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. 98-5

THE MEASUREMENT OF RETAIL OUTPUT AND THE RETAIL REVOLUTION

Leonard I. Nakamura
Federal Reserve Bank of Philadelphia

March 1998

WORKING PAPER NO. 98-5

The Measurement of Retail Output and the Retail Revolution

Leonard I. Nakamura
Federal Reserve Bank of Philadelphia

March 1998

The views expressed here are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or of the Federal Reserve System. I would like to thank Erwin Diewert, Alice Nakamura, and Pierre Fortin for comments and encouragement, Marshall Reinsdorf, Rick Lang, Loretta Mester, and two anonymous referees for detailed comments, and participants at seminars held by the Federal Reserve Bank of Philadelphia, the U.S. Department of Agriculture's Economic Research Service, 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 held in Ottawa, Canada, for their comments and support. Denise Duffy and Victoria Kruglikov 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.

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)

NOTE: This paper makes references to figures not available with this electronic version. If you'd like a hard copy of this paper, please call the Philadelphia Fed's Publications Desk at 215-574-6428 or send an email to Lois.Newell@phil.frb.org.

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, 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. But if retail outlets are faced with allocating substantial fixed costs across commodities, retail pricing practice may result in widespread violation of the law of one price. And this may lead conventional price and output measures to be systematically biased.

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 (an early discussion is in 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 not particularly relevant to the production cost of the product are used to allocate costs to appropriately elastic consumers. For example, whether 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 the four-day weekend means that shoppers generally have higher elasticities because they have more time for search. In both examples, from the perspective of productivity measurement, as a first approximation 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.

This paper builds on the work of Bliss (1988), who argues that competitive retailers offer Ramsey pricing schedules to consumers because the retailers have lump-sum costs of staying in operation which they must distribute over consumers. Consumers have fixed costs associated with visiting a particular retailer; the retailer must overcome these costs by offering a basket of goods that the consumer finds justifies the trip. Intertemporal and interstore price dispersion are natural outcomes of these constraints.

In this paper, I investigate the consequences of Ramsey pricing and its consequences for measured inflation and output, primarily using the example of grocery stores. What makes the problem difficult is that price dispersion forces us to explicitly consider consumer heterogeneity. In practice, the U.S. Bureau of Labor Statistics has typically finessed this problem by selecting the highest price among the dispersed prices rather than the average or unit price. The theoretical section shows that the BLS practice creates first order distortions, while the average or unit price method creates only second order distortions. The third section shows that the difference between the CPI and the average or unit price for retail food pricing has had very substantial consequences for our estimates of real output and inflation measures for food.

II. MODEL

In this section we set forth a model in which a change in technology permits a store to differentiate a product and charge different prices for the two versions of the product. The product differentiation does not affect the cost of producing the good or its value in consumption. Consider a store with two types of customers, *A* and *B*. Type A customers (who have mass 1) have utility equal to $u_A(q) + x$, where x is the numeraire good, and type B customers (mass= 1) have utility equal to $u_B(q) + x$, where we assume that the u_i are twice differentiable and locally quasiconcave. Each type of customer is endowed with E units of the numeraire good. We assume that at any given price, type B customers have a more elastic demand than type A customers. In the base period, date 0, the store sells its product, q, whose cost is c, to both types of customers for the same price, p. Then a new technology arrives, at date 1, which allows the retailer to differentiate the product for a fixed cost D (which we shall generally set equal to zero) and sell it at different prices, p_A and p_B, to the two types of customers.

Let us illustrate the general mode of solution, using the base period with a single price. We can determine Marshallian demand functions $h_i(p)=q$ for each type of customer by taking the partial differential of the utility function, setting it equal to price, and inverting. The indirect utility functions are $v_i(p,E) = u_i(h_i(p))+E-ph_i(p)$. The expenditure function for reaching utility u at price p is $e(p,u) = u - u_i(h_i(p))+ph_i(p)$. The store earns profits $(p-c)(h_A(p)+h_B(p))$. These profits are returned to the shareholders as dividends. Each type of customer owns half the shares. The total utility of customers will be

$W = 2 E + \sum u_i(h_i(p)) - ch_i(p)$. The total utility of purchases at the retail stores will be $\sum u_i(h_i(p))$, and the real contribution of the retail stores to utility can be measured as $\sum u_i(h_i(p)) - ch_i(p)$.

We will consider two polar cases: the store has monopoly power over its market, on the one hand, and it faces competition, on the other. However, in both cases the store must cover a fixed cost, R, out of its sales margins.

Monopoly power over market. Before the new technology arrives, the store maximizes profit = $(p-c) (h_A(p)+ h_B(p))$ by the standard method of setting $(p-c)(h_A'(p)+ h_B'(p))/ (h_A(p)+ h_B(p)) = - 1$.

Now consider that the store has the option of differentiating its product. It will do so if, setting $(p_i-c)h_i'(p_i)/h_i(p_i) = - 1$ for $i = A, B$, $(p_A-c) h_A(p_A)+(p_B-c) h_B(p_B) - D>$ $(p-c) (h_A(p)+ h_B(p)) > R$. Since type A's demand is more inelastic, $p_A > p > p_B$. For small enough D, differentiation will always pay in the monopoly case.

In this case, the monopolist uses the expenditure D to extract rents from the two types of customers. In extracting rents, it makes type A customers worse off and type B customers better off. Depending on the relative sizes of the two groups, it is possible that this rent- seeking behavior leaves customers much worse off and does not improve the monopolist's profit very much.

Limited market power. In this case, the store maximizes the consumer surplus of customers subject to a revenue constraint. This is the Ramsey pricing problem as discussed in

Bliss (1988).[1] Before the new technology arrives, the store minimizes p subject to $(p\text{-}c)((h_A(p)+ h_B(p)) = R$. Afterwards, the store maximizes an aggregate of the utilities of the two types of customers ($W=W(v_A,v_B)$) subject to the revenue constraint $(p_A\text{-}c)\ h_A(p_A)+(p_B\text{-}c)\ h_B(p_B)\text{-}D = R$.

In this case, the store is no worse off, and the consumers, at least as measured by the store's aggregate welfare measure, are better off. Indeed, if type B consumers have elastic demand at the initial price, the store raises its operating income by lowering p_B and it could then also lower price p_A. In that case, there could be a clear Pareto improvement from being able to differentiate the products. On the other hand, if both types of customers have inelastic demand, price differentiation must result in higher prices to the type with more inelastic demand.

What happens to price and output measures? Consider a measure of unit price. Initially, the price per unit of the good is p. Afterwards, the price per unit (or unit value) is $(p_A\ h_A(p_A)+p_B\ h_B(p_B))/(\ h_A(p_A)+ h_B(p_B))$. The ratio of the unit prices is therefore $(p_A\ h_A(p_A)+p_B\ h_B(p_B))/p(h_A(p_A)+ h_B(p_B))$. This is a Paasche price index, and thus a lower bound on a true cost of living increase as measured by Diewert's (1983) Paasche-Pollak cost of living index.[2] To that extent, unit price measures may be a reasonable approximation of a true cost of living

[1]Retailers must have some market power to engage in Ramsey pricing. But market power is limited by the possibility of entry. The constraint on retailer profits can be viewed either as a measure of market power or as an outcome of equilibrium entry with sunk costs.

[2]The theoretical Paasche-Pollak index is a ratio whose numerator is the sum of the expenditure functions of the two types at final period prices and whose denominator is the sum of the expenditure functions at base period prices, evaluated at the utility levels of the final period. Note that the Paasche index here is a ratio of observables: sales and quantity data from period 1, and price data from period 0; one does not need to know the schedules for the demand functions $h_i(p_i)$. The necessary data can be obtained from, e.g., the Nielsen data used by MacDonald (1995.)

index, although biased downward.

A Laspeyres price index could be constructed if we had information about $h_A(p)$ and $h_B(p)$.[3] Then we could use $(p_A\, h_A(p)+p_B\, h_B(p))/p(h_A(p)+ h_B(p))$ as an upper bound on a true cost of living index as measured by Pollak’s (1981) Scitovsky-Laspeyres index.[4] These bounds hold true for both market power conditions discussed above, as can be easily verified by the expenditure functions.

Similarly, it is straightforward to show that the total output ratio is a Laspeyres quantity index and an upper bound on a true welfare improvement measure defined as the ratio of expenditure functions for base period prices, with the numerator evaluated at final period utility and the denominator at base period utility. In the absence of information about $h_A(p)$ and $h_B(p)$, it is difficult to construct the Paasche output index that is a lower bound on a true welfare improvement measure using base period prices.

How does an agency like the U.S. Bureau of Labor Statistics measure the price ratio? It depends on how the "new" goods q_A and q_B are perceived, which, in turn, depends on the details of how the product differentiation is carried out. In many cases, the product differentiation involves restrictions on the conditions under which good q_B is sold, and in this case the BLS typically takes for its price ratio the change from p to p_A, which is a strongly upwardly biased measure.

In figure 1 we set forth three examples of price differentiation, using linear demand. In

[3]In contrast to the Paasche index, the Laspeyres index requires knowledge of the demand schedules to compute $h_A(p)$ and $h_b(p)$, period 1 demands evaluated at period 0 prices.

[4]The theoretical Scitovsky-Laspeyres index is a ratio whose numerator is the sum of the expenditure functions of the two types at final period prices and whose denominator is the sum of the expenditure functions at base period prices, evaluated at the utility levels of the base period.

all three examples, the type A consumers have utility $U_A = 15\ q_A - 1/8\ q_A^2 + x$, and type B consumers have utility $U_B = 11\ q_B - 1/40\ q_B^2 + x$, and c = 7. The first two columns show the impact on a monopoly retailer of the opportunity to differentiate the product; the first column shows the single price monopoly profit maximization, and the second, the two price monopoly profit maximization. In the monopoly case with linear demand, total demand does not change as price differentiation is permitted, but the distribution across consumers changes for the worse. However, the impact on utility is minor, roughly 1 percent.[5] The small impact on utility can be understood in light of the fact that, if we hold quantity fixed, the distribution under one price is optimal. The envelope theorem tells us that a small change in price in the vicinity of the optimum has no first order effects.

In columns 3 and 4, we illustrate the competitive case with large fixed costs. In this case, the fixed costs equal the monopoly profits of the first case, and fixed costs are 25 percent of total revenues. This case approximates the actual margins as a percent of sales reported for food stores in the U.S. Census of Retail Trade. The third column shows the single price case, and the fourth column shows what happens when the retailer can, with price differentiation, maximize the sum of the utilities. By relaxing the price restraint, the retailer is able to set Ramsey prices, raising total quantity sold by 32 percent. Total quantity has a first order impact on utility of the retail good, which rises 25 percent.

Not all of this utility gain is attributable to the retail sector, however, as the additional

---

[5]Note that this case has been constructed so that utility is cardinal -- marginal utility is held constant by the nonretail good's linearity. The remarks that apply to this cardinal utility measure can easily be seen to apply to an expenditure function ratio measure of the type discussed above.

quantity sold requires additional production at a cost of 7 units of the numeraire each. The net utility gain is just over 30 units.

In columns 5 and 6, we illustrate the competitive case with smaller fixed costs. Here fixed costs are 12.5 percent of total revenue. The retailer's flexibility does not have as large an impact on quantities and utility, which rise 6 percent and 5 percent, respectively. Again we see that the impact of the quantities on utility is first order.

Figures 2, 3, and 4, show the demand curves for the two types of customers and the demand curve for the one price case. An x marks the total demand for the two price case.

In Figures 5, 6, and 7, we examine measures of real output and retail value added, using two measures of price: unit price and the price of good A. Figure 5 shows the monopoly case. Column 1 shows the base period nominal expenditures with a single price for both types of customers. Column 2 shows nominal expenditures in the case with price differentiation. Unit price has risen to 9.57 from 9.33 as a consequence of the greater monopoly rents the retailer is able to extract. Column 3 deflates column 2 to base period prices. The price of the goods purchased by the retailer (the cost of goods in row 2) is unchanged from the base period, and units are unchanged from the base period, so retail output is measured to be unchanged in real terms. In fact, if we recall that consumer utility has fallen 1 percent, or 7 units of the numeraire, this overestimates the contribution of the retailer to welfare, but the mismeasurement is relatively small; 7 units on a base of 130.67 is roughly 5 percent. Column 4 uses the BLS procedure of deflating using the higher price, p_A, which has risen to 11 from 9.33. Deflating revenue with the higher price implies that real retail value added has fallen by more than half, or 68 units -- an exaggeration by an order of magnitude.

Figure 6 shows the first competitive case. Deflating using unit values results in a retail value added, in real terms, of roughly 40 units, or an overstatement of about 10 units. Deflating using the high price leads to an implied loss in retail value added of 30 units, an understatement of about 60 units. In figure 7, deflating by unit value in the second competitive case gives us an overstatement of 2 units, while deflating by the high price gives an understatement of 40 units.

These figures illustrate that while unit prices provide modest overestimations of utility, the utility losses associated with price differentials in this range are trivial compared with the first order errors created by deflating using the BLS procedure. This underscores the problems that our price measures have had coping with technological change in retail sales.

An obvious example is airfares, where airlines differentiate between low elasticity business travelers and high elasticity vacation travelers by requiring a Saturday night stayover. In the United States, until the deregulation of airfares in the late 1970s, airlines were compelled to charge uniform fares to passengers. Once airfares were deregulated, major airlines instituted computerized reservation systems that permitted extensive price dispersion. The Bureau of Labor Statistics called the unrestricted fare, p_A, the same good as the standard coach fare, p, and interpreted the discount fares, $p_{B,}$ as new goods that had no weight in the index. Later, the discount fares were included in the index, at a revised fixed weight (eventually raised as high as 90 percent). As can be seen in table 1, the CPI relatively closely tracked the full fare from 1978 to 1996, rather than the unit price per passenger mile.

In general, when a generic or house brand alternative to a brand-name good is offered as a means of price discrimination, the new product is designated a new good even though its contents may be identical to the national brand. For example, until a change made in 1994, the

BLS considered brand-name prescription drugs a different good from the generic prescription drugs that were permitted once the brand-name drug's patent expired (Scherer, 1993). Since 1994, the BLS has adopted a procedure for pricing prescription drugs that should approximate unit pricing.

In grocery stores, with the advent of scanners and on-shelf pricing, rapid price changes for purposes other than cost changes have become widespread. Stores have very widely adopted a form of product differentiation called high-low pricing, in which goods are given two or more prices, and prices change frequently between the two. At any given moment, a good has a fixed price, but the price may change from week to week. More purchases are made at lower prices than higher prices; some households buy at the lowest price in the local area, but most households have, in the short run, store loyalty that evolves dynamically over time (Slade, 1998). In this case, the BLS price inspectors report the price at the time of their visit to the store. Suppose a good is sold at the low price, p_B, during a proportion t of all weeks, and at the high price, p_A, during the rest of the time. Then, on average, the inspectors will report $t\ p_B + (1-t)\ p_A$. But if demand is responsive to price, the unit price will be

$(t\ q(p_B)p_B + (1-t)\ q(p_A)\ p_A)/(t\ q(p_B) + (1-t)\ q(p_A))$, which will always be lower.

A related problem is that BLS price inspectors price products during the workweek. Some stores, such as department stores, concentrate their discounts on holidays and weekends, when customers' time costs are lower. As such discounts have increased over time (Pashigian, 1988), this practice will contribute to bias.

Most recently, grocery stores have been developing programs that permit them to identify the shopper who buys the goods by offering special discounts to shoppers who identify themselves by becoming members of a "club." This permits the grocery stores to provide prices for goods tailored to the characteristics of the individual customer. Private colleges and universities similarly tailor their prices to the characteristics of the purchaser, using financial aid packages. The BLS price measures track the tuition charge, without allowing for the average discount, which has been increasing over time.

The failure to account for increasing price discrimination -- in the sense that closely related goods are sold for very different prices -- is widespread. 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 George 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 improvements 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.37)

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. 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.

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 also applies to Griliches and Cockburn's (1994) work on brand name and prescription drugs, Pashigian's (1988) work on department store pricing, Shepard's (1991) work on gasoline stations, and to telecommunications and the fast-food market.

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.

Technology and price dispersion. Price adjustment in supermarkets is extensive. Levy et al. (1997) report that in 1991-92 a group of four supermarket chains reported between 3228 and 4278 weekly price changes per store, or roughly 13 to 17 percent of items (estimated to average 25,000). Most of these price changes appear to be due to reasons other than cost changes. Three of the chains reported the proportion of price changes due to cost *increases* to be between one-sixth and one-quarter of all price changes (they do not report the proportion due to cost decreases). Thus, most of the price changes appear to be due to pure price discrimination motives.[6] If we assume that cost increases and decreases were equal in number during this period, the supermarket chains average between 4 and 6 price changes per item annually for reasons other than cost.

The same study reports that a fifth supermarket chain, which faced higher costs of changing prices (2 1/2 times as much) because of a regulation that required prices on every item in addition to the shelf price tag, changed prices on only 6 percent of items each week. Although this supermarket chain did not report cost changes, it presumably faced much the same mix of cost changes as the others. The most likely implication appears to be that the majority of its price changes were due to cost reasons -- probably between 1 and 2 price changes annually. Prior to the widespread adoption of scanners, the marking of prices on each individual package was the norm. If we can infer the impact of the change in costs over time from this cross-sectional comparison, the adoption of the new technology resulted in a substantial menu cost decline and greater price changes.

Margaret Slade (1998) studied the dynamic consumption and pricing of saltine crackers at 10 chain grocery stores in Williamsport, PA, using data collected from households in 1984-85. She estimated that price changes cost, on average, about $2.72, considerably more than estimated by Levy et al. Prices in her sample change roughly 1 time every 8 weeks, roughly as frequently as in the Levy et al. study, and costs change essentially not at all. Thus the price changes in her study are also purely strategic.

---

[6]Price increases were likely somewhat more prevalent than price declines. The years 1991 and 1992 saw low but positive inflation. From December 1990 to December 1992, the CPI for food commodities rose at an annualized rate of 1.7 percent while the PPI for consumer foods was unchanged.

BLS procedures and the 1978 revision. 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, 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).

Critiques of CPI data

Average price comparison. 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 with the very narrow product-store combinations priced in the CPI. The AP series gives the prices for these products in cents per physical unit (typically pounds). The prices are weighted by the relative sales of the outlets at which they are collected. However, the AP price measure weights prices by *base period* sales and not by actual current sales, so it is not a unit price measure in the sense described in the model section. The AP series is piggybacked on the CPI data, in the sense that the basic data in the AP series are taken from the CPI collections.

To illustrate the difference between the two series, suppose an existing store sells Gold Medal flour for 20 cents a pound, and a new store starts up that sells the same flour for 15 cents a pound. If the BLS adds the new store's flour price to the data collected for the CPI, its lower price level has no effect on the measured rate of inflation. Only price changes after the item is included affect the measured rate of inflation. But the cheaper flour will lower the AP series.

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 (52 food items) diverge by roughly 2 percentage points a year, with the CPI series rising faster than the AP series. As can be seen in table 2, 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 3, the roughly 2-percentage-point a year divergence between the two series continues to January 1996. Reinsdorf (1994) reweighted the AP series to make it comparable to the total food-at-home category and found that the price divergence shrank but remained 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 1½ 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.

Cost comparisons: Producer price indexes. One possible reason for the CPI to rise more rapidly than average prices is if consumers are shifting to lower quality foods. One means of detecting quality shifts is to look at prices at the wholesale level, to see whether there is a comparable shift in the cost of goods to the retailer. For this, we can turn to the producer price index (PPI). 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 4).

Food retail services. 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 stores, as shown in table 5, 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 6), stock more items (table 7), and have more employees (table 8). 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 that 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 9, is shown in table 10, when we use this "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 an annual rate of 7.7 percent. This is absurd, because, as I have shown, food store output has been increasing along a variety of dimensions.

In Reinsdorf’s studies, 16 of the 52 food items covered by the average price series are fresh fruits and vegetables. The evidence is that much of the discrepancy, at least for fresh fruits and vegetables, is caused by problems associated with price variability and price dispersion. Fresh fruits and vegetables are seasonal products, and their prices rise and fall dramatically from month to month, if the item is available at all. Moreover, their perishability can cause prices to vary dramatically across stores. The formulas that the BLS introduced in 1978 were apparently very vulnerable to these fluctuations and provided upwardly biased measures because of them. But the problems are not confined to fresh fruits and vegetables.

Supermarket "tape" data. A source that permits us to obtain true unit price data is data collected from supermarkets by survey companies like A.C. Nielsen. MacDonald (1995) used the data for nonperishable food products in a study that compared CPI data with supermarket checkout (scanner) data for 1989-94. MacDonald used A.C. Nielsen Company data that report the quantity sold nationwide in a given month for a particular item, as well as the total dollar sales for the item.The advantage of Nielsen data is that they report the quantities sold at different prices, while the BLS's price inspectors report only the particular price they observe, not the amount sold at that price. This permits MacDonald to measure the unit price of the goods studied and compare them to the BLS measures.

MacDonald first analyzed items for which the BLS product categories and the A.C. Nielsen product categories closely correponded, from 1988-91. For each of these 14 groups, the CPI inflation measures were consistently higher compared to the unit price; the average gap was 1.4 percentage points a year. He then looked at a wider array of classes of nonperishable products, comparing annual price changes for the leading brand in each of 323 product classes between April 1988 and April 1993 with the BLS price indexes for these product classes. For this group, the CPI grew at an annual rate of 3.7 percent per year, compared with 1.9 percent for the unit prices of the Nielsen items -- the CPI showed an upward bias of 1.8 percentage points a year. This finding shows that the bias is not confined to seasonal products.

Pounds of fruits and vegetables. Another test of the accuracy of the CPI is to compare nominal measures deflated using the CPI with direct measures of quantity. This is implicitly a unit price comparison. If CPI-deflated output grows more slowly than a pure measure of quantity, we have strong evidence that the CPI is biased.[7]

The U.S. Department of Agriculture computes implicit quantities of U.S. food consumption by weight by adding U.S. production, imports from abroad, and carryover inventory from the previous year, and subtracting exports, processing, nonfood uses, and final end-of-year inventory. These measures are called disappearance estimates. Over the period 1978 to 1988, disappearance data imply that per capita consumption of fresh fruits and vegetables measured in pounds rose 25 percent, or 2.3 percent a year (MacDonald). But deflating U.S. domestic expenditures on fresh fruits and vegetables by the CPI measures for these categories implies that consumption of fresh vegetables *declined* 1.2 percent a year and consumption of fresh fruits declined 0.2 percent a year. Thus, when compared with measures based on disappearance data, the CPI-based measures implicitly underestimate output growth by over 2 percent a year. This discrepancy is a strong argument that the CPI overstated inflation during this period.

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.

---

[7]This assumes that the real value of a unit of output was constant or increased over the period. This seems reasonable, since quality has been rising.

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 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 11). 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 toward them. The view that retail services in grocery stores 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 revision 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 with which economists and government statisticians have had the most experience; these goods are the ones for which we have the best data and on which we 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. 3 Examples of Price Dispersion Effects

	Monopoly over A and B		Competitive Case I		Competitive Case II	
	One price	Two prices	One price	Two prices	One price	Two prices
P Average (unit) price	9.3	9.57	9.3	8.8	8	7.96
p_A Inelastic price	9.3	11	9.3	9.8	8	8.5
q_A Inelastic demand	22.7	16	22.7	20.9	28	26
p_B Elastic price	9.3	9	9.3	8.4	8	7.8
q_B Elastic demand	33.3	40	33.3	52.2	60	65
Q Total demand	56	56	56	73.1	88	91
U_A Inelastic utility	276	208	276	259	322	306
U_B Elastic utility	339	400	339	506	570	609
U Total utility	615	608	615	765	892	915
U-cQ Net utility gain	223	216	223	253	276	278

Figure 5. Monopoly case, Measures of real output				
Prices	1 Price	2 Prices		
Deflation	Undeflated	Undeflated	Unit price	BLS
PQ Revenue	523	536	523	455
cQ Cost	392	392	392	392
Retail Value Added	131	144	131	63

Figure 6. Competitive case I, Measures of real output				
Prices	1 Price	2 Prices		
Deflation	Undeflated	Undeflated	Unit price	BLS
PQ Revenue	523	642	682	613
cQ Cost	392	511	511	511
Retail Value Added	131	131	171	102

Figure 7. Competitive case II, Measures of real output				
Prices	1 Price	2 Prices		
Deflation	Undeflated	Undeflated	Unit price	BLS
PQ Revenue	704	725	728	686
cQ Cost	616	637	637	637
Retail Value Added	88	88	91	49

<table>
<tr><th colspan="7">Table 1. Airfares</th></tr>
<tr><th></th><th></th><th>1964</th><th>1978</th><th>1996</th><th>Annual growth, 1964-78</th><th>Annual growth, 1978-96</th></tr>
<tr><td>CPI, annual average</td><td>1982-84=100</td><td>23.7</td><td>45.5</td><td>192.5</td><td>4.8 %</td><td>8.3 %</td></tr>
<tr><td rowspan="3">yield, cents per passenger-mile</td><td>full fare</td><td rowspan="3">6.1 ¢</td><td rowspan="3">8.5 ¢</td><td>38.9 ¢</td><td rowspan="3">2.4 %</td><td>8.8 %</td></tr>
<tr><td>average</td><td>13.7 ¢</td><td>2.7 %</td></tr>
<tr><td>restricted</td><td>12.0 ¢</td><td>2.0 %</td></tr>
</table>

Sources: BLS and Air Transport Association.

Table 2. 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 3. 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 4. 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 5. 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 6. 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 7. New Product Introductions and Number of Types of Items Stocked, Grocery Supermarkets

Year	New Product Introductions	Items per store	Items stocked Independents	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 8. 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 9. 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 10. 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 11. 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:

Bliss, Christopher. 1988. "A Theory of Retail Pricing," Journal of Industrial Economics 36, June, pp. 375-92.

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.

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

Diewert, W. E., 1983. "The Theory of the Cost of Living Index and the Measurement of Welfare Change," in W.E. Diewert and C. Montmarquette, eds., Price Level Measurement, Ottawa: Statistics Canada, 162-233.

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.

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.

Levy, Daniel, Mark Bergen, Shantanu Dutta, and Robert Venable, 1997, "The Magnitude of Menu Costs: Direct Evidence from Large U.S. Supermarket Chains," Quarterly Journal of Economics, August, 791-826.

MacDonald, James M., 1995, "Consumer Price Index Overstates Food-Price Inflation," FoodReview, September-December, pp. 28-32.

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.

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.

Pollak, Robert A., 1981, "The Social Cost of Living Index," Journal of Public Economics 15, June, 311-336.

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.

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

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

Slade, Margaret E., 1998. "Optimal Pricing with Costly Adjustment: Evidence from Retail-Grocery Prices," Review of Economic Studies 65, 87-107.