. Because price indexes are typically some function of weighted averages of price change, a price index that

considers the two drugs as separate products (i.e., using the NDC definition) will show no price change because prices of the individual drugs did not change. This means that the reduction in revenues will be attributed to a drop in quantities, even if the number of patients or prescriptions did not change. At the other extreme, one can define the branded and generic versions as the same product (i.e., use of the GPI definition) and define the price as the *average* revenue per day of prescription, for example. With this definition, the decline in revenues will translate into a decline in the average revenue, or the price.

[CHART 1]

The substitutability between the branded and generic versions of a drug is the key issue. To the extent that patients only care about the active ingredient, the GPI definition, using the molecule definition, is the correct one. One caveat, however, is that the inert ingredients are often different for branded and generic drugs. As these inert ingredients often involve different side effects, many patients may well view the branded and generic versions of a drug as different goods and the GPI definition would not be appropriate.

We illustrate the numerical importance of this issue using a dataset containing drug claims submitted to 12 HMO plans from 2003:1-2005:4. The data are from Pharmetrics and are described in more detail in Aizcorbe and Nestoriak (2008). We use 200 million claims for oral medications that were purchased at retail pharmacies by patients covered by these HMOs.

In our data, prices for generic drugs are, on average, 30% lower than the prices of their branded counterparts, suggesting that how one handles switches from branded to generic drugs could be numerically important. (We obtained this estimate from a regression using data from 2003 that explained 82% of the variation in NDC prices using dummy variables for the molecule, strength, dosage form, manufacturer, type of health insurance plan, and branded versus generic status of the drug.)

Following the literature, we constructed price indexes that represent the two extremes, one based on the NDC definition and the other on the GPI definition. We found that the NDC-based price indexes grew about twice as fast as the molecule-based index: 2.6 versus 1.2 compound annual growth rates (CAGR) over 2003:1 to 2005:4.

This difference is large, and reflects the influx of generics into the market over this period. In our sample, the number of prescriptions filled for generic drugs grew from about 58% of all oral prescriptions in 2003 to over 60% in 2005.

[TABLE 1]

Exercises like this that compare the two extremes demonstrate the potential importance of the issue. The literature has provided some methods for better folding generics into price indexes but some problems still remain. The methods that have been proposed thus far tend to rely on consumer optimization problems and COLI theory: Fisher and Griliches (1995) and Griliches and Cockburn (1995) rely on estimates of reservation prices to fold in the new generic goods and Feenstra (1994) uses estimates of elasticities of substitution in an exact price index. Even if these methods could, in principle, provide suitable first-order approximations to the problem, subsequent studies that attempted to implement these methods had difficulties:

> …. the use of econometric methods in constructing price indexes that incorporate the effects of new goods requires considerably more experimentation, perhaps with other data sets and families of products, and with specifications that include non-price factors affecting demand functions. .. .Future research should focus on the conditions under which the Feenstra, the Griliches and Cockburn or some other method is more likely to yield robust and plausible findings. (P. 263, Berndt, Kyle, and Ling 2003).

For now, we note that price indexes defined on a GPI basis are likely to show slower price growth than those defined on an NDC basis. Because the sensitivity of price indexes to different definitions often depends on whether the good is defined on an NDC or GPI basis, we will report both sets of indexes in this discussion.

Defining units of "quantity"

How one defines the unit of measurement also has empirical implications for price measurement. Among the definitions for price that are typically used are 1) price per day of treatment, 2) price per prescription, and 3) price per package. The choice of definition is often influenced by the available data. Price per day of treatment is a widely used definition, used both in studies of cost decompositions and hedonic studies, but

requires information on the number of days of treatment associated with each prescription, as found in claims data. Price per prescription is the basis for some IMS statistics (e.g., IMS National Prescription Audit); it also underlies the Consumer Price Index for prescription drugs and has been used in various decomposition studies. Of these two, health economists typically view the price per day as the superior choice (Crown, Ling and Berndt 2002) because price per prescription confounds changes in utilization and price (increases in the number of days per prescription are represented as an increase in price per prescription, for example).

Finally, other IMS data are reported at the "pack" level (e.g., a particular container with a specified number of tablets sold to pharmacies) and this price per package has also been used to represent price (e.g., IMS National Sales Perspectives, which tracks sales from wholesalers to pharmacies and other outlets). This is the definition underlying the Producer Price Index for drugs and the price index used in the pharmaceutical components of the Federal Reserve Board's Industrial Production Index. Because tracking the price of each package holds constant the number of medications (e.g., tablets) in the container, tracking changes in the price per package is tantamount to tracking changes in the "price per tablet."

Our data allow us to construct the three measures and compare them. For this comparison, we use only claims where all three pieces of information are available. Consistent with the discussion in Merlis (2000), price indexes based on price per day tend to show slower price growth than those based on price per prescription. This holds true regardless of which formula is used (Fisher price index or Laspeyres) or which product definition (NDC or GPI). In our sample, the differences are large: using the NDC definition, indexes based on price per prescription grow nearly twice as fast as those based on price per day; the differences are even larger using the GPI definition. Indexes based on price per package also grow faster than the preferred price per day definition, but the differences are less pronounced.

[TABLE 2]

It is troublesome that changing the units of measurement can have such an impact on measured price growth. To the extent that price per day of treatment is the preferred

definition, it is also unfortunate that the necessary data to measure it this way are not as readily available as data on number of prescription or packages sold.

3. Aggregation Issues

Once the product and unit of measurement are defined, one needs a formula to aggregate price changes across the individual products to obtain an aggregate statistic. The best practice method is to construct chained Fisher price indexes. The Fisher formula is a superlative index number formula that has been shown to be superior to other aggregation formulas (Diewert 1992). As discussed below, "chaining" indexes provides a way to bring new goods into the indexes more rapidly and, thus, more closely track the composition of goods sold in the market. We also discuss the Laspeyres index, as it is often used in official price indexes and cost decompositions.

Price Indexes

Price indexes provide a way to measure aggregate price change over some period by comparing the cost of purchasing a market basket at different points in time. The simplest formula is the familiar Laspeyres index, which is usually written:

$$I^{L}_{0,1} = [\Sigma_i P_{i,1} Q_{i,0} / \Sigma_i P_{i,0} Q_{i,0}] \quad (1)$$

where 0 and 1 denote two periods in time, a base and current period, respectively, and i indexes goods that are sold in both periods. The Laspeyres tracks the cost of buying the $Q_{i,0}$ basket at period 0 prices with the cost of buying it at period 1 prices. The index can also be written as a weighted average of price change:

$$I^{L}_{0,1} = \Sigma_i w_{i,0} P_{i,1} / P_{i,0} \quad (2)$$

where the weights, $w_{i,0}$, are the base period expenditure shares and the price relatives, $P_{i,1} / P_{i,0}$ measure the price changes for individual drugs. The weights, or shares, are often called "relative importances" and have been the focus of much of the work in the literature. Written this way, it is easy to see that products in the base period market basket are only included in the index if they are sold in both periods (i.e., if one observes both $P_{i,0}$ and $P_{i,1}$). That is, the index does not include price change for new goods—

goods that entered the market between the two periods—or for goods that exited the market after the base period. Moreover, for goods that were sold in both periods, the Laspeyres fixes the relative importance of these goods at the base period levels and therefore does not reflect any changes in the composition of goods sold over time.

A Fisher Ideal index provides relative importances that are more closely aligned with the composition of goods sold over time. It is normally written as:

$$I^{F}_{0,1} = \{ [\Sigma_i P_{i,1} Q_{i,0} / \Sigma_i P_{i,0} Q_{i,0}] [\Sigma_i P_{i,1} Q_{i,1} / \Sigma_i P_{i,0} Q_{i,1}] \}^{1/2} \qquad (3)$$

It is an average (a geometric average) of the Laspeyres index—the first term—and the Paasche index—the second term.[3] The Paasche index is similar to the Laspeyres except that it uses a different market basket to measure price change—it compares the actual cost of buying the bundle in period 1 ($\Sigma_i P_{i,1} Q_{i,1}$) to what it would have cost to buy that bundle at period 0 prices ($\Sigma_i P_{i,0} Q_{i,0}$).

The Fisher index may also be written as a ratio of weighted averages:

$$I^{F}_{0,1} = \{ \Sigma_i w_{i,0} P_{i,1} / P_{i,0}] / [\Sigma_i w_{i,0} P_{i,0} / P_{i,1}] \}^{1/2} \qquad (4)$$

with the Laspeyres in the numerator and the inverse of a Paasche in the denominator. Here it is easy to see that, unlike the Laspeyres, the Fisher uses expenditure shares from both periods. So, as market shares change over time, the Fisher places a higher weight on goods that are gaining market share whereas the Laspeyres does not.

Just like the Laspeyres, however, this index ignores the entry of new goods and the exit of older goods. In a dynamic industry such as pharmaceuticals, the omission of new and exiting drugs can have important empirical implications. For drugs, the evidence is that pricing for new drugs can be very different from that of older, more established drugs, indicating that an index that includes new drugs will likely show different price growth than one that does not (Berndt 2002).

One way to better incorporate any price change for new drugs is to construct indexes over shorter spans of time and to cumulate, or chain, the resulting price indexes.

[3] A geometric average of A and B is $[AxB]^{1/2}$, and numerically gives similar answers as the usual arithmetic mean (A+B)/2.

For example, suppose we are measuring price change from 2003 to 2005 with annual price indexes. One could construct two Fisher price indexes, one for price change from 2003 to 2004 ($I^{F}_{2003,2004}$) and another for price change from 2004 to 2005 ($I^{F}_{2004,2005}$). One can then cumulate the growth in the two indexes to obtain a chained Fisher price index over the entire period:

$$I^{CF}_{2003,2005} \equiv I^{F}_{2003,2004} \times I^{F}_{2004,2005} \qquad (5)$$

The resulting index includes more spending on new drugs than the unchained version in (4). While the only new drugs included in (4) are those introduced in 2003, the chained index includes drugs introduced in 2004 in the $I^{F}_{2004,2005}$ index. Chained indexes thus provide a way of folding in new goods more quickly and so the index more closely tracks prices for the goods actually sold in the market. One can construct a chained Laspeyres in a similar manner.

One paradigm that has been used to justify the superiority of the Fisher Ideal index over others is cost of living index (COLI) theory. However, as discussed earlier, it is widely understood that the applicability of this theory in the health care setting is tenuous at best. Fortunately, there are other criteria that one can use to compare the relative merits of these formulas. In his "axiomatic approach," Diewert (1992) considers about 20 properties that one would like to see in a price index. For example, one property is a time-reversal test which requires that if the prices and quantities in the two periods being compared are interchanged the resulting price index is the reciprocal of the original price index. Diewert showed that the Fisher index formula met this and other criteria better than other available formulas.

Empirical results

An important contribution of the empirical literature was to demonstrate that the choice of formula and chaining method matters. The Fisher formula takes into account any changes in the relative importance of drugs over time, whereas the Laspeyres formula does not. Chaining indexes brings new goods into the index more rapidly.

The differences in these indexes can be positive or negative. For example, in their study of drugs sold by four companies making up about 25% of the market, Berndt,

Griliches and Rosett (1993) found that price growth in chained indexes was slower than that in fixed-based indexes. But, in their study of antidepressant drugs, Berndt, Cockburn and Griliches (1996) found the opposite—chained Laspeyres tended to show faster price growth than the unchained counterparts. Which way it goes depends on how fast prices for new goods grow relative to established goods, and how the composition of drugs in the market is changing over time.

We illustrate these points using our data. For our comparison, we used the preferred price per day for price, and did the calculations for both the NDC and GPI definitions of the product.

Using the NDC definition, the chained indexes grow at CAGRs that are about .4 percentage points slower than their unchained counterparts (about 10 percent of the unchained growth rate). This says that folding in new goods into the index more quickly yields indexes that grow slower and suggests that, in our sample, prices of new drugs grow slower than those of older drugs. With respect to choice of formula, the Laspeyres and Fisher indexes grow at very similar rates, whether the indexes are chained or not, under the NDC definition.

[TABLE 3]

A similar comparison using the GPI definition looks very different. First, the chained price indexes show faster (not slower) price growth than the unchained ones. This reflects the fact that prices for new molecules grow faster than those of older molecules that include generics: as molecules lose patent protection, the diffusion of the less expensive generics pushes down the price of the molecule. Hence, folding in new molecules faster—as the chained indexes do—yields an index that includes molecules with faster price growth and so the chained index grows faster. As before, the chained versions of the Laspeyres and Fisher are very similar.

The unchained indexes show an interesting pattern. They track prices for goods sold in 2003:1. The unchained Laspeyres—the dotted line in chart 2—grows until mid-2004 and then exhibits a declining trend through the last quarter in our data. This contour is driven entirely by the influx of generics into the market over this period. The pattern we see in the price index is mirrored in the number of generic prescriptions as a share of total: the rise in prices in the earlier period is associated with a decline in the generic share and the subsequent decline in the price index coincides with sustained increases in

the generic share. By 2005:4, over 60% of prescriptions were for generic drugs, up from 58% in 2003.

[CHART 2]

The unchained Fisher shows a similar pattern but it is less pronounced. The patterns in the Fisher and Laspeyres indexes are similar because both indexes include the same molecules (both exclude entry and exit). But, the Fisher shows slightly faster price growth because molecules with the fastest price growth also gain market share over time and thus have a bigger weight in the Fisher index than in the Laspeyres.

Summing up, chained and unchained indexes can show very different rates of price growth. In our data, the differences are particularly large for drugs defined as molecules (GPI definition) where the growth rates differ by orders of magnitude.

Implications for decompositions of spending growth

Spending on prescription drugs grew 30.9% from 2003:1 to 2005:4, or at a 9.4% CAGR. Because different price indexes yield different growth rates for measured price change, they also yield different growth rates for the implied growth of quantities, or "utilization." We illustrate this by decomposing the 9.4% CAGR into price and quantity components using the eight possible price indexes discussed above. In each case, we deflate the growth in spending using the price index to obtain the implied growth in utilization: growth in spending / growth in prices = growth in quantities. So that for the chained Laspeyres index that uses the NDC definition (in the top left portion of the table), the calculation is: (1.0938/1.0262)=1.0659.

The use of different price indexes does not materially change the qualitative conclusion that all of the indexes attribute most of the growth in spending to growth in utilization, not price: Of the 9.4% growth in spending, growth in utilization contributes 6 to 9 percentage points, depending on the index. Indexes that use the GPI definition of product attribute more than those based on the NDC definition: the unchained indexes that use the GPI definition essentially attribute all of the growth in spending to increases in utilization because measured price growth is essentially flat.

For national accounting purposes, however, the differences of an implied growth rate for quantity of 8% CAGR—the chained Fisher using the GPI definition—and 6%

CAGR—the unchained Laspeyres using the NDC definition—are large and the implied trends will be very different. After five years, the level of one index would be 7 times the level of the other if these rates were sustained.

We close by noting that this type of decomposition is similar in many respects to the decompositions that health economists use to parse out the drivers of cost growth[4] but there are differences. In those decompositions, they first decompose overall spending growth into spending on new drugs versus growth in spending on established drugs (drugs that were sold in both periods). They separate spending on new drugs because that spending growth is qualitatively different. For example, increases in spending that are due to spending on new, better, drugs have a different implication for patients than increases in spending that arise from higher prices for older drugs. They then construct a price index for the established drugs (drugs sold in 2003:1 that were also sold in 2005:4). For this purpose, unchained indexes are the relevant ones to use. They use this price index, just as we have, to estimate the growth in spending on established drugs and separate it into price and utilization components. Although the price index literature argues that the Fisher formula is superior to the Laspeyres, empirically, the choice of formula does not make much of a difference in our sample: the Fisher and Laspeyres price indexes are very similar and, so, the implied growth rates for quantity are also very similar.

[TABLE 4]

4. Measuring Quality Change

Ideally, one would want a price index that would take into account changes in the quality of goods. For example, while the average price of computers has stayed fairly constant over the last two decades, the performance (speed of processor, data storage capabilities, etc.) has greatly improved. Price indexes that take these enhancements into account should, and do, show rapid declines, reflecting rapid improvements in quality. These indexes are typically constructed by relying on market prices as a gauge of the market's valuation of quality differences across goods to measure the value of quality indirectly—as is done in price indexes—or more directly—as in a hedonic regression.

[4] See Merlis 2000 for a review.

These techniques have been applied to drug markets with mixed results. At the end of the day, many believe that the complicated features of medical care markets do not allow the interpretation of prices as a gauge of patients' valuations of drugs and hence question the ability of methods like price indexes and hedonics to adequately capture the quality of goods. As we show here, the rates of quality change implied by standard methods are quite low. Assuming that the quality of goods is improving over time, price growth measured using these techniques should perhaps be viewed as an upper bound on true price change, where the "true price change" would account for increases in quality over time. Indeed, this is the view taken in studies that aim to assess the biases in official statistics. (Lebow and Rudd 2003).

4.1 Quality in Matched-Model Indexes

We begin by asking "what do standard price indexes assume about quality?" The average quality of goods increases both when existing goods get better over time and when new, better, goods enter the market. The price indexes described above can control for the first issue of quality change in existing goods if the market and data allow one to track identical goods over time. Markets where goods are "custom"—housing, for example—present difficulties because the nature of the good makes it difficult to track identical products over time. But, this is not the case for drugs, where the available data have sufficient detail on the products so that one can track products with identical physical attributes over time. If there are unobserved attributes that change over time (e.g. perceived efficacy or experience with the drug), these indexes will count any price increases associated with these changes as increases in price, not quality.

The indexes also involve an implicit adjustment for quality change when new goods are introduced. It can be shown that these indexes value the quality differences across goods as the difference in market prices that prevailed at the point of entry (Aizcorbe 2006). Specifically, standard price indexes implicitly compare prices of new and incumbent goods and attribute that gap in prices to the market's valuation of the quality differences in the goods.

One problem with this kind of implicit valuation is that, as mentioned earlier, it is not clear that a comparison of prices provides patients' valuations of the benefits of new drugs over established ones. Another unsettling feature of this quality valuation is that it

is applied only at the period of introduction. Because the diffusion of new drugs is slow, the market share for the new drug is relatively small in the period it enters the market and grows over time. If so, including the new drug in an index as soon as possible may imply a smaller quality estimate than bringing it in later. Griliches and Cockburn (1995) discuss this issue in the context of new generics and show that different ways of handling diffusion can generate very different price indexes. This remains an unresolved issue.

4.2 Hedonic Price Indexes

Hedonic regression techniques provide an explicit way to control for quality change when constructing price indexes.[5] A hedonic regression relates variation in prices, both across goods and over time, to differences in the goods' attributes: bigger houses sell for more, higher resolution printers are more expensive, etc. To the extent that these attributes are related to price, a hedonic regression can be used to capture these relationships and to construct price indexes that control for changes in these attributes or, changes in "quality."

The Regression

Hedonic studies for drugs have typically applied the regression on pooled data and used time dummies to form a "dummy variable" (DV) price index. As argued in Schultze and Mackie (2000) and Pakes (2003), this method constrains parameters to be fixed over time whereas the underlying parameters may well change over time. However, in cases where the available data do not allow estimation in each period, pooling the data and using the DV index is the only option. This was, indeed, the case for hedonic studies of specific drugs, where the focus on narrowly defined medications did not typically yield sufficient observations to run cross-sectional regressions.

We, thus, focus on the pooled specification and the DV price index. The pooled hedonic regression explains the prices of each product that is sold at time t ($P_{i,t}$, i = 1 … N) as a function of the quantities of its characteristics ($C_{k,i,t}$, k = 1, …K) and time dummy variables ($D_{i,t}$, t = 1, …T). The regression is usually specified in semi-logarithmic form:

$$ln\ P_{i,t} = \Sigma_k \beta_k C_{k,i,t} + \Sigma_t \delta_t D_{i,t} + \varepsilon_{i,t} \qquad (6)$$

[5] See Berndt (1996) and Triplett (2006) for a full discussion of hedonic techniques

where $D_{i,t}$ = 1 if a price for product m is observed at time t, and
= 0 otherwise,

and β_k, δ_t, and $\varepsilon_{i,t}$ are econometric estimates. Each product has K characteristics that can influence its value, and, in general, the quantity of each characteristic in a product can change over time. The characteristics typically are numeric values (such as number of milligrams of active ingredient), but they can also be dummy variables that designate the presence or absence of an attribute of the good in a particular product (such as whether the drug is the extended release version).

There are a number of econometric issues in implementing hedonic regressions, including heteroskedasticity, unobserved characteristics, choice of functional form and imprecise estimates owing to collinearity (Berndt 1996). The omitted variable issue was revisited by Bajari and Benkard (2005) and Pakes and Erickson (2009). Bajari and Benkard argued that the existence of these unobservable characteristics pose problems for hedonic techniques that are made evident in the low explanatory power one typically obtains in these regressions. Their work and that in Pakes and Erickson (2009) develop new methods that account for these unobserved characteristics and shows that accounting for them not only improves the explanatory power of the regression but also the inferences that one draws from them.

On the interpretation of hedonic coefficients, Pakes (2003) argues that the hedonic regression should be interpreted as a reduced form, where the coefficients can reflect changes in both demand- and supply-side factors. For drugs, demand-side factors include factors that increase the prevalence of some conditions and, hence, the demand for medications to treat them or new knowledge about the efficacy of drugs; supply factors can include the rising cost of research and development, or variation in marketing expenditures. Under Pakes' interpretation, there is no reason to expect coefficients associated with "good" outcomes to have positive signs.

The $\Sigma_k \beta_k C_{k,i,t}$ terms control for differences in products' qualities, the regression delegates all other influences on prices to the time dummies and the (assumed normally distributed) residuals. The time dummy coefficients, δ_t, capture the average value of the other influences for each time period, and are estimates of the aggregate constant-quality price *level* (rather than price relative) for the good at time t.

Empirical results

To date, there are only five studies that have used hedonic techniques to construct price indexes for drugs: Suslow (ulcers), Berndt, Cockburn and Griliches (depression), Cockburn and Annis (arthritis), and Lucarelli and Nicholson (colorectal cancer). All of these studies show that price indexes that control for differences in attributes across drugs and over time show substantially slower price growth than average prices.

The kind of drug attributes that they used included features such as the efficacy of the drug (like healing or survival rates), ease of administration (number of daily doses needed for treatment), as well as the unwanted presence of side effects and interactions with other medications. An advantage of this regression approach over matched-model price indexes is that it can accommodate attributes that change over time, something that matched-model indexes cannot. For example, Cockburn and Anis (1998) include variables to reflect new information on old drugs from clinical trials—that is, what is known about drugs changes over time and that can be incorporated in the hedonic regression. Similarly, Berndt et al. (2001) include an indicator for experience with the drug (cumulative sales), another variable that changes over time.

Empirically, hedonic techniques applied to drugs have failed to find an overwhelming connection between the attributes and price. This result has also been reported in demand studies where the coefficient on price tends to be insignificant (Cockburn and Anis 1998 and Lucarelli and Nicholson 2009). Some have noted that this might be a reflection that patients and doctors are not very sensitive to price. Surveys cited in Suslow (1992) suggest that patients ranked affordability fourth in importance in the list of factors they look for in anti-ulcer medications, behind "Be safe," "Make you feel better quickly," and "Be convenient to take." Similarly, though doctors do include price in the list of factors they consider when prescribing drugs, it is not the most important thing. This is consistent with the observation that prices can be fairly non-responsive to relatively large changes in markets, including drugs coming off patent and the subsequent entry of generics.

Dummy Variable Index

The DV measure of price *change* (in logs) is the difference between the estimated time dummy coefficients for time 0 and time 1. When there is no entry and exit—i.e., all N goods were sold in both periods—the (logged) price index may be written:

$$\ln I^{DV} = \underline{\delta_1} - \underline{\delta_0}$$

$$= \Sigma_i (ln\ P_{i,1} - \Sigma_k \underline{\beta_k} C_{k,i,1}) / N_1$$

$$- \Sigma_i (ln\ P_{i,0} - \Sigma_k \underline{\beta_k} C_{k,i,0}) / N_0 \qquad (7)$$

where the N's denote the total number of products and we ignore the (mean zero) residuals. Note, then, that the (logged) DV measure for price change from time 0 to time 1 is the difference of two (logged) geometric means: the mean of quality-adjusted prices for products that exist at time t, the first term, and that of products that exist at time 0, the second term.

How does the DV index deal with new goods? This issue was studied by Silver and Heravi (2002) and Aizcorbe, Corrado and Doms (2003). Consider the introduction of a new good (call it "**n**") at time 1. The DV price measure provides an explicit imputation for these missing prices. To see this, augment (7) to include the new good. After some tedious algebra, the DV estimate for price change from t=0 to t=1 can be expressed as:

$$\ln I^{DV}_{1,0} =$$

$$\underline{\delta_1} - \underline{\delta_0} = (N_0 / N_1) [\Sigma_i (ln\ P_{i,1} - ln\ P_{i,0}) / N_1]$$

$$+ (1 / N_1) [(ln\ P_{\mathbf{n},1} - \Sigma_k \underline{\beta_k} C_{\mathbf{n},k,1}) - \Sigma_i (ln\ \underline{P}_{i,0} - \Sigma_k \underline{\beta_k} C_{k,i,0}) / N_0] \quad (8)$$

Equation (8) shows that the DV measure may be written as a weighted average of a price measure for continuing goods (the first term) and one for the new good (the second term) where the weights implicitly used by the hedonic regression are the share of observations of each type. For continuing goods, the DV measure uses a geomean price index. For the new good, the hedonic regression imputes a price relative as the difference between the quality-adjusted price for the new product at time 1 ($ln P_{\mathbf{n},1} - \Sigma_k \underline{\beta_k} C_{\mathbf{n},k,1}$) and

the average quality-adjusted price for all observed products in the prior period (Σ_i ($ln\ P_{i,0} - \Sigma_k \beta_k C_{k,i,0}$) / N_0).

This is similar to the implicit valuation in price indexes discussed above in that both use the period of introduction to account for quality; the DV price index thus shares the problem that standard price indexes have in trying to measure quality change in the presence of diffusion. It differs, however, in that the hedonic explicitly estimates quality differences based on the hedonic coefficients. Empirically, the DV measure typically gives slower price growth than price indexes (See, for example, table 4.4 in Berndt 1996). Some think that this is because the hedonic is better at capturing changes in quality than standard price indexes (Triplett, 2006). Others have noted that these differences in price growth might arise from the fact that the DV price indexes are unweighted (like geomeans) whereas price indexes typically use expenditure weights (Aizcorbe and Pho 2005).

In our data, a DV index shows slower price growth and faster quality growth than either the chained Fisher or Laspeyres indexes. One way to calculate the quality change implied by price indexes is using the identity: dln(average price) = ln(price index) + dln(quality change). There are choices to make about what to use as the average price and those choices could yield different measures for implied quality. For a geomean price index, for example, if one measures the change in the average price as the change in geometric means of the logged price levels for prices of all goods sold in each period, then, the implied change in quality has a clean interpretation as the difference between the (logged) price of the new goods and the average (logged) price of all goods sold that period (Aizcorbe 2006). For other index formulas, the implied quality term does not have a tidy interpretation. Nonetheless, as a first cut, we do the calculations using differences in a geomean of the (logged) price levels to get a rough gauge of how much quality growth is implied by the different indexes.

To obtain a DV price index, we estimate a hedonic regression that uses fixed effects to control for quality differences across drugs and relegates all other influences to the time dummies—we cannot do more than this because we do not observe the typical attributes used in hedonic regressions, like efficacy. However, the drug-specific fixed effects will control for any of those attributes that are fixed over the life of the drug. The

raw data are logged prices at the NDC level and the fixed effects for drugs' characteristics are at the GPI level.

The first three columns of table 5 show the usual result that the DV price index attributes more of the spending growth to quality than to price. The DV price index is essentially flat, actually falling at about a 1/4 percent CAGR and the implied rate of quality growth is 3%. The quality growth implied by the Laspeyres and Fisher indexes is about one-half the growth of quality implied by the DV index: about 1-1/2 % CAGR.

These differences are not necessarily related to the fact that the DV index is based on a hedonic regression while the others are based on standard price index methods. The last column of the table shows growth rates for a price index generated using an unweighted geometric mean formula. The growth in that index is very similar to that of the DV index, perhaps because their functional forms are so similar—both are unweighted indexes. However, the similarity only holds when the geomean index is chained, thereby including new goods quickly.

[TABLE 5]

A final point about these estimates of quality growth is that they seem small when compared with estimates for other goods. For example, standard price indexes for Intel's microprocessors implied quality growth of over 20 percent *per quarter* over the 1990s (Aizcorbe 2006). Similarly, Bils and Klenow (2001) estimate that average quality of over 60 categories of durable goods grew 3-3/4 percent per year over the 1980-96 period.

Compared with these rates of quality growth, the estimates for quality growth for drugs seem small and suggest that the methods discussed above do not adequately measure the value of new pharmaceutical innovations. This probably reflects, in part, the inability of prices to provide a good gauge of patients' valuations. To the extent that the average quality of drugs improves over time, price indexes generated using standard methods are perhaps best viewed as upper bounds to an unobserved price index that takes these quality improvements into account.

Measuring quality directly

Health economists view the output of medical services as the incremental improvements to health status that result from treatment. Cost effectiveness studies try to measure this directly by comparing the incremental cost of transitioning from new to

current treatments against the incremental benefits in terms of health outcomes or effectiveness with a cost-effectiveness ratio:[6]

$$\text{CE ratio} = [\ \text{cost}_{new} - \text{cost}_{current}\ ] \ / \ [\ \text{effect}_{new} - \text{effect}_{current}\ ] \quad (7)$$

A seminal paper by Cutler, McClellan, Newhouse and Remler (1998) provided a cost of living interpretation to this notion of cost effectiveness by tying it back to utility analysis. For heart attacks, several different types of treatments are given at the same time (e.g. surgery and drugs) so their model compares the entire cost of treating the condition with the attendant outcomes.

A close cousin of cost-effectiveness analysis is called "cost utility analysis," where the cost-effectiveness ratios are expressed in terms of cost per quality-adjusted life year (QALY), an outcome measure that incorporates the value people place on different outcomes (Drummond et al. 1997).

Similar cost-effectiveness calculations have been done to assess the quality of new drug treatments[7]. These calculations have been done using different data sources. For example, for colorectal cancer drugs, Lucarelli and Nicholson (2009) use industry data to estimate the incremental cost of new chemotherapy regimens (the numerator) and data from clinical trials to estimate the increase in life expectancy from the treatment, while Howard et al. (forthcoming) used retrospective survey data at the patient level to calculate the cost of chemotherapy and to estimate survival curves for calculating the incremental benefits in terms of increased longevity.

This would seem to be a promising method particularly when drugs are the only treatment (e.g. chemotherapy for certain cancers). When drugs and other treatments are substitutes, however, new drugs can involve cost offsets, such as when a new drug makes the utilization of other treatments are no longer necessary, that should figure into this calculation. Similarly, when drugs and other treatments are complements, it will be difficult to parse out the marginal improvements to health from drugs as opposed to other treatments.

[6] see Garber (2000) for a review of this literature

[7] see Neumann et al. (2000) and Crown, Ling, and Berndt (2002) for a summary of the various methods that have been used to study the effectiveness of drugs

Although these direct methods are likely to have any number of methodological issues, they do provide a way to assess the value of drug improvements in a more transparent way than index methods.

6. Price Indexes in Cross-Country Comparisons

Price indexes have also been used to compare drug prices in different countries. There, the question is "Are drugs more expensive in the U.S. than in other countries?" rather than "Are drugs more expensive today than they were yesterday?" For pharmaceuticals, there are both informal comparisons based on individual drugs (U.S. General Accounting Office 1994) as well as more formal calculations that apply price indexes or regression techniques to more comprehensive data (Danzon and Chao 2000, for example). Because the questions are very similar, many of the issues that arise in the context of the temporal price indexes discussed above also arise in the cross-country context.

Prices and utilization patterns for drugs vary greatly across countries so that cross-country comparisons can give very different results depending on which drugs are included and how much weight each drug is given. Perhaps the most vexing problem is that drugs sold in one country are often not sold in others so that the comparisons are necessarily incomplete. For example, using a comprehensive dataset for seven countries, Danzon and Chao (2000) found that less than one-third of the molecules sold in seven countries are present in all seven markets. Moreover, when making comparisons across pairs of countries, they found that over 40% of total retail pharmacy sales in their dataset could not be included. This is the analog to the "new goods" problem in the temporal context and makes it very difficult to boil down differences in drug prices across countries into one summary statistic.

For drugs common to the countries, comparisons based on price indexes are sensitive to choice of index formula. The larger problem around the choice of formula is that, unlike in the temporal context, different formulas answer different questions: a Laspeyres formula that uses the US as the base country tells you how much the US market basket of drugs would cost if one had to pay the other country's prices. The Paasche index tells you how much the other country would have to pay to buy their market basket at US prices. The Fisher index—that gives an average of these two—has

not been viewed as particularly informative in cross-country comparisons of drug prices. Moreover, the Laspeyres and Paasche indexes do not have the usual bounding interpretation because, as in the temporal context, the COLI theory does not apply here.

There is a fairly large literature devoted to indexes that may be used to do cross-country comparisons or, more broadly, spatial comparisons. Indexes have been used compare prices across regions of the US at a point in time (Kokoski 1991 and Aten 2008), across countries at a point in time (Kravis, I.B., A. Heston and R. Summers 1982), and across both space and time (Hill 2003). The studies use both index number approaches (Diewert 1999) and regression-based approaches (Summers 1973).

7. Summary

Existing work in this area has gone a long way toward improving our understanding of price indexes and the kinds of questions that they can and cannot address. The lessons from this literature are numerous.

First, much has been learned about the relative merits of different aggregation schemes. Indexes that more-closely track the composition of products sold in the market are better than those that do not: chained indexes are better than unchained ones and the Fisher formula is better than the Laspeyres.

Second, different ways of dealing with the entry of generic drugs can yield very different price indexes. Price indexes based on a GPI definition grow much slower than those based on the NDC definition. The GPI indexes are viewed as the better way to define the product, with the caveat that perhaps patients do not always view branded and generic versions of drugs as perfect substitutes. More work is needed to pin this down further.

Third, perhaps the most daunting problem is that existing methods do not provide an adequate way to deal with improvements provided by new drugs. At the end of the day, many believe that the complicated features of medical care markets do not allow the interpretation of prices as a gauge of patients' valuations of drugs and, hence, question the ability of methods like price indexes and hedonics to adequately capture the quality of goods. What we're left with is an interpretation of these indexes as upper bounds to true price change: Assuming that the quality of drugs is improving over time, price indexes

that do not adequately account for these better outcomes are overstating the price per quality unit to patients. More work is needed to find alternative methods that allow one to infer the benefit to patients of new drugs. Merging what is known about price indexes with cost effectiveness methods would seem to be a promising avenue of research.

Finally, the problem of accounting for quality change is part of a broader issue. Namely, there are difficulties in tying these indexes to consumer welfare or , more specifically, to the improved outcomes from medications. Without a COLI or similar paradigm, it is difficult to interpret changes in measured price growth as good or bad. Structural demand estimation that allows one to estimate utilities—like Lucarelli and Nicholson (2009)—is a promising line of research that could fill this gap.

References

Abraham, Katharine, and Christopher Mackie, eds. *Beyond the Market: Designing Nonmarket Accounts for the United States*: The National Academies Press, 2005.

Aizcorbe, Ana. "Why Did Semiconductor Price Indexes Fall So Fast in the 1990s? A Decomposition." *Economic Inquiry* 44 (2006): 485-96.

Aizcorbe, Ana, Carol Corrado, and Mark Doms. "When Do Matched-Model and Hedonic Techniques Yield Similar Measures?" In *Working Papers in Applied Economic Theory*: Federal Reserve Bank of San Francisco, 2003.

———. "When Do Matched-Model and Hedonic Techniques Yield Similar Results?"no. WP 2003-14 (2003), http://www.frbsf.org/publications/economics/papers/2003/wp03-14bk.pdf.

Aizcorbe, Ana, and Nicole Nestoriak. "The Importance of Pricing the Bundle of Treatments." In *BEA Working Paper Series*: Bureau of Economic Analysis, 2008.

———. "Pce Deflators for Medical Care: A Progress Report." In *Academic Advisory Committee of the Bureau of Economic Analysis*. Washington, DC: Bureau of Economic Analysis, 2005.

Aizcorbe, Ana, and Yvonne Pho. "Differences in Hedonic and Matched-Model Price Indexes: Do the Weights Matter?" In *BEA Working Paper Series*: Bureau of Economic Analysis, 2005.

Aten, Bettina H. "Estimates of State and Metropolitan Price Levels for Consumption Goods and Services in the United States, 2005." In *BEA Paper*, 2008.

Bajari, Patrick, and C. Lanier Benkiard. "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach." *Journal of Political Economy* 113 (2005): 1239-76.

Berndt, Ernst R. "The Measurement of Quality Change." In *The Practice of Econometrics*. Reading: Addison-Wesley Publishing Company, 1996.

———. "Pharmaceuticals in U.S. Health Care: Determinants of Quantity and Price." *Journal of Economic Perspectives* 16 (2002): 45-66.

Berndt, Ernst R., Alicia M. Busch, and Richard G. Frank. "Medicare Part D, Prescription Drug Prices, and Treatment Quality." *American Economics Association* 4 (2007): 1-41.

Berndt, Ernst R., Iain M. Cockburn, Douglas L. Cocks, A. Epstein, and Zvi Griliches. "Is Price Inflation Different for the Elderly? An Empirical Analysis of Prescription Drugs." *Frontiers of Health Policy* 1 (1998): 33-75.

———. "Prescription Drug Prices for the Elderly." *Monthly Labor Review* 121 (1998): 23-34.

Berndt, Ernst R., Iain M. Cockburn, and Zvi Griliches. "Pharmaceutical Innovations and Market Dynamics: Tracking Effects on Price Indexes for Antidepressant Drugs." *Brookings Papers on Economic Activity: Microeconomics* 2 (1996): 133-88.

Berndt, Ernst R., and P.E. Greenberg. "An Updated and Extended Study of the Price Growth of Prescription Pharmaceutical Preparations." In *Competitive Strategies in the Pharmaceutical Industry*, edited by R.B. Helms, 35-48. Washington, DC: American Enterprise Institute, 1995.

Berndt, Ernst R., Zvi Griliches, and J.G. Rosett. "Auditing the Producer Price Index: Micro Evidence from Prescription Pharmaceutical Preparations." *Journal of Business and Economic Statistics* 11 (1993): 251-64.

Berndt, Ernst R., Margaret K. Kyle, and Davina C. Ling. "The Long Shadow of Patent Expiration: Generic Entry and Rx-to-Otc Switches." In *Scanner Data and Price Indexes*, edited by Robert C. Feenstra and Matthew D. Shapiro, 229-74. Chicago: The University of Chicago Press, 2003.

Bils, Mark, and Peter Klenow. "Quantifying Quality Growth." *American Economic Review* 91 (2001): 1006-30.

Boskin, Michael J., E. Dulberger, R. Gordon, Zvi Griliches, and Dale Jorgenson. "Toward a More Accurate Measure of the Cost of Living." Advisory Commission To Study The Consumer Price Index 1996.

Chernew, Michael, Dean Smith, Duane Kirking, and A. Mark Fendrick. "Decomposing Pharmaceutical Cost Growth in Different Types of Health Plans." *American Journal of Managed Care* 7 (2001): 667-741.

Cockburn, Iain M., and A. Anis. "Hedonic Analysis of Arthritis Drugs." In *Medical Care Output and Productivity*, edited by David M. Cutler and Ernst R. Berndt, 439-61. Chicago: The University of Chicago Press, 1998.

Court, Andrew. "Hedonic Price Indexes with Automotive Examples." In *The Dynamics of Automobile Demand*, 99-117. New York: General Motors Corporation, 1939.

Crown, William H., Davina C. Ling, and Ernst R. Berndt. "Measuring the Costs and Benefits of Pharmaceutical Expenditures." *Expert Review Pharmaeconomics Outcome Research* 2 (2002): 467-75.

Cutler, David M., and Mark McClellan. "Productivity Change in Health Care." *American Economic Review* 91 (2001): 281-86.

Cutler, David M., Mark McClellan, Joseph P. Newhouse, and Dahlia Remler. "Are Medical Prices Declining? Evidence from Heart Attack Treatments." *Quarterly Journal of Economics* 113 (1998): 991-1024.

Danzon, Patricia M., and Li-Wei Chao. "Cross-National Price Differences for Pharmaceuticals: How Large and Why?" *Journal of Health Economics* 19 (2000): 159-95.

Danzon, Patricia M., and Jeong D. Kim. "International Price Comparisons for Pharmaceuticals: Measurement and Policy Issues." *Pharmaeconomics* 14 (1998): 115-28.

Diewert, Erwin W. "Axiomatic and Economic Approaches to Elementary Price Indexes." National Bureau of Economic Research, 1995.

———. "Axiomatic and Economic Approaches to International Comparisons." In *International and Interarea Comparisons of Income, Output and Prices*, edited by A. Heston and R.E. Lipsey, 13-87. Chicago: University of Chicago Press, 1999.

———. "Exact and Superlative Index Numbers." *Journal of Econometrics* 4 (1976): 115-45.

———. "Fisher Ideal Output, Input and Productivity Indexed Revisited." *Journal of Productivity Analysis* 3 (1992): 211-48.

Drummond, M.F. et. al. *Methods for Economic Evaluation of Health Care Programmes*. New York: Oxford University Press, 1997.

Dubois, Robert W., Anita J. Chawla, Cheryl A. Neslusan, Mark W. Smith, and Sally Wade. "Explaining Drug Spending Trends: Does Perception Match Reality?" *Health Affairs* 19 (2000): 231-39.

Erickson, Timothy, and Ariel Pakes. "An Experimental Component Index for the Cpi: From Annual Computer Data to Monthly Data on Other Goods." Harvard University, 2008.
Feenstra, Robert C. "Generics and New Goods in Pharmaceutical Price Indexes: Comment." *American Economic Review* 87 (1997): 760-67.
———. "New Product Varieties and the Measurement of International Prices." *American Economic Review* 84 (1994): 157-77.
Fisher, Franklin M., and Zvi Griliches. "Aggregate Price Indices, New Goods, and Generics." *Quaterly Journal of Economics* 110 (1995): 229-44.
Fisher, Irving. *The Making of Index Numbers*. Boston: Houghton Mifflin, Co., 1922.
Garber, Alan. "Advances in Cost-Effectiveness Analysis of Health Interventions." In *Handbook of Health Economics*, edited by A.J. Culyer and Joseph P. Newhouse, 181-221. North Holland: Elsevier, 2000.
Gordon, Robert J., and Zvi Griliches. "Quality Changes and New Products." *American Economic Review* 87 (1997): 84-88.
Griliches, Zvi. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change." In *The Price Statistics of the Federal Government*, 173-96. New York: National Bureau of Economic Research, 1961.
Griliches, Zvi, and Iain M. Cockburn. "Generics and New Goods in Pharmaceutical Price Indexes." *American Economic Review* 84 (1994): 1213-32.
Hill, Robert J. "Constructing Price Indexes across Space and Time: The Case of the European Union." *American Economic Review* 94 (2003): 1379-410.
Hill, T.P. "Consumer Price Index Manual: Theory and Practice ", edited by T.P. Hill: International Labor Organization, 2004.
Howard, David H., John Kauh, and Joseph Lipscomb. "The Value of New Chemotherapeutics for Metastatic Colorectal Cancer." *Archives of Internal Medicine* (Forthcoming).
Kelly, Greg G. "Improving the Ppi Samples for Prescription Pharmaceuticals." *Monthly Labor Review* 120 (1997): 10-17.
Kokoski, Mary F. . "New Research on Interarea Consumer Price Differences." *Monthly Labor Review* 114 (1991): 31-34.
Konus, A.A. "The Problem of the True Index of the Cost of Living." *Econometrica* 7 (1939): 10-29.
Kravis, I.B., and Robert Summers. *World Product and Income: International Comparisons of Real Gross Product*. Baltimore: Johns Hopkins University Press, 1982.
Lebow, David, and Jeremy Rudd. "Measurement Error in the Consumer Price Index: Where Do We Stand?" *Journal of Economic Literature* 41 (2003): 159-201.
Lucarelli, Claudio, and Sean Nicholson. "A Quality-Adjusted Price Index for Colorectal Cancer Drugs." National Bureau of Economic Research, 2009.
Merlis, Mark. "Explaining the Growth in Prescription Drug Spending: A Review of Recent Studies " In *Pharmaceutical Pricing Practices, Utilization and Costs* Leavey Conference Center, Georgetown University 2000.
Neumann, Peter J., Eileen A. Sandberg, Chaim M. Bell, Patricia W. Stone, and Richard H. Chapman. "Are Pharmaceuticals Cost-Effective? A Review of the Evidence." *Health Affairs* 19 (2000): 92-109.
U.S. General Accounting Office. "Prescription Drugs: Companies Typically Charge More in the United States Than in the United Kingdom." Washington, DC, 1994.

Pakes, Ariel. "A Reconsideration of Hedonic Price Indexes with an Application to Pc's." *American Economic Review* 93 (2003): 1578-96.
Pollak, Robert A. *The Theory of the Cost-of-Living Index*. New York: Oxford University Press, 1989.
Schultze, Charles, and Christopher Mackie. *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes*, Panel on Conceptual, Measurement, and Other Statistical Issues in Developing Cost of Living Indexes. Washington, DC: National Academies Press, 2002.
Schweitzer, Stuart O. *Pharmaceutical Economics and Policy*. 2nd ed. New York: Oxford University Press, 2007.
Summers, Robert. "International Price Comparisons Based Upon Incomplete Data." *The Review of of Income and Wealth* 19 (1973): 1-16.
Suslow, Valerie Y. "Are There Better Ways to Spell Relief? A Hedonic Pricing Analysis of Ulcer Drugs." University of Michigan, 1992.
Triplett, Jack. *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products*. Paris: Organization for Economic Co-operation and Development, 2006.
U.S. General Accounting Office. "Prescription Drugs: Companies Typically Charge More in the United States Than in the United Kingdom." Washington, DC, 1994.

Table 1. Alternate Definitions of Product
(compound annual growth rates)

Product Definition	Laspeyres	Fisher
NDC	**2.62%**	**2.60%**
GPI	**1.18%**	**1.20%**

Note: All indexes are chained from 2003:1 - 2005:4

Table 2. Growth in Chained Price Indexes, 2003:1-2005:4: Definition of Quantity (compound annual growth rates)

	NDC		GPI	
	Laspeyres	**Fisher**	**Laspeyres**	**Fisher**
Days of Treatment	**2.62%**	**2.60%**	**1.18%**	**1.20%**
Package	**3.78%**	**3.49%**	**2.40%**	**2.14%**
Prescription	**4.51%**	**4.42%**	**3.16%**	**3.11%**

Note: All indexes are chained from 2003:1 to 2005:4.

Table 3. Effect of Chaining Price Indexes (compound annual growth rates)

	NDC		GPI	
	Laspeyres	**Fisher**	**Laspeyres**	**Fisher**
Chained	**2.62%**	**2.60%**	**1.18%**	**1.20%**
Unchained	**3.01%**	**2.96%**	**0.06%**	**0.53%**

Note: Prices are defined using days of treatment.

Table 4. Alternate Decompositions of Growth in Spending
(compound annual growth rates)

	NDC		GPI	
	Laspeyres	**Fisher**	**Laspeyres**	**Fisher**
Growth in Spending	**9.38%**	**9.38%**	**9.38%**	**9.38%**
	Chained			
Price	**2.62%**	**2.60%**	**1.18%**	**1.20%**
Quantity	**6.59%**	**6.61%**	**8.11%**	**8.08%**
	Unchained			
Price	**3.01%**	**2.96%**	**0.06%**	**0.53%**
Quantity	**6.18%**	**6.24%**	**9.32%**	**8.80%**

Note: All prices are defined as price per day of treatment

Table 5. Alternate Decompositions of Growth in Average Prices (compound annual growth rates)

	Hedonic	Price Indexes		
	DV	Geomean	Laspeyres	Fisher
Average Price	**2.81%**	**2.81%**	**2.81%**	**2.81%**
Price Index	**-0.25%**	**-0.18%**	**1.18%**	**1.20%**
Implied Quality	**3.06%**	**2.99%**	**1.63%**	**1.61%**

Note: All prices are defined as price per day of treatment; all indexes are chained from 2003:1 to 2005:4

Chart 1. Effect of Brand-Generic Substitution on Price of Molecule

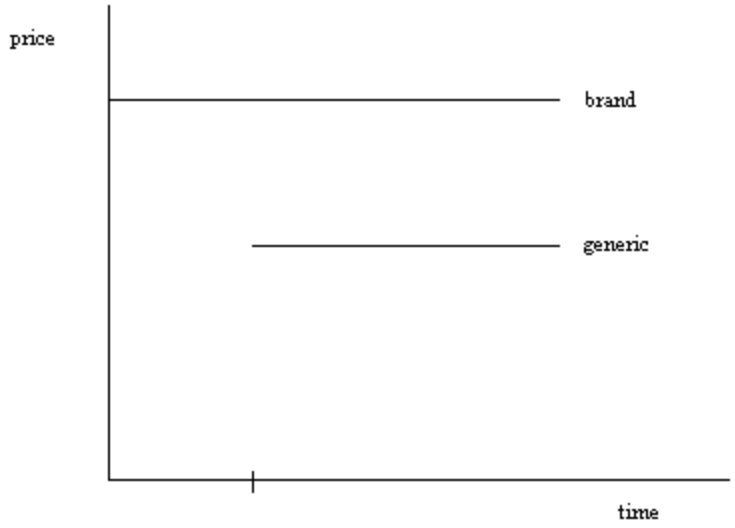

Chart 2. Effect of Generic Diffusion on Unchained Laspeyres Price Index