FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall
Philadelphia, Pennsylvania 19106-1574
(215) 574-6428, www.phil.frb.org

Working Papers

Research Department

WORKING PAPER NO. 97-4

THE MEASURMENT OF RETAIL OUTPUT AND THE RETAIL REVOLUTION

Leonard I. Nakamura
Federal Reserve Bank of Philadelphia

May 1997

WORKING PAPER NO. 97-4

The Measurement of Retail Output and the Retail Revolution

Leonard I. Nakamura
Federal Reserve Bank of Philadelphia

May 1997

Address editorial correspondence to:
Leonard I. Nakamura
Research Department
Federal Reserve Bank of Philadelphia
10 Independence Mall
Philadelphia, PA 19106
215-574-3804 (office)
215-574-4364 (fax)
nakamura@frbphil.org (email)

 I would like to thank Erwin Diewert and Pierre Fortin for comments and encouragement, Marshall Reinsdorf, Rick Lang, and Loretta Mester for detailed comments, and participants at seminars held by the Federal Reserve Bank of Philadelphia and Hoenig and Co. for comments. I am extremely grateful to the organizers and participants at the preconference and conference on Service Sector Productivity and the Productivity Paradox for their comments and support. Denise Duffy provided excellent research assistance on this project.

ABSTRACT

The Measurement of Retail Output and the Retail Revolution

The computerization of retailing has made price dispersion a norm in the United States, so that any given list price or transactions price is an increasingly imperfect measure of a product's resource cost. As a consequence, measuring the real output of retailers has become increasingly difficult. Food retailing is used as a case study to examine data problems in retail productivity measurement. Crude direct measures of grocery store output suggest that the CPI for food-at-home may have been overstated by 1.4 percentage points annually from 1978 to 1996.

The Measurement of Retail Output and The Retail Revolution

I. INTRODUCTION

This paper is an exploration of the law of one price and the consequences of its violation for the measurement of output and price. Jevons's law of one price is simply that "in the same open market, at any one moment, there cannot be two prices for the same kind of article." (Jevons,1879, in Wilson, 1964) This unique price in turn will reflect resource cost in a competitive market. Thus the law of one price implies that prices are a useful measure of resource cost.

The paper focuses on the retail revolution as the source of the violation of the law of one price and the difficulties this creates for measuring retail productivity. The retail revolution, which began in earnest in the late 1970s, is the rapid automation of retail transactions processing made possible by computerization (the lineaments of this revolution were first descried in the work of Bluestone et al, 1981). Computerization of retail transactions -- a process drastically accelerated by the widespread adoption of scanners by retailers over the course of the 1980s -- has facilitated the ability of retailers to i) cheaply and efficiently vary prices, ii) offer an increasing variety of products, and iii) analyze in detail the price elasticities of demand for products. As a consequence, computerization has accelerated a process of product differentiation in which characteristics of products that are not particularly relevant to the production cost of the product are used to allocate costs to appropriately elastic consumers. For example, whether or not the two halves of a round trip by air are separated by a Saturday night is scarcely relevant to the production cost of the flights, but this restriction separates price inelastic business travelers from price elastic vacation travelers. Another example is Black Friday, the day after Thanksgiving in

America when winter goods are widely discounted. If anything, direct sales costs are higher on this traditional vacation weekend, but economies of scale for shoppers on that day make it a focal point to separate shoppers with low search costs from shoppers with high search costs.

In both the cited examples, from the perspective of productivity measurement, the correct price is the weighted average price rather than either the high price or the low price, as will be shown in the next section. Correctly measuring productivity requires knowing the quantities sold at the two prices, particularly when pricing practices are changing.

Price discrimination has become very widespread in retailing. Other rapidly changing aspects of retailing include longer hours of operation, increases in product variety (rapid increases in store-keeping units and UPC codes), information exchange technology (scanners and electronic data interchange), inventory management (just-in-time inventory techniques and inventory management by manufacturers), retail outlets (buying clubs and category killers), and retail environments (regional malls and selling floor space). The speed of these changes in retailing, which themselves are in large part due to reduced costs of information processing, communication, and transportation, weakens the a priori case for the standard method of deflation.

When inflation is incorrectly measured, real output measures of retail stores are misleading. This can be corrected, although only crudely and in part, by looking at quantitative measures of retail service provision. Accurate measurement of inflation and productivity is likely to require ongoing research efforts as retail practices evolve.

One aspect of this problem that has been repeatedly recognized is the product life cycle. The 1960 NBER Price Statistics Review Committee (1961), chaired by Stigler, wrote:

> New products are usually introduced at relatively high prices and their prices fall as they gain acceptance, owing to economies of producing them on a larger scale and to improvement in the technique of production that come with time and experience. The price of a mature product or service is likely to be at the lowest level in its history relative to other prices. Finally, in the "old age" of a product, its relative price will often tend to rise as the scale of production contracts and economies of scale are reduced. (p.39)

The "old age" phase is one in which, although the product remains in "competition" with new products that are replacing it, its rising price is not a symptom of a general price rise or of an increase in the cost of living, but of its failure to compete successfully. Here the law of one price fails to hold between the mature product and similar competitors; this is discussed further in section II. The retail revolution has the effect of accelerating the rate of product introduction and speeding up the product life cycle, producing an acceleration in price mismeasurement in addition to that due to the multiple pricing at any given stage of the product life cycle discussed above.

II. THEORY WITH ILLUSTRATIVE EXAMPLES

The importance of the law of one price for the measurement of price and output can be seen in the following example. Suppose there are two goods, an old good and a new good. In 1980, the old good has a price of $10 and the new good has a price of $5. Ten years later, the old good's price has risen to $12 while the new good's price is unchanged. If the law of one price holds, then the old good and the new good are evidently different goods since they have different prices. But if the law of one price fails to hold, because the old good and the new good are cases of Coca Cola being sold by discounters, one efficient and growing (WalMart, for example) and the other inefficient and failing, then they are not different goods.

Assume that consumers have shifted from the old good to the new good as shown in Table 1. The question is, how should we measure inflation and output growth? If the two different products are different goods, whose relative value to the consumer is reflected by their relative prices and quantities, then the price level has risen as in line 4 of Table 2, where we use the 1980 quantity weights to weight the price movements. If the two products are the same good, then the price level has fallen because the average unit price has fallen (last line, Table 1). How do we know when goods are the same good or different goods? The answer is that, if we cannot rely on the law of one price, we need additional information about the products and how consumers value what differentiates them. But the choice makes a substantial difference in how we measure output (Table 3).

As discussed in the next section, the US Bureau of Labor Statistics (BLS) considers items sold at different outlets to be different goods when calculating the Consumer Price Index (CPI). When a store with an obsolete retail technology that sells goods at a stable high price is replaced by one with a superior technology that sells them at a stable low price, then the switch implies, given the BLS methodology, a decline in output with a stable price rather than a decline in price. Thus retail efficiency is mistaken for inefficiency. This point was made by Denison (1962) and Reinsdorf's (1993) seminal article picked up this theme with respect to grocery store prices. Dulberger (1993), in the same NBER volume as Reinsdorf, made the same point for semiconductors, where inflation mismeasurement has been spectacular. It applies to airfares and prescription drugs, Pashigian's (1988) work on department store pricing, telecommunications, Shepard's (1991) work on gasoline stations, and to the fast food market.

One firm, one good. Under the law of one price, there is a single price, p, with resource

cost c(q). Total revenue, given quantity q is pq. Profit is pq - qc(q). If the zero profit condition holds, then p=c(q), and price is a direct measure of resource cost. If the law of one price fails, then the firm may sell the good for several different prices, p(i), i =1 . . . n, with quantities q(i) sold at each price. Then profit is $\sum$ p(i)q(i) - qc(q), where q = $\sum$ q(i). By assumption, cost does not vary with price: that is the essence of the violation of the law of one price. Under the zero profit condition, the *weighted average* price equals resource cost, $\sum$ p(i)q(i)/q =c(q).

Suppose the zero profit condition does not hold. In this case $\sum$ p(i)q(i) $\geq$ qc(q). The weighted average price is an upper bound to resource cost in the multiple price case, just as the single price is when the law of one price holds.

Suppose a good begins with one price and then has two. The methodology of the BLS in constructing the CPI has generally been to treat the second price as a new good, and the price difference as a reduction in utility, rather than a reduction in resource cost. That is, the CPI in effect treats the unrestricted or undiscounted price as the resource cost of production.

Is that the case here? No. The resource cost is better measured by the weighted average price. Unfortunately, to know the weightedc average price we need to know the quantities sold at each of the several prices. In the absence of data on the quantity sold at each price, any selection rule is likely to be biased.

Why should the number of prices associated with a particular good increase? Deregulation and computerization are two reasons. Regulation often enforces a single price regime artificially, so that with the lifting of regulation a multiplicity of prices appears. As argued above, computerization may make it easier to price discriminate. This is true of airfares, car rentals, and hotel rates, where computerized reservation systems permit discrimination across

customers, dates, and bundles of purchases.

An example: Air travel. In 1978, when our current methodology for the CPI was put in place, there was only one round-trip coach fare on most routes, as fares were regulated by the Civil Aeronautics Board. At present, by contrast, dozens of different fares are available with a variety of restrictions on every route, and the fare structure changes by the minute.

Between 1978 and 1996, the average price paid per mile by passengers grew at a 2.7 percent annual rate (Table 4). The CPI-U for airfares grew at an 8.3 percent annual rate, a difference of 5.6 percentage points. If we use the CPI to deflate airline revenues from passenger travel, we find that "real" airline passenger travel output fell from 1978 to 1996. But, in fact, passenger miles on airlines more than doubled.

How can such a substantial gap have been sustained for so long? The reason is the dispersion of fares. Full fare for unrestricted travel has risen at nearly a 9 percent annual rate, and the CPI for airfares has basically tracked the full fare. The average restricted (discount) fare has increased at only 2 percent a year. The average domestic unrestricted fare is now 3.1 times as much as the average restricted fare. But only 7 percent of passenger miles are flown at full fare (full fares account for 20 percent of passenger revenues, because the average restricted fare is only one-third as much as the average full fare).

It can be argued that this constellation of fares is highly efficient. In essence, airlines have divided their customers into two broad groups: business travelers and vacation travelers. Business travelers, who care most about saving time, want to have the maximum possible flights to a wide variety of destinations. They are willing to pay more for this privilege. Vacation travelers, on the other hand, care most about saving money and are often flexible about exactly

when they fly. The airlines accommodate both types of travelers, providing a multitude of flights for business travelers and filling the seats with vacation travelers. The typical restriction on discount flights -- a Saturday overnight stay, which vacation travelers can often easily accommodate -- is used to separate business travelers from vacation travelers.

The proliferation of fares and restrictions also has costs. Consumers have to be careful shoppers, planning their trips in advance and having to guess when to lock in a nonrefundable fare. Changing plans has become more costly. There are ways to estimate the relative costs and benefits of the proliferation of fares, but they are not simple. One approach is to directly measure the cost of the restrictions to the flier, such as how much it costs to change the schedule (if a flier could have a nonstop flight instead of a multistop flight for $50 more, the cost of that restriction must be less than $50 to the flier). Another approach is time diary measures, which track shopping times (the amount of time consumers spend searching for the lowest fares) and compare them to estimated waiting times (the amount of time consumers save by having more flights available). A third approach is to measure the elasticities of demand for different types of travelers, which gives us estimates of the travelers' tradeoff between time and price. A fourth approach is to estimate the market structure of airline competition, to see where airlines have market power and where competition reigns; this enables us to see how large the departure from the zero profit condition might be. A fifth approach is to analyze and measure the role of the various contributors to airline output: airplane manufacturers, travel agents, airlines, airports. Using multiple approaches permits a more precise estimate of the inflation rate.

Consumer surplus, one good, two prices. We use Figure 1 for the discussion. Suppose a good is sold at a single price, P^0 at time 0. Nominal sales are represented by the rectangle P^0Q^0.

Consumer marginal utility is reflected in the demand curve, through points A and B.[1] Then at time 1, it is sold at two prices, P^0 and P^1, because the sellers are able to use some device to separate buyers for the goods. If the sellers are able to perfectly price discriminate, then $Q^1 - Q^0$ purchases are made at the lower price. Consumer surplus will be ABD under perfect price discrimination. P^0ABDP^1 represents the maximum consumer surplus under imperfect price discrimination; it represents the consumer surplus if the price actually falls to P^1. Note that to the extent consumers actually pay the lower price, the arithmetic is precisely that of any other decline in price.

The BLS methodology in this case would be to view the good with price P^0 as a different good from the good with price P^1 and find no decline of price. For example, before 1978 airfares were regulated and a single coach roundtrip fare between Los Angeles and New York prevailed. But since then, deregulation and the institution of computerized reservation services have permitted airlines to charge multiple fares. The BLS methodology was to measure the inflation rate of the unrestricted roundtrip fare, considering the restricted roundtrip fares sold to vacation travelers a separate good. This practice undermeasures consumer surplus. The BLS practice of ignoring the price decline is a correct application of the Laspeyres price index only if price discrimination is perfect.

A lower bound to inflation, on the other hand, can be found if we consider the original single price as the reservation price for the new good. Using the change from the original price to the new lower price as the rate of inflation will overmeasure consumer surplus. As the airfares

---

[1]The diagram assumes unitary demand elasticity, which implies under perfect competition that changes in price for this product affect the demand for only this product. In turn, it follows that the demand curve measures consumer surplus.

example shows, these upper and lower bounds can easily drift far apart. This is why a reasonably precise measure of inflation cannot be obtained in the multiple price case without additional information about the transactions.

The clearest examples of goods that have multiple prices are discount goods. Pashigian studies the issue of why goods are sold at discount and argues that the data support either a thesis of demand uncertainty (Lazear, 1986) or price discrimination (Chamberlin, 1948). Interestingly, Pashigian's data suggest that markdowns accelerated for goods in the 1980s. The dollar value of total markdowns as a percentage of dollar sales increased from 5.2 percent in 1955, to 6.1 percent in 1965, 8.9 percent in 1975, and 16.1 percent in 1984. Compared to the undiscounted price, the average price thus fell at a 0.1 percent annual rate from 1955 to 1965, at a 0.3 percent annual rate from 1965 to 1975 and a 0.8 percent annual rate from 1975 to 1984.

Pashigian also reports that, according to data for 1986 from the Market Research Corporation of America (11,500 consumers, panel, diary data) on men's dress shirts, between 55 and 78 percent of all shirts were sold on sale, with an average percent markdown of between 34 and 48 percent for those sold on sale. Percent sold on sale peaks in January and July, with percent markdowns highest in February and August. Thus sales are greatest at the end of the fall-winter and spring-summer seasons. At a minimum, this suggests roughly a 19 percent markdown on the average purchase.

Two firms, one good. The issue discussed above concerns price differentials by the producer of a differentiated good under conditions of free entry, a version of monopolistic competition. Price differentials can also, of course, arise in the competition across similar goods produced by different firms. The quote from the Stigler Price Review Committee in section I

refers to such price differentials.

Identical or nearly identical products can sell at sharp price differences because one product has a brand name and the other is a generic or discounted product. Steady increases in cigarette prices during the 1980s, for example, led to the introduction of off-brand cigarettes. Consumers who choose a generic product over the branded product may experience a substantial gain in consumer surplus or a little one.

A striking example of the competition between a brand name and a discounted product arises in the case of pharmaceuticals after the patent has expired on a brand name drug. In this case, in the U.S., the generic competition has the authority of the Food and Drug Administration attesting to its therapeutic equivalence with the brand name drug.

Griliches and Cockburn (1994) have closely examined at the wholesale level a small group of pharmaceuticals whose patents have expired. The typical behavior is that the price of the brand-name patent drug stays relatively fixed and may even rise, while the generic competition comes in at a substantially lower price and declines over time. The generic competition captures most of the market, but the brand name retains a substantial fraction of the market. In a detailed analysis of one of the drugs, Griliches and Cockburn show that the price for Keflex (brand name for cephalexin) continued to rise after it went off patent, rising 24 percent, despite competition from generic cephalexin. Generic cephalexin started at a price that was 54 percent of Keflex's, and over the next three years fell by nearly half, to less than one-quarter of Keflex's price at that date. Generics captured five-sixths of the market in quantity terms, but Keflex retained just over 45 percent of the revenue stream.

The BLS, until January 1995, treated the generic drug as a separate product worth only a

fraction of the brand name product (the relative price of the generic at introduction) and would likely have shown a rising price index for this drug. But many purchasers are probably indifferent to the brand name and would have experienced a pure price decline when the generic became available. An analysis of this example by Fisher and Griliches (1995) argues that the most reasonable measure of the consumer surplus realized by purchasers who switched would lead to a price index that declines 48 percent, compared with the Laspeyres approach, which would show a price increase of 14 percent. Over the 45 months covered in the study, this represents an annual rate difference of 19 percentage points, for a product whose wholesale value was 0.7 percent of the prescription market, so correcting the price mismeasurement on this one drug alone could lower the annual inflation rate of the prescriptions index by 0.1 percentage point.

Broadly speaking, drugs no longer under patent account for 70 percent of the dollar value of the prescription market, so that these effects may be large (Caves et al, 1991). Broader studies of generic substitution by Caves et al (1991), and Grabowski and Vernon (1992) confirm the outlines of the Keflex study for a wider range of drugs. It is quite conceivable that redefining the value of generics would lower the price level for prescription drugs 3 percentage points annually. Scherer (1993) estimates conservatively that the treatment of generics produces an annual upward bias of 1.2 percent. A further source of bias arises from drug manufacturers switching prescription drugs to over-the-counter status, which also induces substantial consumer surplus gains (Temin, 1992). As with outlet differentials, unit values can be used to estimate an upward bound on the bias in failing to link generic products with brand-name products. Possible intermediate assumptions are discussed in Fisher and Griliches (1995). The marketing of generic

drugs in substantial numbers dates only from the passage of the Drug Price Competition and Patent Term Restoration Act in 1984, so this phenomenon is an acceleration of mismeasurement since that time. In January 1995, the BLS revised its procedures for generic drugs. Now the BLS records a price decline from the brand name product to the generic when the generic predominates. This procedure should eliminate this part of the measurement bias in prescription drugs. No steps have been taken so far for drugs whose status changes from prescription to nonprescription.

Are there important biases in the other direction? Shapiro and Wilcox (1996) argue that there are important biases in the other direction. They discuss the problem of the failure of the law of one price across firms as "the new-outlets effect." They discern five cases in which an incumbent is faced by an entrant.The entrant can enter with a low price product with high quality (LH), with a low price product with low quality (LL), or a high price product with differentiated quality (HD). The incumbent provides a high price product with high quality (HH). In response to entry, the incumbent can stand pat (HH), lower price and maintain high quality (LH), lower price and lower quality (LL).

The five cases they consider among the possibilities are listed in the matrix in Figure 2. They ask: when should a lower price on the part of a new outlet be considered a decline in price? Their response is that much of the time, the lower price is properly captured in the CPI. They admit that in case 1, when the entrant has a low price and the incumbent has a high price, the lower price is a decline in price, and would not be properly captured by the CPI. This is clearly what was happening in the generic drug case.

What about cases 2 and 5? In these cases, the entrant lowers price, but the incumbent can

duplicate the technology of the entrant. In case 2, when both the entrant and the incumbent lower price and quality remains high, they argue that the CPI properly captures the decline in price, so there is no bias. But they claim in case 5, when both entrant and incumbent lower price by lowering quality, the CPI is biased downwards, registering a downward shift in price when the fall in quality means there should be no such decline. Thus cases 2 and 5 imply that the example of case 1 exaggerates the bias in the CPI.

But this argument rests on the assumption that the incumbent can easily duplicate the technology of the entrant, so that the law of one price holds. And if the incumbent can duplicate the entrant's technology there is no reason for successful entry. The thought experiment in cases 2 and 5 is simply an assertion that the law of one price holds. This is an empirical issue.

In cases 3 and 4, in contrast, there is a rationale for successful entry, as the entrant increases variety. In case 3 the entrant provides a lower cost product with lower quality, while in case 4 the entrant provides a high cost product with differentiated quality. In both cases the entrant must be providing consumer surplus if it is displacing the incumbent, and in neither case does the CPI register a decline in price, so the CPI is upwardly biased. Is the average displacement price an accurate measure of the missing consumer surplus? If we believe that on average unmeasured quality is deteriorating, then the average price overstates consumer surplus. But is this the case? With average income per person or per household rising, one would expect that on average unmeasured quality was improving. Thus on average, replacements by the entrant are likely to be of *rising* quality, so that consumer surplus is understated.

An advance in retail technology may very rapidly shift sales from one set of store-item combinations to another, reducing per unit costs. For example, from 1987 to 1992, conventional

department stores' (including independents and national chains) share of general merchandise sales fell from 44.2 percent to 34.7 percent, while discount stores' (discount department stores and miscellaneous general stores) share rose from 52.3 percent to 61.6 percent (the balance was variety stores). Discount stores typically provide fewer sales personnel than the conventional department stores. But the rapid increase in share of discount stores argues that, as with McDonald's burgers, consumers received substantial consumer surplus. From 1980 to 1990, the share of conventional supermarkets in grocery sales fell from 73.1 percent to 34.9 percent, while the share of combination stores (stores that combine additional departments, such as drug, delicatessen, and bakery units) rose from 21.7 percent to 47.5 percent and that of warehouse stores rose from 5.2 percent to 17.6 percent.

Moulton (1996) makes a calculation that the change in outlets in retail food markets has had only a minimal effect (0.1 percent annually) on food price inflation, using prices at warehouse outlets and multiplying by the change in shares of such outlets. But the big change in retail direction has been not to warehouses but to superstores, as we shall see below. And it is possible for superstores to offer greater variety with higher markups on new items and low average prices on commodity items, through devices such as specials and double coupon discounting.

III. GROCERY STORES

Among all consumer prices, food prices have been collected for the longest time by economists. The following extended example shows that our measures of food prices (narrowly defined here as food purchased for consumption at home) went dramatically awry beginning in 1978. The argument takes the form of a reductio ad absurdum: our official statistics imply that

the real output of retail services at supermarkets fell dramatically, but direct measures of supermarket services rose substantially over this period.

The Bureau of Labor Statistics has been collecting monthly data on food prices since World War I, when the CPI, then called the cost of living index, was institutionalized. Prior to 1978, the prices collected were for the same goods and services across all the cities surveyed. Price inspectors throughout the country would collect prices for "milk, delivered, glass bottles," or "bacon, first quality, hand sliced." Imposing a uniform definition nationally poses some problems. Over long periods of time the quality of these goods might well vary, and indeed the products might disappear altogether. Milk might be rich or watered or sour; first quality bacon in one city might be second quality in another. And delivered milk has become a rare commodity in most cities.

In 1978, a new methodology replaced this uniform national specification of products by decentralized specification of products. Price inspectors were asked to define detailed product specifications in the field. The price inspectors were given broad product definitions, such as flour and prepared flour mixes, and a store location based on a nationwide survey called the Consumer Point of Purchase Survey. For example, the Survey and the randomization process might result in the choice of the Acme supermarket at Germantown and Sedgwick in Philadelphia. Then the price inspector, with the help of store personnel, would choose several possible items, and using scientific sampling pick one, say, Betty Crocker Chocolate Fudge cake mix. For the next five years the item priced by the price inspector would be that particular item at that particular store (unless the store stopped carrying that item or closed).

The BLS also collects and publishes average price (AP) data on a selected group of foods.

This is a separate series that prices products (such as flour, white, all-purpose) that are relatively broadly defined when compared to the very narrow product-store combinations priced in the CPI. The AP series gives the average unit prices for these products (typically, prices per pound). The prices are weighted by the relative sales of the outlets at which they are collected. The AP series is apparently piggybacked on the CPI data, in the sense that the basic data in the AP series are, to the extent possible, taken from the CPI collections.

The AP series, it should be pointed out, is essentially what economists have typically collected historically. The AP series (except for a break from 1978 to 1980) is available going back to 1890 (for nine foods).

Before the introduction of this new methodology in 1978, the CPI series and the AP series showed no systematic tendency to diverge. An economist at the BLS, Marshall Reinsdorf, published an article in 1993 that has become one of the seminal articles in the area of CPI price mismeasurement. He discovered that from 1980 to 1990, the CPI and AP series for comparable products diverge by roughly 2 percentage points a year, with the CPI series rising faster than the AP series. As can be seen in Table 5, the divergence over a recent 6-year period is quite substantial for many of the products -- and the divergence is almost universally in the same direction. And as seen in Table 6, the roughly 2 percentage point a year divergence between the two series continues to January 1996. Reinsdorf (1994) reweights the AP series to make it comparable to the total food-at-home category and finds that the price divergence shrinks but remains substantial at 1.4 percentage points a year.

In principle, there are two reasons the CPI and the AP series might diverge. One is that customers may be switching to lower quality goods within each product category. The other is

that customers may be switching to less costly outlets for goods. And there is an additional technical reason: the method that the BLS used to reweight goods when it updated its sample was biased in the absence of the law of one price. This so-called "formula bias," which apparently accounted for 1/2 percentage point a year of the 1 ½ to 2 percentage point annual divergence, was corrected in January 1995. Formula bias itself is a product of the failure of the law of one price.

One possible reason for the CPI to rise more rapidly than average prices is if consumers are shifting to lower quality foods. We would have evidence of a switch to lower quality goods if the CPI rate of increase were mirrored by an increase in the PPI for comparable goods. It is not. The CPI series for food at home grows 1.4 percentage points faster from 1977 to 1992 than does the PPI series for consumer food (Table 7).

Another possibility is that supermarkets' retail services could be declining rapidly, if, for example, variety were decreasing or service personnel were declining or if stores were becoming more cramped as a result of changes in format. This is also not the case. There has been some switch to discount warehouse type stores, as shown in Table 8, but the greater switch has been to the superstore format, in which the supermarket sells extensive additional lines of goods, such as drugs, and provides additional services, such as a deli counter, fresh fish, flowers, and even banking. In this enlarged format, supermarkets are larger (Table 9), stock more items (Table 10), and have more employees (Table 11). While some of the growth in number of products is due to a shift toward more drugs and other nonfood products, most of it appears to be due to an increase in variety of food products.

Consider the following. We can use the CPI for food commodities to deflate food store sales for 1992 to measure the real value of food products and retail services delivered to

consumers. Similarly, we can use the PPI for finished consumer foods to deflate 1992 food store goods purchases to get a measure of the real value of products farms and manufacturers delivered to food stores. The difference should be real retail services added by the food stores: the economic contribution of supermarkets. This calculation, based on Table 12, is shown in Table 13, when we use this so-called "double-deflation" methodology to estimate the real contribution of supermarket output. The implication of our official statistics is that food store output has been declining at a 7.7 percent annual rate. This is absurd, because as I have shown along a variety of dimensions, food store output has been increasing.

In short, the CPI attributes declining real output to a retail segment that by every conceivable measure has been rapidly providing an ever greater abundance of value-added services. This unreasonable result is the outcome of the clash between the CPI methodology put in place in 1978, and the fact that foods do not obey the law of one price in our current retail environment.

Independents vs. chains. Between 1954 and 1974, the shift from independent ownership of supermarkets to chain ownership proceeded very slowly. The sales share of independents declined from 42 percent to 38 percent, or roughly 10 percent. From 1974 to 1994, that sales share declined from 38 percent to 26 percent or by nearly one-third.

During the past decade and a half, chains have adopted information technology more rapidly, notably indexed by more rapid adoption of scanners. Chains also have expanded their hours of operation and sales floor area more rapidly and increased the amount of employment per transaction (Table 14). Thus all these indicators suggest a steady improvement in the service provided to shoppers, rather than a decline in such service. And the chains, which provided

more of these services, expanded sales at the expense of the independents.

In sum, grocery stores have exhibited a high rate of innovation, and this innovation has taken the form of increased services to customers. Customers have rewarded the innovators by shifting swiftly towards them. The view that grocery stores retail services have been declining seems simply untenable, and appears to be a product of substantial mismeasurement.

IV. CONCLUSION

Price discrimination has become very widespread in retailing. Other rapidly changing aspects of retailing include hours of operation, increases in product variety (rapid increases in store-keeping units and UPC codes), information exchange technology (scanners and electronic data interchange), inventory management (just-in-time inventory techniques and inventory management by manufacturers), retail outlets (buying clubs and category killers), and retail environments (regional malls and selling floor space). The speed of these changes in retailing, which themselves are in large part due to reduced costs of information processing, communication, and transportation, weakens the a priori case for the standard method of measuring inflation.

The computerization of retailing has made price dispersion a norm in the United States, so that any given list or transactions price of a product is an increasingly imperfect measure of its resource cost. As a consequence, measuring the real output of retailers has become increasingly difficult. Indeed, the very substantial revison of the CPI in 1978 may have worsened our estimates of the inflation rate because it failed to take sufficient account of the failure of the law of one price. Food retailing is used as a case study to examine data problems in retail productivity measurement. Crude direct measures of grocery store output suggest that the CPI for

food-at-home may have been overstated by 1.4 percentage points annually from 1978 to 1996. Food-at-home is the area of pricing where economists and government statisticians have had the most experience; these goods are the ones where we have the best data and have concentrated most of our efforts in pricing. Errors in other areas of pricing are likely to be even larger; preliminary studies of other areas tend to confirm this a priori estimate (Nakamura, 1997).

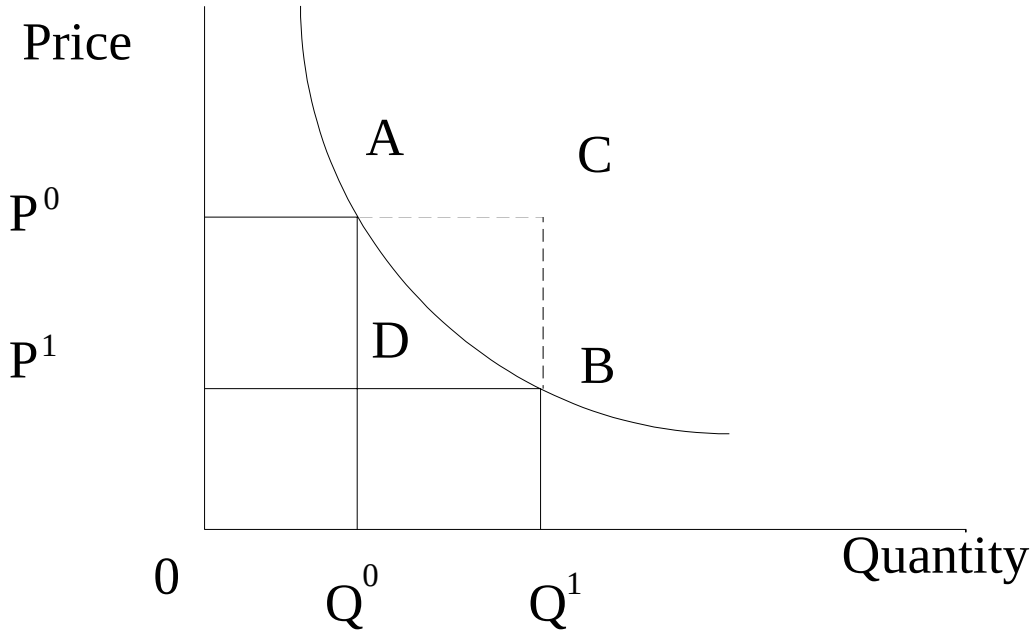

Figure 1. Marshallian consumer surplus

Figure 2 . Matrix of price-quality combinations from Shapiro Wilcox				
		Incumbent		
		HH	LH	LL
	LH	1	2	
Entrant	LL	3		5
	HD	4		

Table 1. Actual Purchases						
	1980			1990		
Item	Price	Quantity	Total	Price	Quantity	Total
Old Good	$10	1000	$10000	$12	5	$60
New Good	$5	12	$60	$5	2000	$10,000
Total			$10,060			$10,060
Average (unit) price	$9.94	1012	$10,060	$5.02	2005	$10,060

Table 2. CPI with and without law of one price						
	1980			1990		
Item	Price	Quantity	Total	Price	1980 Quantity	Total
Old Good	$10	1000	$10,000	$12	1000	$12,000
New Good	$5	12	$60	$5	12	$60
total		1012	$10,060		2005	$12060
CPI with law of one price	1980 set to 100		100			119.9
CPI without law of one price (unit price)	1980 set to 100		100			50.5

Table 3. Output with and without law of one price, 1980 = 100		
	1980	1990

Item	Price	Quantity	Total	Price	Quantity	Total
Industry with law of one price	100	$10,060	$10,060	119.9	$8390	$10,060
Industry without law of one price	100	$10,060	$10,060	50.5	$19,921	$10,060

Table 4. Airfares

		1964	1978	1996	Annual growth, 1964-78	Annual growth, 1978-96
CPI, annual average	1982-84=100	23.7	45.5	192.5	4.8 %	8.3 %
yield, cents per passenger-mile	full fare	6.1 ¢	8.5 ¢	38.9 ¢	2.4 %	8.8 %
	average			13.7 ¢		2.7 %
	restricted			12.0 ¢		2.0 %

Sources: BLS and Air Transport Association.

Table 5. Average Prices of Foods Consistently Rise Less than the Consumer Price Index for the Same Foods

Selected Foods Average Prices Per Pound, In Dollars				Consumer Price Index	
Category	Jan 1989	Jan 1996	% increase	Category	Jan 1989 to Jan 1996 % increase
Flour, white, all purpose	$0.23	$0.26	14.9%	Flour and prepared flour mixes	27.7%
Ground chuck, 100% beef	$1.81	$1.80	-0.4%	Ground beef, excluding canned	7.9%
Bacon, sliced	$1.81	$2.14	18.5%	Bacon	33.9%
Chicken, fresh, whole	$0.91	$0.94	4.0%	Fresh whole chicken	9.4%
Eggs, grade A, large	$0.94	$1.15	22.7%	Eggs	30.1%
Apples	$0.73	$0.88	20.3%	Apples	39.4%
Oranges,navel	$0.52	$0.56	7.7%	Oranges, including tangerines	46.4%
Lettuce, iceberg	$0.79	$0.77	-3.1%	Lettuce	12.2%

Source: U.S. Bureau of Labor Statistics, CPI Detailed Report, January 1989 and January 1996.

Table 6. Average Prices compared to Consumer Price Index

	1980 to 1989 (Reinsdorf)	1989 to 1996 (Nakamura)
Average Prices, Selected Foods	2.1%	1.2%
CPI, Same Selected Foods	4.2%	3.3%
Difference	2.1%	2.1%

Source: Reinsdorf, 1993, and U.S. Bureau of Labor Statistics, CPI Detailed Report, January 1989 and January 1996.

Table 7. Comparison of CPI and PPI for foods

	PPI, consumer foods 1977=100	CPI, food at home 1977 = 100	PPI, annual rate of growth from previous period	CPI, annual rate of growth from previous period
1959	47.4	46.7		
1977	100	100	4.2%	4.3%
1992	168	205	3.5%	4.9%

Source: Economic Report of the President, 1997.

Table 8. Grocery Supermarkets by type

Percent of total	1980	1990	1993	1994
Conventional	73.1%	34.9%	28.0%	28.2%
Superstore	21.7%	47.6%	55.2%	56.6%
Warehouse	5.2%	17.6%	16.8%	15.2%
Total (billion $)	$157	$260	$281	$289

Source: Statistical Abstract of the U.S., 1996

Table 9. Selling Floor Space

(million sq ft)

	1972	1977	1987	1992
Grocery	545.7	606.1	747.6	844.1

Source: U.S. Census of Retail Trade, various years

Table 10. New Product Introductions and Number of Types of Items Stocked, Grocery Supermarkets

Year	New Product Introductions	Items per store	Items stocked Ind	Items stocked Chains
1960		6000		
1964	1281	6900		
1970	1365	7800		
1975	1831			
1980	2689	9400		
1982			9339	11382
1983			9629	10883
1985	7330			
1990	13244	16500	11611	17901
1992	16790			
1993			15,751	20,299
1994		19,612	15,957	21,949

Source: Progressive Grocer, various issues, U.S. Statistical Abstract, 1996, and Moody, 1997.

Table 11. Employment in Grocery Store Retail Industry, thousands			
	1983	1993	% Change
Total	2234	2852	27.6
Exec and admin	175	122	-30.1
Sales	933	1243	33.2
Admin support	611	770	26
Service occup	185	315	69.6
Other	329	402	22.2

Source: Moody, 1997

Table 12. Food Stores, Sales, Margin and Payroll (Millions of Dollars)						
	Sales	Gross Margin	Annual Payroll including fringe benefits	Margin as Percent of Sales	Payroll as Percent of Sales	Non-Payroll Margin as Percent of Sales
1977	157,940	36,651	18,565	23.2%	11.8%	11.4%
1982	240,520	58,623	32,433	24.4%	13.5%	10.9%
1987	301,847	77,200	39,202	25.6%	13.0%	12.6%
1992	377,099	96,206	52,373	25.5%	13.9%	11.6%
1992 in 1977 dollars	179,115	11,116				

Source: Census of Retail Trade, U.S. Department of Commerce

Table 13. Measures of Output and Hours: Food Stores Annualized Growth Rates in Percent				
	BLS Hours	BLS Output	Double Deflation Output	Double Deflation Output with 3.5 % CPI inflation rate for food
1977-92	1.7 %	0.9%	-7.7 %	4.1 %

Source: BLS, Productivity Measures for Selected Industries and Government Services, July 1996, Bulletin 2480, and author's calculations.

Table 14. Performance Measures for Grocery Stores (Independents/Chains)					
	1982	1983	1990	1993	1994
Scanners	18/26	22/38	61/80	75/91	80/95
Hours Per Week	89/102	93/107	102/125	103/130	102/131
Selling Area (000 sq ft)	13.1/20.6	13.3/21.3	14.8/25.3	15.9/29.1	16.4/31.6
Weekly transactions per full-time equiv employee	253/255	257/245	231/214	233/196	228/202

Source: Moody, 1997

References:

Bluestone, Barry, Patricia Hanna, Sarah Kuhn, and Laura Moore, 1981. The Retail Revolution: Market Transformation, Investment, and Labor in the Modern Department Store, Auburn House (Boston) 160 pp.

Caves, Richard E., Michael D. Whinston, and Mark A. Hurwicz, 1991. "Patent Expiration, Entry, and Competition in the U.S. Pharmaceutical Industry," Brookings Papers on Economic Activity, Microeconomics, 1-66.

Chamberlin, Edward Hastings, 1948, The Theory of Monopolistic Competition, 6th edition, Harvard University, Cambridge, Mass.

Denison, Edward, 1962, The Sources of Economic Growth in the United States and the Alternative Before Us. New York: The Committee for Economic Development.

Dulberger, Ellen R. “Sources of Price Decline in Computer Processors: Selected Electronic Components,” 1993, in Murray F. Foss, et al., eds., Price Measurements and Their Uses, NBER (Chicago) 103-124.

Fisher, Franklin M. and Zvi Griliches, 1995. "Aggregate Price Indices, New Goods, and Generics," Quarte

rly Journal of Economics 110, February, 229-244.

Grabowski, Henry G. and John M. Vernon, 1992. "Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act," Journal of Law and Economics 35, October, 331-350.

Griliches, Zvi, and Iain Cockburn, 1994. “Generics and New Goods in Pharmaceutical Price Indexes,” American Economic Review 84, December, 1141-56.

Jevons, W. Stanley, “The Theory of Political Economy,” in George W. Wilson, ed. Classics of Economic Theory, 1964, Indiana University, Bloomington, Ind., 537-566.

Lazear, Edward P., 1986, “Retail Pricing and Clearance Sales,” American Economic Review 76, March, 14-32.

Moody, Gerald R., 1997, “Information Technology and the Economic Performance of the Grocery Store Industry,” U.S. Economics and Statistics Administration, Office of Policy Development, mimeo, January.

Moulton, Brent R., “Bias in the Consumer Price Index: What is the Evidence,”1996. Journal of Economic Perspectives, Fall, 159-178.

Nakamura, Leonard I., 1997. “Is U.S. Economic Performance Really That Bad?” Federal Reserve Bank of Philadelphia,

mimeo, February.

Pashigian, B. Peter, 1988, “Demand Uncertainty and Sales: A Study of Fashion and Markdown Pricing,” American Economic Review 78, December, 936-953.

Pashigian, B. Peter and Brian Bowen, 1991. “Why are Products Sold on Sale?: Explanations of Pricing Regularities,” Quarterly Journal of Economics 106, November, 1015-1038.

Price Statistics Review Committee, 1961. The Price Statistics of the Federal Government, NBER, New York.

Reinsdorf, Marshall, 1993. “The Effect of Outlet Price Differentials on the U.S. Consumer Price Index,” in Murray F. Foss, et al., eds., Price Measurements and Their Uses, NBER (Chicago) 227-254.

Reinsdorf, Marshall, 1994, “Price Dispersion, Seller Substitutions and the U.S. CPI,” BLS Working Paper 252, March.

Reinsdorf, Marshall, Paul Liegey, and Kenneth Stewart, 1996, “New Ways of Handling Quality Change in the U.S. Consumer Price Index,” BLS Working Paper 276, Novem

ber.

Scherer, F. M., 1993. "Pricing, Profits, and Technological Progress in the Pharmaceutical Industry," Journal of Economic Perspectives 7, Spring, 97-115.

Shapiro, Matthew D. and David W. Wilcox, 1996, “Mismeasurement in the Consumer Price Index: An Evaluation,” NBER Macroeconomics Annual.

Shepard, Andrea, 1991, “Price Discrimination and Retail Configuration,” Journal of Political Economy 99, February, 30-53.

Stigler, George J., 1961. “The Economics of Information,” Journal of Political Economy 69, 213-225.

Temin, Peter, 1992. "Realized Benefits from Switching Drugs," Journal of Law and Economics 35, October, 351-370.