NBER WORKING PAPER SERIES

BEST PRICES: PRICE DISCRIMINATION AND CONSUMER SUBSTITUTION

Judith A. Chevalier
Anil K Kashyap

Working Paper 20768
http://www.nber.org/papers/w20768

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2014

The views expressed here are ours and not those of any institutions with which we are affiliated or the NBER. We acknowledge support from the NSF through a grant administered by the NBER. Kashyap thanks the Chicago Booth Initiative on Global Markets and the Center for Research on Securities Prices for research support. We thank David Argente, Conor Devitt, Cecilia Gamba, Aaron Jones, Bruno Pellegrino, Ashish Shenoy, and Rasool Zandvakil for research assistance and many seminar participants for comments. We are grateful to the SymphonyIRI Group for data. As a condition of use, SymphonyIRI reviews all papers using their data to check that the data are not described misleadingly. All analyses in this paper based on SymphonyIRI data are the work of Chevalier and Kashyap, not SymphonyIRI Group Inc. For information on Kashyap's outside compensated activities see http://faculty.chicagobooth.edu/anil.kashyap/. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w20768.ack

Best Prices: Price Discrimination and Consumer Substitution
Judith A. Chevalier and Anil K Kashyap
NBER Working Paper No. 20768
December 2014, Revised August 2015
JEL No. C43,D11,D12,D4,L81

ABSTRACT

We propose a method for constructing price indices when retailers use periodic sales to price-discriminate amongst heterogeneous customers. To do so, we introduce a model in which Loyal customers buy one brand and do not strategically time purchases, while Bargain Hunters always pay the lowest price available, the "best price". We derive the exact price index and demonstrate empirically that accounting for our best price construct substantially improves the match between conventional price indices and actual prices paid by consumers. We demonstrate that our methodology improves inflation measurement without imposing an unrealistically large burden on the data-collection agency.

Judith A. Chevalier
Yale School of Management
135 Prospect Street
New Haven, CT 06520
and NBER
judith.chevalier@yale.edu

Anil K Kashyap
Booth School of Business
University of Chicago
5807 S. Woodlawn Avenue
Chicago, IL 60637
and NBER
anil.kashyap@chicagobooth.edu

I. Introduction

Since the 1970s, a series of technological innovations have lowered the cost to retailers of offering a wide variety of products at rapidly changing prices (see Holmes 2001, Basker et al. 2010). In particular, the adoption of the product scanner has enabled supermarkets and other retail stores to increase the number of stock-keeping units per square foot and to introduce and evaluate strategies such as "High-Low" pricing. Ellickson and Misra (2008) estimate that seventy-two percent of retail supermarkets in 1998 described themselves as employing some type of "High-Low" pricing strategy. As Basker (2013) argues, forty years after the introduction of the scanner, the extent to which the technology is being exploited separates the top-performing retailers from their competitors. Retail firms (and their suppliers) are continually innovating in their attempts to exploit consumer data in order to improve pricing and promotion strategies. In a series of papers (Nakamura (1998), Nakamura (1999)), Leonard Nakamura refers to innovations that enable firms to charge different prices for identical or similar products as the "retail revolution".

While high frequency price variation is one important price discrimination strategy in traditional retail stores, it is far from the only routinely practiced price discrimination strategy. For example, it has become increasingly common for retailers, credit card companies, and travel services to use the internet to target offers and discounts to consumers based on browsing history and geolocation. The increase in IT-enabled price discrimination opportunities has led to a substantial literature on market equilibria in the presence of rich customer data (e.g. Acquisti and Varian 2005, Taylor and Taylor 2004, Burke et al. 2012). Temporary discounts create a challenge for price measurement when varieties are close substitutes. If individual retailers stocked one variety of each product and did not engage in temporary discounting strategies, aggregating prices to obtain a price series for a particular product would be relatively straightforward. The prices for a given product across retailers would be expected to covary positively, as price movements would be driven largely by common cost and demand shocks. When retailers stock a large number of varieties and strategically set prices for price discrimination purposes, the relative prices of different varieties can be quite volatile. Thus, the aggregation methodology will have important implications for price measurement.

In this paper, we explore theoretically and empirically the importance of retailer price discrimination strategies for measuring prices. Different methodologies of accounting for sales lead to different estimates of the cost of living and of inflation rates. Most cost of living discussions take an exact index as a benchmark. An exact index tracks the relative cost of obtaining a given level of utility at different points in time. Constructing an exact index in an environment with high frequency price variation due to sales is challenging in several respects. First, it is commonly understood that the observed phenomenon of frequent temporary discounts reverting to a regular price is likely the result of price discrimination. Price discrimination generally only makes sense when there are different types of consumers to discriminate among. But, different consumer types imply that the price of maintaining a given level of utility over a specific period of time is not well-specified without aggregating utilities across consumers. Furthermore, the measurement horizon can be complicated to define. If consumers respond to price discounts by stockpiling goods, the appropriate period of time over which to measure prices and construct a price index is also not obvious. An important conceptual question, then, is how to weight sales prices vs. regular prices in constructing price indices when aggregating both across products and through time in order to mimic an appropriate exact index.

It is well-understood that a fixed weight (Laspeyres) price aggregation methodology substantially underestimates the importance of periodic sales in reducing a consumer's cost of living. This occurs because a Laspeyres index supposes that there is no substitution in response to discounts, i.e. that the elasticity of substitution between alternative products is zero. In order to better control for consumer substitution, during the 1990s, statistical agencies in a number of countries, including the Bureau of Labor Statistics (BLS) in the U.S., abandoned the Laspeyres formulation and adopted the geometric mean formula in the calculation of the most basic components of the Consumer Price Index (CPI). The geometric mean formula is a constant elasticity of substitution index that implies an elasticity of substitution of one between varieties. However, because different brands of the same commodity experience periodic discounts, past evidence (as well as our own new evidence) suggests that cross-price elasticities of substitution are often much greater than one in absolute value.

This observation has led to many suggested alternatives to the BLS approach. Caves et al. (1982) show, for a representative consumer with a fairly general form of homothetic utility, that a Tornqvist index approximates the exact cost of living index. The Tornqvist index combines the current period expenditure weights with a base period expenditure weights. Feenstra and Shapiro (2003*a*) furthermore propose calculating the Tornqvist over a series of periods to account for the tendency of shoppers to store goods.

Nakamura and Reinsdorf (2014), Diewert (1995), Reinsdorf (1999), and Feenstra and Shapiro (2003*b*) endorse the approach suggested by Walsh (1921), Davies (1924), and Davies (1932) who advocate the use of unit values (the current period sales-weighted average of prices) to aggregate prices of very similar products (such as different brands of a consumer product). This unit value approach corresponds exactly to an exact cost of living index only when all consumers have identical preferences and experience no disutility in substituting between alternative brands, sizes, and varieties. Additionally, the unit value approach could overstate the cost of living in the presence of promotional activity because this approach does not directly account for the disconnect between the decision to purchase and to consume that could arise from strategic storage. Nakamura (1999) demonstrates assumptions under which the unit value representation would, even with heterogeneous consumers, more closely reflect variations in consumer utility than would the BLS' measurement approach.

There are two important limitations to the existing literature on aggregating prices for price indices in the presence of price-discriminating discounts. First, the approaches advocated by Diewert (1995), Reinsdorf (1999), and Feenstra and Shapiro (2003*a*) arose in part out of enthusiasm for the possibility that statistical agencies would begin to exploit the same scanner datasets that retailers and manufacturers were using. However, this enthusiasm appears to have waned in recent years. While researchers routinely analyze high frequency quantity data from scanner datasets, statistical agencies such as the BLS have chosen not to rely on commercial provision of high frequency quantity data. The impediments to employing commercial scanner data, some of which are detailed in Bradley et al. (1997), Reinsdorf (1999) and Triplett (2003), have not yet been resolved. Hence, the BLS continues to primarily obtain retail data by having BLS agents visit stores to sample prices.[1] Thus, while the suggestion of using variable weight indices at the lowest aggregation level (such as in constructing unit values or the Tornqvist) is simple in theory, it is difficult to implement in practice because the CPI data collection methodology does not involve frequently gathering quantity or expenditure data. Any proposed alternatives to the current CPI methodology will be more practical if they do not rely on using high frequency quantity or expenditure data.

Second, the industrial organization literature explains sales as a vehicle for price discriminating amongst consumers with different preferences. However, the extant utility maximization models that provide foundation for various cost of living constructions rely on representative consumer models. The existing literature is silent

[1]Scanner data is used to construct a CPI elsewhere, as in the construction of the Dutch CPI.

on the question of how to specify an exact cost of living index in an environment with heterogeneous consumers; neither the Tornqvist nor the unit value construction necessarily approaches an exact index with heterogeneous consumers. The only treatment of this issue of which we are aware is Nakamura (1999); he uses a two-type consumer model to demonstrate the poor performance of the fixed weight index relative to usage of unit values in measuring real output, but he does not provide a specification of an exact cost of living index. Furthermore, while our approach is informed by the optimal retailer response to consumer heterogeneity (second degree price discrimination), the existing approaches are not utilizing the supply side to inform the price aggregation methodology.

In order to address the question of how to summarize prices in the presence of intermittent discounts and heterogeneous consumers, we present a very simple model of sales motivated by price discrimination. Our model of sales is similar in spirit to Varian (1980), Salop and Stiglitz (1982), Sobel (1984) and Pesendorfer (2002). In particular, in our model, we posit that that some consumers are active shoppers who chase discounts, substitute across products in a narrowly defined product category, and potentially use storage to maintain smooth consumption whilst concentrating their purchases to take advantage of discounts. Other customers are passive, and retailers will employ strategies to charge these two groups different prices. Due to the actions of these strategic consumers, we find that weighted average prices paid differ substantially from posted prices. In contrast to the previous literature on sales, we focus explicitly on a retailer controlling the prices of multiple products. The robust implication of this kind of model is that store pricing patterns ought to reflect the presence of different consumers and be strategically coordinated across products. The model can account for frequent temporary discounts and long periods of constant regular prices.

We then examine the empirical validity of our model using detailed store-level data. We use a national dataset of supermarket prices for 2001 to 2011 provided by Symphony IRI. The IRI dataset covers stores in 50 markets around the country. Prices for individual products at many of the IRI stores display the now-familiar pattern of very infrequent regular price changes combined with frequent temporary discounts. Using the IRI data, and focusing on two representative categories of products which the BLS also tracks, peanut butter and coffee, we conduct a store-level analysis in different cities across the country. This analysis is designed to test three implications of the model.

First, our model implies that, depending on the functional form of storage costs, unit values aggregated over time are, or approach, the exact index. The intuition behind this result is clear. The retailer's second degree price discrimination motive induces the retailer to charge prices that minimize the possibility of (and in our stark model eliminate) consumers with strong brand preferences choosing to switch to less-preferred goods to save money. Thus, substitution between brands, when it is actually observed, is not associated with any utility consequences.

Empirically, we show that some consumers chase discounts, and thus, actual prices paid are substantially lower than regular prices, and even substantially below a fixed weight index of the average posted price of items in a category. This finding suggests that any price summary methodology that relies on average posted prices will overstate the price level that is experienced by consumers. We show that the geometric mean methodology adopted by the BLS for CPI calculation in 1999 also overstates the average price paid by consumers.

Second, we introduce the concept of a "best price", defined as the lowest price charged for any good in the narrow product category during a short multi-week time window. The model predicts that "best prices" should be the relevant prices for discount-chasing consumers. We show that the actual price paid tracks the "best price"; the average price paid is well-approximated by an appropriately weighted average of the best price and the fixed weight price index. Thus, the data match the structural form of our model.

Third, we examine the implication of our results for price index construction. While the price paid (average price constructed with variable weights, or, unit values) would appropriately summarize the utility implications of prices generated by our model, a real time variable weight index is impractical for a statistical agency to construct. We demonstrate that our structural model implies a simple methodology to approximate price paid without the need for high frequency quantity data.

Our second type of analysis uses a wider set of products and a nationally representative sample of prices to investigate the model's implications for inflation measurement. The store-level results imply that the level of prices as conventionally measured will be overstated, but need not imply that inflation estimates will be incorrect. The accuracy of the inflation estimates will instead depend on whether bargain-hunting behavior by consumers and price discrimination strategies by retailers are stable. Previous research, some described above already and more that is reviewed below, gives many reasons to doubt that these choices would be stable. But, ultimately this is an empirical question that is best assessed by directly measuring inflation.

We demonstrate that inflation for the prices paid by consumers between 2001 and

2011 differs systematically from the inflation rate that is implied by the BLS methodology for estimating inflation. Importantly, we find that inflation in unit values (the prices actually paid by consumers) is well-approximated by a combination of inflation as measured by a fixed weight index and best price inflation. Moreover, if we depart from the prediction of our model and replace the fixed weight inflation index with inflation measured by a BLS-style geometric mean index, inflation in best prices continues to be significant. So the cyclical properties of inflation that is actually experienced by consumers depends on how best prices are changing.

These results are closely related to findings by Handbury et al. (2013). They examine biases in the official Japanese price index relative to a Tornqvist index constructed using scanner data. They find that the error in the official index relative to the Tornqvist is substantial and that is nonlinear in the level of inflation. The inflation index overstates Tornqvist inflation when inflation is low and more closely approximates inflation when inflation is high. They determine that the problem stems largely from formula bias at lowest level of aggregation, which is also our area of focus. While the BLS's methodology of using a geometric mean index reduces the measured formula bias somewhat relative to the Japanese practice of using the simple average of prices (the Dutot index), the BLS methodology still results in a nonlinear bias relative to the Tornqvist. Roughly speaking, this suggests that, while the BLS methodology allows for more substitution at the lowest item level relative to the Dutot, the actual quantity data implies different substitution patterns than the geometric mean formula allows.

The research by Handbury et al. (2013) leads back to the conundrum that our paper is attempting to address. Statistical agencies cannot (or at least have not) made use of high frequency quantity data. This makes direct construction of any variable weight index (such as the Tornqvist) impossible. However, in theory, the adoption of a model-driven weighting scheme that approximated substitution patterns better than the Dutot or the geometric mean methodology would improve the accuracy of the price index. Our paper proposes such a weighting scheme to accommodate the specific (and important) issue of temporary discounts.

Our paper proceeds as follows. Section 2 presents our simple model of a price-discriminating retailer facing two types of customers. We derive testable empirical implications of the model. Section 3 describes the data. Section 4 establishes a number of new facts about pricing and purchase patterns that are consistent with the model. Section 5 concludes.

II. A Model of Price Discriminating Retailers and Heterogeneous Consumers

We begin by presenting a simple model that is similar in spirit to Varian (1980), Salop and Stiglitz (1982), Sobel (1984), and Pesendorfer (2002). The model takes consumer heterogeneity as its primitive. Consumers are heterogeneous in their brand loyalty and in their willingness to engage in product storage. The firm knows about this heterogeneity and accounts for it in price setting. We show conditions under which, even in a stable demand environment, intermittent price discounts arise as an optimal strategy. We also show that, although the firm bears no costs of any kind in of changing prices, when consumer preferences are stable, the firm will iterate between a small number of prices, even in the face of changing marginal costs and some types of demand changes. The "regular" price will change infrequently but temporary discounts will be utilized to price discriminate. Our goal is to use this model to characterize the appropriate price index in this environment.

A. Model Assumptions

Consider a retailer selling two substitute differentiated products, A and B. We will focus on a single retailer for simplicity. However, we note that it would be fairly straightforward to embed our model into a model of two retailers competing in geographic space. In such a model (see, for example, Lal and Matutes 1994, Pesendorfer 2002, Hosken and Reiffen 2007), consumer reservation prices would be determined by the price that would trigger consumer travel to another store. Thus, for tractability, we focus on a single retailer, but a monopoly assumption is not necessary. We discuss the competitive case as one of the possible extensions after we have analyzed the basic set up.

Assume that all customers have unit demand in each period but are differentiated in their preferences for the two substitute goods. A fraction $\alpha/2$ of the customers value product A at V^H and product B at V^L, where $V^H > V^L$. We call these consumers the Loyal A types. For convenience, we consider the symmetric case where a fraction $\alpha/2$ of the customers, the Loyal B types, value product B at V^H and product A at V^L. The remaining share of consumers $(1-\alpha)$, the Bargain Hunters, value both products at V^M=$(V^L$+$V^H)/2$. That is, the two groups have the same mean willingness to pay for the substitute products, but one group has a preference between the two products and the other is willing to substitute between them freely. We normalize the total number of consumers to be 1 and suppose that the consumers can shop for N periods. Thus, in each period, the Loyal i types have utility $U^H = \max(0,\ V^H\text{-}P^i,\ V^L\text{-}P^{-i})$. The Bargain Hunter's utility from consuming is $U^B = \max(0, \max_i(V^M\text{-}P^i))$.

We allow for the possibility that the Bargain Hunters may choose to strategically engage in storage. So their total utility consists of the utility from consuming minus any disutility associated with storage. The Loyal types do not strategically shop and thus, do not engage in storage. Thus, we are assuming that the Loyal types are inactive shoppers — they do not wait for and/or stock up during bargains, while the Bargain Hunter types do. In this sense, our model reflects well the empirical facts described in Aguiar and Hurst (2007); they document that some consumers in a local area pay systematically lower prices for the same goods than other consumers. That is, some consumers are strategic in Bargain-Hunting, and others are not.[2]

The seller has a constant returns to scale technology of producing A and B and the marginal cost of producing either is c.

Consider the shopping problem for the Loyals. Each period, they arrive at the retailer to shop. If the price of either good is less than or equal to their reservation value, they purchase one unit of the good that delivers more surplus — but, they never engage in storage. If neither good is priced low enough, the Loyals will not purchase, returning the next period with unit demand.[3]

In contrast, the Bargain Hunters engage in storage whenever it pays to do so. To build intuition about the nature of pricing problem we begin with an assumption that all the consumers have perfect foresight regarding future prices; we will eventually relax this to assume instead that they merely form rational expectations about future prices. We assume that Bargain Hunters experience a disutility from storage of $\delta(k)$, where k is the number of periods over which units are stored (equivalently, the number of units stored in this unit demand model). We assume that $\delta(k)$ is nondecreasing and weakly convex in k. This assumption is motivated by the observation that most consumers have a fixed amount of storage space to devote to product inventories.

Following Salop and Stiglitz (1982), we assume that consumers will only buy units for storage if their net utility from doing so is positive. Thus, for example, if a Bargain Hunter enters the second to last period with no inventory, and expects

[2] As in Pesendorfer (2002), we combine Bargain-Hunting behavior with low brand preference. We could provide a more detailed model with more types—brand Loyals who are willing to intertemporally substitute purchases, brand Loyals who do not intertemporally substitute, non-Loyals who are willing to intertemporally substitute and non-Loyals who do not intertemporally substitute. Most of the interesting implications are evident with these 4 types collapsed into the two extremes.

[3] In our model, as long as costs are less than V^H, the firm will always set price so that the high types purchase.

P^A=P^B=V^H in the final period, then the Bargain Hunter will purchase two units if $P^A < V^M - \delta(1)$ or $P^B < V^M - \delta(1)$, but only one unit if $P^A = V^M$ or $P^B = V^M$. That is, the Bargain Hunter will purchase one unit if the price of at least one good is low enough to induce a one period purchase, but not low enough to induce storage. Finally, if the Bargain Hunter enters period t with no inventory, and anticipates that prices will remain high until period $t+k$, the Bargain Hunter will buy k units if the price of at least one good is less than or equal to V^M - $\delta(k)$. If the price posted is low enough to induce storing for k periods, then the net utility from buying k units is (at least weakly) higher than buying any fewer than k units.

Total profits for the retailer over a horizon of N periods depend on the total amount of A and B sold and the prices charged. The retailer has three basic choices: (i) the retailer can always charge high prices and service only the Loyals, foregoing any potential margins to be earned on the Bargain Hunters; (ii) the retailer can charge a low price for at least one good each period and serve both types of customers, thus foregoing the extra willingness to pay that could have been extracted from the Loyals; (iii) the retailer can strategically iterate between high and low prices in an attempt to capture the potential demand from the Bargain Hunters while exploiting some of the extra willingness to pay of the Loyals.

In our analysis, we adopt the following definition of an equilibrium.

Equilibrium Definition: An equilibrium consists of a sequence of prices for both goods A and B from period 1 onwards announced at date 1 by the retailer and demand functions for both type of consumers such that:

1) *The consumers' demand functions maximize their expected utility taking the prices as given;*

2) *The retailer's profit is maximized at announced prices taking the consumers' demand functions as given;*

3) *The retailer doesn't want to deviate from the announced prices at any later date.*

B. Model Results - Retailer Behavior

Our goal is to provide intuition for when one would expect to see retailers offering periodic discounts and then to help frame a discussion for how to construct price indices in this kind of environment. Depending on the parameters, strategy (i), (ii), or (iii), described above can be optimal. We will pay particular attention to parameter values under which (iii) is optimal, since the pricing behavior associated with (iii) is roughly consistent with our empirical observation of occasional sales at supermarkets. We characterize the retailer's behavior in several steps.

Proposition 1: For V^H-V^L large enough, it is never optimal for the retailer to charge less than V^H for both A and B in the same period.

Sketch of Proof: The only reason not to charge V^H for both goods in all periods is to induce the Bargain Hunters to purchase. Consider a period in which the seller finds it optimal to charge a price below V^H for at least one good. In order for this to induce the Bargain Hunters to purchase, this price must be less than or equal to V^M. If the Loyals value their preferred good sufficiently more than their less-preferred good, charging a low price for the second good in the same period leads to a loss of margins on the Loyals who prefer the second good, but produces no offsetting demand increase for the Bargain Hunters (who have already been induced to purchase by the discount on the first good). Note, however, that if V^H-V^L is small enough, there is an incentive compatibility constraint for the Loyals that may bind. Specifically, the Loyal i types will buy their "preferred" good i as long as V^H-$P^i > V^L$-P^{-i}, or equivalently, V^H-$V^L > P^i$-P^{-i}. If V^H-V^L is small enough such that this incentive compatibility constraint for the wrong-type Loyal binds, it may be optimal to provide a large discount on one good and a small discount on the other. In the examples below, we return to the conditions under which the incentive compatibility constraint does not bind.

Proposition 2: A sufficient condition for the retailer to charge V^H for both goods every period and not serve the Bargain Hunter type at all is:

$$\frac{\alpha}{2}(V^H - c) - \frac{\alpha}{2}(V^M - c) > N(1-\alpha)(V^M - c) \tag{1}$$

Sketch of proof: Consider the situation in which the storage cost is zero so that the retailer can service the Bargain Hunters with a single discounted price or sale in the first period. This makes the Bargain Hunters maximally profitable to serve because the discount needed to induce them to purchase is as low as possible. The retailer will not hold any sales if the incremental revenues earned from the Loyals from forgoing a sale (the left hand side of (1)) is greater than the incremental revenues that would be earned from the Bargain Hunters by holding a sale (the right hand side).

This limiting case is helpful for building intuition about the efficacy of discounting. From (1), notice that it is optimal not to serve the Bargain Hunters if they are not numerous (if α is big relative to $1-\alpha$). They will also be ignored if the required discount is large (V^M is small relative to V^H and c). With exactly zero storage costs, it is more desirable to ignore the Bargain Hunters if the number of periods is smaller. More generally, if the Bargain Hunters could store the good for a long time without having to be compensated for storage costs, the retailer can earn the margins of serving the Bargain Hunters without providing many periods of windfall discounts to the Loyals.

Now consider the conditions determining when it will be profitable for the retailer to use discounts to serve Bargain Hunters.

Proposition 3: It is profitable for the retailer to service the Bargain Hunters by inducing them to consume in every period if the condition below is satisfied:

$$\frac{\alpha}{2}(V^H - c) + (1 - \frac{\alpha}{2})(V^M - c) > \alpha(V^H - c) \tag{2}$$

Sketch of proof: In order to prove this proposition, we need to show that, when (2) holds, for any pricing plan that leaves the Bargain Hunters not consuming in some period, there exists an alternative plan that makes the retailer strictly better off. Note that, if under a plan, Bargain Hunters do not have units in storage and do not purchase at date t, keeping the pricing plan the same before t and changing the pricing plan from t onwards will not have any effect on Bargain Hunter consumption and storage up to that point. That is, if they entered a period t in which a discount was not expected with no storage, it must be because buying a unit in some prior period and storing the good until period t would have given them negative utility. This argument relies on the fact that Bargain Hunters are always better off when they choose to consume instead of forgoing consumption. So assessing a deviation requires only that we check that the retailer gains and that the Bargain Hunter's reservation price is met.

We prove this proposition by contradiction. Suppose that under the announced plan, there is no discount expected in period N, the final period, yet Bargain Hunters' enter period N with no units in storage. If (2) holds, the retailer will find it in his interest to change his plan for period N and offer a sale of $P^A = V^H$, $P^B = V^M$ or $P^A = V^M$, $P^B = V^H$. As discussed above, such a change in plans will not change the Bargain Hunters consumption plan in periods prior to N. However, offering such a sale in period N increases retailer's profit for period N (reflected by the left hand side of (2)) relative to the profit without the sale (given by the right hand side of (2)).

We can use this intuition to construct a proof using backward induction. Next find the period before N and closest to N in which Bargain Hunters do not consume and call this period E. The same intuition applies again and suggests an alternative pricing plan that increases retailer's profits. Specifically, keep the pricing scheme under the new plan the same for periods 1 to E-1 and set the pricing scheme for E to N-1 the same as the pricing scheme from E+1 to N under the old plan. The new plan leaves Bargain Hunters' purchases and consumption unchanged in periods 1 to E-1 and only changes Bargain Hunters' purchases for periods E to N. This shift leaves two possibilities for period N: either the Bargain Hunters enter period N with no storage or the Bargain Hunters enter period N with a unit in storage.

Consider first the possibility that under the new plan Bargain Hunters' consumption in the last period is zero, then the actions of the Bargain Hunters in periods E to N-1 with this new plan are the same as their actions in periods E+1 to N. Also offering a sale in period N will not change the Bargain Hunters' action in periods E to N-1 for the reasons given above. Therefore the retailer's profits in periods E to N-1 are the same as his profits from E+1 to N with the old plan. But, the retailer's profit with the new plan in period N strictly dominates his profit in period E with the old plan. Hence, the retailer has increased his total profits in periods E to N with this new plan.

Next consider the alternative case in which Bargain Hunters consume in the last period under the new plan. In this case, with the same number of sale periods from E to N the retailer has sold an extra unit to the Bargain Hunters and hence has made positive profits on that extra unit.

Working backward this way towards the first date and changing the retailer's pricing plan as proposed means that the retailer will make sure that the Bargain Hunters consume at least one unit each period. Since Bargain Hunters can consume at most one unit per period, the retailer's optimal plan must induce the Bargain Hunters to consume exactly one unit every period.

Proposition 4: If Bargain Hunters enter period N-1 with no inventory, and the condition in Proposition 3 holds, it is more profitable for the retailer to induce two-period storage in period N-1 than it is to hold sales in both period N-1 and N if:

$$\frac{\alpha}{2}(V^H - V^M) > (2 - \frac{3\alpha}{2})\delta(1) \tag{3}$$

Sketch of proof: Total profits in period N-1 and period N are the sum of the Bargain Hunter and Loyal profits in each period. If a discount is offered each period, the retailer can offer a smaller discount because the price does not have to be low enough to induce storage. However, half of the Loyals are able to buy at the sale price each period. Inducing storage is profitable if the losses associated with the discount needed to induce storage (on the right side of (3)) is less than the extra revenues earned from the Loyals by being able to charge V^H for both goods in period N rather than the price that would induce the Bargain Hunters to buy in the period N (on the left hand side of (3)).[4]

The analysis so far had assumed that the consumers knew the price path in advance. But we described things this way for simplicity. A rational expectations equilibrium will obtain in this environment.

Proposition 5: If (2) holds, the retailer's optimal plan is time consistent; that is the retailer finds it optimal to announce a plan at period 1 and stick to it until the end without surprising the Bargain Hunters.

Sketch of Proof: If (2) holds the best thing the retailer can do is to make sure that the Bargain Hunters consume one unit each period. Therefore, not holding a sale in a period that the Bargain Hunters entered with nothing in storage because they expected a sale is never optimal. Also, holding a sale in a period that the Bargain Hunters have goods in storage already is never optimal. Due to the storage cost, the Bargain Hunters' valuation for the next consumption unit is at least weakly smaller compared to a period when they have no units in storage and they can consume at most N units. The only circumstance in which a retailer can be made strictly better off by holding a surprise sale is if, under the announced plan, Bargain Hunters arrive in a period with no units in storage. However such a sale could have been announced from the beginning with the same exact outcome as holding a surprise sale. Therefore, the retailer can maximize his profits by announcing a plan in period 1 and sticking to it until the end; for an optimal plan, there is no gain to changing what was announced in any later dates.

Also note that, due to the weak convexity of the storage costs, a Bargain Hunter is at least indifferent to buying inframarginal units and will earn rents if storage

[4]The right hand side of (3) follows because the Bargain Hunters each buy two units, one to consume and one to store, while the same-type Loyals buy one unit in the period with the discount so that total units sold at the discounted price is $2(1-\alpha)+\alpha/2$.

costs are strictly convex. As shown above, in order to increase revenues from the Loyals, it is worth it to the retailer to leave these rents to the Bargain Hunters. Note that the Bargain Hunters are just indifferent, under this pricing scheme, between purchasing the k^{th} unit and not purchasing it.

The above propositions show that there are parameter values for which offering discounts (but not every period) is optimal. Now, consider a retailer who holds a sale every k^{th} period. If $k = 1$, the retailer holds a sale every period. If $k = N$ (as when storage costs are zero), the retailer holds a sale in the first period and satisfies all N periods of demand for the Bargain Hunters in the first period. If $k > N$, the retailer never holds sales. We will focus on interior solutions where $1 < k < N$. We consider the profits of a retailer who charges $P^A = P^B = V^H$ every period except during a "sale" and holds a sale at P=V^M-$\delta(k)$ for one or the other good every k^{th} period, including period N. The undiscounted total profits for this retailer over all N periods are (if N, k, and N/k are integers):

(4)

$$N\frac{k-1}{k}\alpha(V^H - c) + \frac{N}{k}\frac{\alpha}{2}(V^M - \delta(k) - c) + \frac{N}{k}\frac{\alpha}{2}(V^H - c) +$$

$$N(1-\alpha)(V^M - \delta(k) - c)$$

The four terms in (4) are very intuitive. The first piece represents the profits from selling to the Loyals only, which will occur during all the periods with regular prices. The second term is the profits from the Loyals during the periods where they are able to buy their preferred good at a discount. During these sale periods the other Loyals still pays V^H so that explains the third term. The last term is the profits from the Bargain Hunters. Notice as long as the sale price is low enough to induce storage, the Bargain Hunters will buy enough to consume every period, even though purchases take place only every k^{th} period. So k matters only because it governs the disutility of storage. Therefore, so as long as the conditions above for serving the Bargain Hunters hold, the seller will then choose k to maximize (4). That is, the seller will deterministically hold a sale every k^{th} period.

To explore how the model works we consider two simple functional forms for $\delta(k)$. The first is:

(5)

$$\delta(k) = \begin{cases} 0 & \text{if } k \leq k^* \\ \infty & \text{if } k > k^* \end{cases}$$

This formulation implies that Bargain Hunters have no storage costs up to a fixed capacity constraint. In this case, the optimal strategy for the retailer is to charge V^H for one good every period and to charge V^M for the other good every k^* periods as long as:

$$(1-\alpha)k^*(V^M - c) > \frac{\alpha}{2}(V^H - V^M) \tag{6}$$

This expression implies that it is optimal to have a sale every k^* periods if the incremental profits from having one sale over k^* periods to attract the Bargain Hunters is greater than the loss in revenues on the Loyals from holding the sale. Notice that here, since the retailer does not discount either good to a price below V^M, and since $V^H > V^M > V^L$, we need no additional conditions to insure that the wrong-type Loyals will not substitute to the less-preferred good to take advantage of the sale.

Another useful benchmark is linear storage costs, where $\delta(k)$=δk. Assume also that V^H - $V^L > \frac{\alpha\delta}{1-\alpha}$. Here, over N periods, a retailer chooses k to maximize:

$$N\frac{k-1}{k}\alpha(V^H - c) + \frac{N}{k}\frac{\alpha}{2}(V^M - \delta k - c) + \frac{N}{k}\frac{\alpha}{2}(V^H - c) + N\left(1-\alpha\right)\left(V^M - \delta k - c\right) \tag{7}$$

For a retailer with an interior solution for k, the optimal k is:

$$k = \frac{\sqrt{(V^H - V^L)\alpha}}{2\sqrt{(1-\alpha)\delta}} \tag{8}$$

This expression is intuitive. The time between sales, k, is increasing in the willingness to pay a premium by the Loyals, and decreasing in the share of the Bargain Hunters and the storage cost parameter δ. The assumption that that V^H - $V^L > \frac{\alpha\delta}{1-\alpha}$ is required to insure that the incentive compatibility constraint for the wrong-type Loyals is satisfied. That is, the condition insures that a customer who is Loyal to good A prefers it enough to good B to not switch to good B when B is discounted.

From the retailer's point of view the relevant "price plan" is the full sequence of high and low prices that prevail over the cycle of N periods. Note that even with unchanging cost and demand parameters, for many parameter combinations, the firm changes prices from period to period as it optimally iterates between capturing rents from Loyals and capturing the demand of the Bargain Hunters. Fixing tastes and technology in this set up, the main choice variable for the retailer is k. P^A and P^B are choices in each period but are pinned down by the willingness to pay of the consumer types.

It is useful to compare the outcomes of this model to the models proposed in Kehoe and Midrigan (2012) and Eichenbaum et al. (2011). Both of those models can predict that a firm will charge a fixed regular price and sometimes offer a discount. However, in both of these papers, the decision to discount is driven by some change in the cost or demand environment. In our model, a sale would occur every k weeks with no change in the cost or demand environment.

It is also useful to compare this model to Pesendorfer's (2002) microeconomic model of sales. In Pesendorfer (2002), the sale decision is stochastic, but shocks to cost are an important driver of the decision to hold a sale. In our model, discounts occur even if marginal cost is constant.

Another important distinction between all of these models and ours is that none of the other models explicitly examine a retailer managing a portfolio of close substitute products. Indeed, the fact that cost changes are important in these models implies that the prices for close substitute products would tend to be *positively* correlated. In our model, the time series of prices for close substitute products are *negatively* correlated (unless a common shock leads to the change in V^L and V^H).

Guimaraes and Sheedy (2011) offer a model that appears to be very similar to ours. Importantly, discounts in their model are not driven primarily by cost shocks. Their model, like Sobel (1984), examines two competitors (competing brands or competing retailers) chasing a *fixed* pool of Bargain Hunters. Their model presumes a stationary demand from the Bargain Hunters so that competition in each shopping period looks the same: there are some Bargain Hunters in the market and if one store ignores them, the other store has an incentive to serve them. Sales are thus strategic substitutes; firms do not want to have sales if their rivals have them, but do want to have sales if rivals do not. In this model, sales are not much used as a response to shocks in equilibrium because, in effect, the strategic substitutability assumption keeps the total level of sales in the economy nearly constant. This result would not hold if, for example, consumer responsiveness to sales varied over the cycle or if sales were strategic complements, as they are in, for example, Lal and Matutes (1994).

Klenow and Willis (2007) find that sale prices are at least as responsive to shocks as regular prices. The Klenow and Willis result is rationalized by our model, but not by the model in Guimaraes and Sheedy (2011). Wong and Nevo (2014) also find that shopping intensity during the 2008/9 recession varied, so that the effective pool of Bargain Hunters expanded. Kryvtsov and Vincent (2014) study data on CPI for the United Kingdom from 1996 to 2012 and find that the intensity of sales is highly counter-cyclical. For instance, they estimate that that frequency of temporary discounts roughly doubled during the Great Recession. All this evidence casts doubt on the force in the Guimaraes and Sheedy (2011) model that makes discounting insensitive to economic conditions.

It is also useful to enumerate the circumstances in our model that would lead to a change in the regular price. The regular price moves with V^H and is held constant if there are shocks to various parameters (V^L, c, and α) as well as if there are changes in the storage cost function $\delta(k)$. Thus, despite the absence of menu costs in our model, our model accords with the empirical observation in the literature that regular prices change infrequently. If any of those parameters change, the retailer's optimal response is to alter the frequency of sales, the depth of sales, or the decision of whether or not to have sales at all. Thus, if at time zero, the retailer were engaged in a predictable k period sales cycle with fixed "high" and "low" prices, shocks to these demand and cost parameters would lead to changes in the discounting policy that would change the lowest price at which consumers could achieve a fixed level of utility.

Note that, in any individual period, the quantity sold for an individual good could be as low as $\alpha/2$ and as large as $\alpha/2 + k(1 - \alpha)$, if the good is on sale and the Bargain Hunters are storing until a sale expected in k periods. Thus, the volatility in quantity across products and periods occurs despite demand and supply primitives that are constant through time. In our model, the sum total of goods A and B sold over the $k-$period cycle is constant as long as the parameters are such that the retailer chooses to serve the Bargain Hunters.

C. Model Results - Implications for price measurement

The preceding results are also helpful for the construction of price measures. As we demonstrate more carefully below, in this model, due to the assumptions made about product valuations and consumer types, if storage costs are zero or small, measurement of changes in unit values is the appropriate summary statistic for changes in consumer utility. That is, our model corresponds to the scenarios described in Diewert (1995) and Feenstra and Shapiro (2003*b*), where unit values correspond to the exact price index. This is not generally true in models in which consumers and goods are heterogeneous and some consumers have brand loyalty. Here, however, the Bargain Hunter consumers bear no utility cost of switching between products; if storage costs are zero, then each Bargain Hunter has consumer surplus equal to the difference between V^M and the price of the cheaper product. Due to the strategic behavior of the retailer, the Loyal customer never buys the "wrong" product. A Loyal consumer has consumer surplus equal to the difference between V^H and the price of the consumer's preferred product. Total consumer surplus is the sum of the consumer surplus of each type of consumer in the economy. In the equilibrium that we describe above, Bargain Hunters earn no net consumer surplus. They pay their reservation value V^M. However, the presence of Bargain Hunters exerts a positive externality on Loyal consumers. Because of the existence of Bargain Hunters, Loyals sometimes obtain a product at a price lower than their maximum willingness to pay. Storage costs represent the only wedge between the changes in consumer utility and unit values. If storage costs are large, some of the savings that consumers achieve in temporary sales are expended on storage, not captured as increases in consumer utility.

A fixed weight index measuring average price over the entire period would be:

$$\frac{1}{2k}(V^M - \delta(k)) + \frac{2k-1}{2k}V^H \tag{9}$$

This fixed weight price does not capture average price from the retailer's perspective or from the perspective of the Bargain Hunters. The actual price paid or unit values put much more weight on the low price, because unit values reflect the strategic shift of the Bargain Hunters into the low priced product (when one is on sale) and the stockpiling by Bargain Hunters.

This is easiest to see by first assuming the very simple functional form for the storage cost presented in (5) above, the case where Bargain Hunters can store goods costlessly up to a fixed capacity of k units. In this case, average revenue per unit equals:

$$\frac{\alpha + 2k(1-\alpha)}{2k}V^M + \frac{\alpha(2k-1)}{2k}V^H \tag{10}$$

Of course, we know that in general, lower prices expand demand, but in our model, that factor is magnified by the presence of multiple products that Bargain Hunters view as perfect substitutes and because Bargain Hunters stockpile demand.

An important outcome of our model is that, when storage costs are zero (or negligible), the time series of average revenue per unit (unit values), averaged over the sales cycle of k periods, is the exact price index. That is, a change in the unit value index represents a change in the cost of achieving a given level of utility over time. To see this, for example, consider the storage cost specification in (5) above and consider the change in equilibrium if a new storage technology increases the storage capacity of the Bargain Hunters, k. The seller will decrease the frequency of sales. This will increase the average price paid by the Loyal types but will leave unchanged the average price paid by the Bargain Hunters; the Bargain Hunters will simply store the good longer. The average cost over the k period sales cycle of achieving an overall weighted average level of utility of $\alpha\, V^H + (1-\alpha)\, V^M$ per period is given by (10) (the unit value). If k changes, the gross utility achieved by the two types does not change (as they change their consumption), but the cost of achieving that level of utility varies with k according to the formula in (10). Thus, under the simple storage capacity assumption, our model corresponds to the scenarios described in Diewert (1995) and Feenstra and Shapiro (2003*b*), where the unit value index is the exact price index.[5]

It will be very helpful to compare (9) and (10). Recall that (9) is the fixed weight index and that (10) is the average revenue per unit or unit value. Note also that the lowest or "best price" achieved over the k period sales cycle is V^M. The unit value in (10) can, therefore, be rearranged to be equal to:

$$\alpha\left(\frac{1}{2k}V^M + \frac{2k-1}{2k}V^H\right) + (1-\alpha)V^M \tag{11}$$

This expression shows that the unit value in (10) is equal to a weighted average of the fixed weight price index in (9) and the "best price", where the share of Loyals and of Bargain Hunters in the population are the weights. The formula shown in (11) forms the basis of our price measurement proposal. Should the patterns of consumer and retailer behavior sufficiently conform to our model, a price index can be constructed using (i) an estimate of the share of Bargain Hunters in the marketplace (ii) a conventional fixed weight index and (iii) a measure of the best price. By providing a simple functional form for product substitution, the formula does not require time-varying quantity or expenditure weights. To summarize:

[5]The idea here, of thinking of the price as the average over the entire planning horizon, is closely related to an insight first provided in Feenstra and Shapiro (2003). They advocate for the use of the Tornqvist index rather than unit values. Reinsdorf, in his discussion of the paper, notes that, under the storage condition described in (5), unit values are a reasonable alternative. Of course, the Tornqvist does not emerge as an exact index in our model. However, more generally, the primary drawback of the Tornqvist, from our perspective, is the reliance on current period weights which cannot be collected by BLS shoppers.

$$unitvalue_t = \alpha fixedweightindex_t + (1-\alpha) bestprice_t \tag{12}$$

Importantly, unit values will not be the appropriate input for an exact index for more general convex storage cost. In the general case, the weighted average price (unit value) in the case of convex storage costs is:

$$\alpha\left(\frac{1}{2k}(V^M - \delta(k)) + \frac{2k-1}{2k}V^H\right) + (1-\alpha)(V^M - \delta(k)) \tag{13}$$

Here, the fact that the price paid is lower than V^M does not imply that the price of obtaining a given level of utility for the Bargain Hunters is equivalently below V^M because the storage cost is incurred. In general, the higher the storage costs borne by the Bargain Hunters, then the greater the deviation between unit values and the exact index. However, the unit value is still equivalent to the appropriately weighted average of the fixed weight index and the best price as in (13).

Note also from (13) that, as α approaches 0, the weighted average price will approach the "best price"—the lowest price posted for any of the substitute products within the k period planning cycle. If α is fairly small and if storage costs are small and V^H remains constant for a long period of time, then a time series of the weighted average price will resemble a fixed increment over the time series of the "best price". If α is very large, then the weighted average price will more closely resemble the "regular" price. For intermediate values of α, the price paid resembles a weighted average of the "best" price and the fixed price index as illustrated in (13).

These observations also have implications for inflation measurement. If the environment is stable, then the model only predicts that unit values will be lower than standard price indices. The *change* in prices paid might be higher or lower than the change implied by any particular index, so that conventionally measured inflation would not necessarily be too low or too high. However, if there are shocks, for example to the willingness to pay of Bargain Hunters, these shocks can shift the frequency and depth of sales, which would drive a wedge between the unit values and standard price indices.

We can take the log difference of (12) to compute an approximate inflation formula. Specifically:

$$(14) \quad ln(unitvalue_t) - ln(unitvalue_{t-1}) = \\ ln(\alpha fixedweightindex_t + (1-\alpha)bestprice_t) - \\ ln(\alpha fixedweightindex_{t-1} + (1-\alpha)bestprice_{t-1})$$

Both (12) and (14) suggest estimating equations that we will explore in the empirical analysis.

D. Model Extensions

The basic model can be extended in a number of ways. Most importantly, from a measurement perspective, one could incorporate more complex demand. In particular, one could imagine maintaining the assumption that there is a mass of consumers that are indifferent between the substitute products, but build a different model of the cross-price elasticity of the Loyals. In that case, the appropriate price index would still be the weighted average of the two types, but the Loyal type could have a more complex index formulation corresponding to the Loyals' substitution patterns.

Furthermore, there are numerous ways to expand the model to allow for additional types of consumers. In particular, as mentioned above, there may be some consumers who inter-temporally substitute actively, but are highly brand loyal. There may be other consumers who are not brand loyal but do not inter-temporally substitute. In these more complex cases, the formulation and intuition in (9) is particularly helpful. Specifically, we can think of the prices paid for a set of closely related goods over a shopping cycle as being characterized by the weighted average of the prices paid by Loyal shoppers and the prices paid by Bargain Hunters. As shown above, the prices paid by the Loyal shoppers are essentially a fixed weight average of the prices posted. However, the prices paid by Bargain Hunters more closely resemble the lowest price charged for any substitute good over a reasonable shopping horizon. In a model with more types, the pricing outcomes will be as in (9) but the weights on the different prices will reflect the behavior of the different consumer types.

Also, while there is a long tradition of two-type models in the literature on temporary discounts (including Varian 1980, Sobel 1984), the model can be extended to allow for a continuum of consumers' willingness to pay. As in Varian (1980), as long as enough consumers do not move all of their purchases into sale periods, intertemporal variation in prices will enhance the retailer's ability to price discriminate.

Our model also abstracts from competition among multiple stores. Obviously, a model with perfect competition would limit the opportunity to price discriminate. However, an extension of our model along the lines of Lal and Matutes (1994) would preserve the basic insights of our model. Lal and Matutes (1994) model multiproduct retailers with travel costs between stores. The ability of consumers to shop both stores, at a cost, disciplines the overall rents across products that stores can extract from each consumer. However, the ability to extract different levels of rents from different consumer types and from different products is preserved. Ellison (2005), in a different context, provides an extension of Lal and Matutes in which price discrimination across consumer types emerges in an equilibrium with spatial competition. The easiest way to understand the effects of introducing competition into our model is to view competition as impacting consumer willingness to pay. Spatial competition opens additional scope for price discrimination as retailers may benefit from price discriminating between consumers who will and won't travel to obtain a price discount.

Our basic model provides a framework for constructing price indices. Below, we will show empirically that the unit value index is well-approximated by a weighted average of a fixed weight index and the "best price", as in (12). We use these insights to propose a methodology for measuring prices in the absence of high frequency quantity data.

Summing up, the model comfortably explains the familiar price pattern observed for individual goods of a regular price with intermittent sale prices. In addition, it makes the following testable predictions. At the store-level, a disproportionate fraction of goods are sold at temporary discounts. Second, a unit value price index should be well-approximated by a linear combination of a conventional fixed weight price index and the best available price within the group of close substitutes over the course of several weeks. Third, the geometric mean index used by the BLS will not adequately account for the migration of consumers to the "best price". Fourth, the model suggests that inflation as experienced by consumers will also depend on the inflation rate for the "best price".

III. The Data

We are interested in examining both the incidence of sales and consumer responses. We will also measure whether monthly indices using variable quantity weights can be approximated in datasets without access to quantity data. To accomplish this, our project requires a data set that contains both prices and quantities. This requirement eliminates many data sets, most notably those from the Bureau of Labor Statistics that have been the source for many of the most important recent papers on price setting patterns. We instead we rely on scanner data from IRI Symphony that includes quantity information.

Symphony IRI's "IRI Marketing Data Set" is described in Bronnenberg et al. (2008); since the time of that publication, IRI has released additional data so that it covers retailers in 50 market areas for the period 2001 to 2011. We undertake two sets of analyses. First, to explore the details of our model on micro data, we undertake a store-level analysis. For these purposes, we focus on two categories of products: ground coffee and peanut butter. We look at substitution patterns for a single large store in each of nine cities. Second, to more closely approximate BLS inflation measurement procedures, we construct national indices using a broader set of products.

For the store-level analysis, we choose the peanut butter and coffee categories for several reasons. First, the categorization of these products in the IRI data set matches closely a product that is tracked by the BLS. This allows us to mimic the BLS sampling procedures for that particular product.[6] The categories are reasonably representative of all the grocery products in the IRI sample. Bronnenberg et al. (2008) describe the 30 categories tracked by IRI and calculate the percent of total volume in the category sold "on any deal". The median category has 37 percent of its volume sold "on any deal". Bronnenberg et al report that the share sold on deal is 40.8 percent for coffee and 32.9 percent for peanut butter. In our store-level analysis of these two products, we examine results for a sample of nine cities, one chosen randomly from each of the nine Census divisions. In each city, we sample from the largest chain — though we discard any chain where private label brands dominate the national brands. The details of our rules for selecting brands and stores are discussed in the Data Appendix. Here, we summarize the main elements of our procedures.

Whenever possible, we use BLS methodologies as a guide in creating our own methodologies for examining prices. In particular, the IRI data are reported at a weekly frequency; BLS indices are constructed at a monthly frequency. For some of the store-level analyses of substitution patterns, we work with weekly data, but we also aggregate data to a monthly frequency to mimic CPI construction methodologies.

Our store-level analysis proceeds starting with weekly universal product code

[6]Reinsdorf (1999) examines coffee for a similar reason – to match the BLS data to scanner data.

(UPC) level data for coffee and peanut butter. An example of the most disaggregated definition of a good that we would consider would be a "Peter Pan 18 ounce jar of creamy peanut butter". The IRI Symphony data set, like most scanner data sets, contain a large number of products, many of which have a short lifespan. The coming and going of individual products could lead to the prices that are sampled jumping around in ways that can easily generate misleading inferences. For example, if a standard product is unavailable, because it either disappears or is temporarily unavailable, and is replaced by a super-sized version of that product or by a premium version of the product, the price will move purely because of substitution. We take a number of steps to limit this possibility.

First, we restrict attention to benchmark sizes (11 to 13 ounces for coffee and 15 to 18 ounces for peanut butter) and that are most typical in the data. Next, we identify the best selling products and focus on the top 10 UPCs in each category. In calculating market shares we inspect the data and splice together UPCs that are replacements for each other. In addition, we consolidate all UPCs within a given brand that have very high price correlation (as is typically the case for different flavors of the same product). We then retain only UPCs for which we have at least 6 years data. Finally we convert the prices into price per ounce and exclude any private label UPCs or super premium UPCs. Collectively, we believe these steps create coherent categories of similar products where the model's assumptions about substitution are reasonable and where sampling variation will not produce spurious price volatility.

For each store and UPC, our data contain weekly observations on prices and ounces sold. For some of our calculations, we require definitions for "sale prices" and "regular prices". We extend the methodology proposed by Kehoe and Midrigan (2012). Kehoe and Midrigan propose measuring a sale as a price cut which is reversed within five weeks. We adopt a similar definition. However, we note that the data contain some very small apparent price changes; there are cases where the price in a week appears to be less than a cent or two lower than the price in the previous week. As in most scanner datasets, the price series is actually constructed by dividing total revenues by total unit sales. There may be product scanning input errors or situations in which a consumer uses a cents off coupon or a store coupon, and any of these would create tiny shifts in measured prices that do not reflect changes in posted prices. We thus set a tolerance for a price change—requiring the price change to be "large" enough to be considered either a sale price or a change in the regular price. We set this tolerance at \$0.002 per ounce. A regular price is set equal to the actual price if the product is not determined to be on sale using the above methodology. If the item is on sale, the regular price is defined to be the most recent past price for the item for which the item was not on sale.

In our national index analysis, we again seek to replicate the BLS methodology as closely as is possible. Because the BLS samples on a monthly basis, we assign each week in our database to a month. As the BLS does, each month, we will sample from particular stores in a particular week.

Even for specific products that the BLS tracks, the BLS does not provide detailed data about the number of prices sampled, but approximations are possible. For example, roughly 35,354 food and beverage prices are collected monthly by the BLS (see http://www.bls.gov/cpi/cpirr2011.pdf). Peanut butter is the most important item in a stratum called "other fats and oils including peanut butter" that has about 0.78% of the total weight of food and beverages. Thus, the BLS should collect roughly 276 prices per month for the full stratum. We assume that by virtue of its dominant position, 240 of these quotes would be for peanut butter. A similar calculation for coffee leads to an approximation that the BLS would collect 224 coffee prices.

For our national index analysis, we expand our analysis to consider 21 additional products for IRI. These represent all of the categories of products tracked in the IRI data that we viewed as sufficiently homogeneous to allow us to identify a "best price" and for which sufficient continuous data were available.[7] Most of these products do not correspond exactly to a category that the BLS tracks in an identifiable way.[8] Lacking better data for exactly how many prices the BLS would sample in our other categories, we collect 240 prices per month for all categories in our national index analysis.

In order to make sure that our calculations are geographically representative, we divide the US into the four US census regions and allocate our price estimates to each region in the same proportion as they are allocated in the CPI (using data from http://www.bls.gov/cpi/cpiri2011.pdf). In both our store-level and national analyses, we use a methodology similar to the BLS to calculate and update weights for the fixed weight index. We start by identifying base volume weights in a "pre-period" of our dataset. Each quarter, we construct a new weight consisting of 15/16 of the previous weights and 1/16 of the prior quarter's weights.

[7] An example of a category that we excluded was diapers. Ounces seemed like a poor measure for diapers and using the number of diapers in a package is also misleading since diapers are sized like clothing and the products for larger-sized children tend to be more expensive.

[8] The reasons for this are varied. For example, one of the categories that we examine is frankfurters. This does not correspond exactly to a category that the BLS tracks. This is in part because the BLS categorizations classify meats by animal of origin. Thus, frankfurters contribute to the pork, beef, and turkey items, and there is no BLS item for "frankfurters".

As discussed above, we focus on only two categories in our store-level analysis and a broader set of 23 categories in our national analysis. There is always a tradeoff between comprehensiveness of categories and the care that can be paid to the details of each category. For example, Chevalier et al. (2003) document that in some categories UPCs are discontinued only to have the same product appear with a new UPC. Hence, splicing series by hand is the only sure way to capture all the same sales of these types of similar items. But, the task of splicing the data to capture UPC changes, as well as grouping UPCs within brands, requires a substantial investment in learning about the category and cleaning data. Of course, this could in principle be done on a large scale, but it is very costly. The Bureau of Labor Statistics devotes some portion of its field force to cleaning data to create consistency across categories in construction the CPI (see, for example, Bureau of Labor Statistics, 2011.) For our peanut butter and coffee categories for the store-level analysis, we undertake this type of splicing. In our national analysis of 23 categories, we use a series of rules to select the sizes and products that are included. Details can be found in the data appendix.

IV. Results - Store-level data

Before turning to the predictions of our model, we document that high frequency price variation is quite important for understanding the actual prices that consumers pay. Figures 1 and 2 show, for the 2003-2005 period in the Charlotte store, the monthly fixed weight purchase index for peanut butter and coffee, respectively, along with a fixed weight index of regular prices, the unit value, and the best price. The graphs show that both products have a substantial period of time when the regular price is nearly constant. Movements in the fixed weight index clearly reflect the occasional temporary discounts (captured by movements in the best price series). However, as expected, unit values embody temporary discounts much more than the fixed weight index would imply.

Next, we turn to examining the extent to which our model adequately explains the pattern of temporary discounts and purchases. Recall, we identified three store-level predictions of our model. First, a disproportionate fraction of goods should be sold while the price is temporarily low. Second, unit values should be well-approximated by a linear combination of a conventional fixed weight price index and the best available price within the group of close substitutes over the course of several weeks. Third, the geometric mean index used by the BLS should not adequately account for the migration of consumers to the “best price”.

A. *Purchase Responses to Sales*

We first examine the effect of sales on prices paid by consumers and quantities purchased. Table 1 shows total weeks and ounces with temporary discounts for each of the nine cities. It also shows, for weeks in which there is a discount, the mean size of the discount in each city. It is unsurprising that quantity sold increases substantially when a product experiences a price reduction. However, the combination of frequent, staggered discounts along with consumers who readily switch brands and time purchases means that a substantial fraction of all of the units sold are sold at prices below the "regular" price.[9] In all of the cities, the fraction of ounces sold at discount is substantially greater than the share of item-weeks with discounts, suggesting ubiquitous strategic shopping behavior by consumers for both products in all locations. For peanut butter, on average, the fraction of units bought on sale is roughly twice the fraction of item-weeks with sales. For coffee, responsiveness is lower, with the fraction of units sold on sale being roughly 1.3 times the fraction of item-weeks on sale. Under the strong assumption that sale weeks are chosen randomly, the ratio of the quantity sold in sale vs. non sale weeks can be compared to the mean size of discounts during sales to calculate a rough estimate of the elasticity of demand. These calculations suggest an elasticity of demand of roughly 4 for coffee and roughly 6 for peanut butter.

There are a number of reasons why these patterns for coffee and peanut butter may differ. For example, the demographic characteristics of peanut butter buyers may differ from that of coffee buyers or, coffee companies may be more successful in differentiating themselves in a way that creates brand loyalty, and people may be more willing to store peanut butter. The greatest potential impact of sales on price measurement occurs in Hartford for peanut butter, where, over our eleven year sample, only 8 percent of product-weeks have sales, but 50 percent of the total ounces sold are purchased at a sale price.

Table 2 shows brand-level detail for the Charlotte store. The data illustrate that the pattern of sales and regular prices can differ significantly across brands. For the Charlotte store, discounts are slightly less frequent for Skippy than either for

[9] The large share of transactions that take place at sales prices is not surprising. Kehoe and Midrigan (2012) mention this finding as one of their observations about the Dominick's data that is distributed by the University of Chicago. Bronnenberg, Kruger and Mela (2008)'s IRI data set covers 30 categories of goods at grocery chains in 47 different geographical markets (that has since expanded to 50 markets). Their Table 2 shows the fraction of products that are sold on any deal and the mean percentage is 36.8%; more than 30% are sold on deal in 25 of the 30 categories they study. Griffith et al. (2009) also find that about 29.5% of total food expenditure from a large sample of British households is on sale items. Hence, the findings for our categories are very typical of what happens in grocery stores.

Peter Pan or Jif, although the share of units bought on discount is substantial for all three. The deal propensities for coffee are much different, with Maxwell House being discounted about four times as often as JFG. The fact that JFG sells so much of its product at regular prices despite the omnipresent discounts by Maxwell House suggests that there are numerous JFG-loyal customers.

B. Best Prices

We next turn to the most important prediction of our model, that the variable weight (unit value) index can be well-approximated as an appropriately weighted average of the fixed weight index and a measure of the "best price", the lowest price obtainable within a narrow item category (or stratum, to use the BLS language). In order to measure our model empirically, we do need to operationalize some theoretical constructs of our model. First, we must make an assumption about the horizon over which Bargain Hunters can be expected to look for sales and stockpile. Given that the BLS publishes monthly data, we will assume that shoppers operate on a monthly cycle. Second, we must define the universe of items within the store over which to seek the best price. Here, we use all products that meet the popularity and availability screens described above.

We conduct a simple regression of the variable weight price index ("price paid") on the fixed weight index and the "best price" series. Note that if (12) is a good approximation for our data, the weights on the two price series should add up to one and the constant term should be roughly zero. These conditions are *not* hard-wired to hold. Consumer preferences and, therefore, market shares could drift away from the fixed weights of the fixed weight indices; or consumers' demand might be very different than supposed by our model that emphasizes a prominent role for best prices.

Results are shown in Table 3 for both products in each of the cities. The first important finding is that the best price coefficient is positive and significant in all cases. That is, the prices actually paid by consumers (unit values), are not well-approximated by the fixed weight index and instead controlling for the best price adds explanatory power. This is not surprising given our findings in Table 1.

For most of the product/city combinations, the constant term appears small and the fixed weight and best price coefficients roughly sum to one. In about half the cases for each product, the estimated value of the constant term is not significantly different from zero at standard confidence levels. In about half the cases for each product we also cannot reject that the hypothesis that the coefficients for best price and the fixed weight index sum to one. In all cases, the model fits well; the minimum R-squared among the 18 product-city specifications is 0.86.

The coefficient estimates on the best price vary between 0.2 and 0.7, suggesting that bargain hunting tendencies differ noticeably. Interestingly there is a strong positive (0.80) correlation between the best price coefficients for the two products. To the extent that demographic characteristics of the store customers are driving the importance of the best price, this correlation is reassuring.

An important question to consider is whether the BLS adequately accounts for sale-chasing behavior. The BLS constructs a fixed expenditure share geometric mean index within item strata. This methodology, therefore, allows for a limited amount of cross-item substitution. However, this substitution differs substantially from that implied by our model. The BLS methodology effectively assumes a cross-price elasticity of demand of 1 between items within the strata. Whether or not this elasticity is sufficient to capture shopping patterns is an empirical issue, as is question of whether ordinal prices as stressed by our model matter.

To assess the importance of these alternative perspectives, Table 4, shows an alternative set of regression results in which a BLS-type geometric mean index and our Best Price indicator are both allowed to account for movements in the average price paid. In all cases, the best price measure still has significant explanatory power for the average price paid. Indeed, the coefficients for best price when using the geometric weight index in the regression are nearly identical to the coefficients for best price when using the fixed weight index in the regression. This suggests that the BLS methodology does not sufficiently account for deal-chasing behavior on the part of strategic shoppers.

The Best Price coefficient may be positive and significant in these regressions that include the geometric mean index for either of two reasons. First, the substitution elasticity of one that is implicit in the geometric mean index may not be elastic enough to capture the substitution patterns in the data. Second, the CES may not be the correct functional form to capture the substitution patterns in our data. The CES implies quite different consumer behavior than our model. First, whereas the CES implies that cardinal price differences drive decisions, our model suggests that the ordinally lowest price in the category should play an important role in driving purchasing. Second, the CES implies a love of variety that is at odds with the assumptions in our model. The CES implies that if the prices for branded varieties were equal, consumers would want to consumer a little bit of each variety. In our model consumers have clear preferred brands.

We investigate the possibility that a CES price index would perform better if it had a higher elasticity of substitution. We address this issue by considering alternative CES price indices. Following Shapiro and Wilcox (1997), we conduct a grid search to find the CES price index that best matches the unit value index.[10] For each of the city-products, we search over CES price indices with elasticity of substitution parameters from 0.5 to 10 in increments of 0.5 in order to find the CES price index for that city-product that most closely approximates the variable weight (unit value) index. The results are found in Table 5. Not surprisingly, for all city product pairs, the best matching CES index has an elasticity of substitution that is higher than one. For peanut butter, the preferred specification yielded an elasticity estimate greater than 4.5 in each city. For coffee, the best fit CES index has an elasticity ranging from 2 to 10. These findings accord with Ivancic et al. (2009), who suggest (in matching to the Tornqvist) that estimates of elasticities of substitution in the lowest item strata for food products are typically greater than 1 and are often greater than 3. Consistent with the patterns in Tables 1 and 2, coffee is on average less elastic than peanut butter, and there is a positive correlation across the two products within cities (the correlation of the best fitting elasticity of substitution for coffee and the peanut butter within cities is 0.7).

Having identified the best fitting CES price index for each city-product pair, we then examine whether our best prices construct maintains explanatory power for predicting unit values. Specifically, for each city-product pair, we regress the unit value on the Best Price and the optimized CES Price Index. Thus, the regression specification is identical to Table 4 but replaces the geometric mean index with the best-fitting CES index. For all of the city-product pairs except coffee in Chicago, the coefficient on the Best Price measure is positive and statistically different from zero at at least the ten percent confidence level. This suggests that this ordinal price metric is typically required to match the substitution patterns in the variable weight index. In the one exception for the sixteen city-product pairs, coffee in Chicago, the variable weight index is closely approximated by a CES price index with an elasticity of substitution equal to seven.[11]

As a final point of comparison, in unreported results, we also compute Tornqvist indices for peanut butter and coffee in each of our cities. In our model, a Tornqvist does not approximate an exact index due to consumer heterogeneity. However, in many models with homogeneous consumers, a Tornqvist index would approximate

[10] In their study, Shapiro and Wilcox match CES price indices to the Tornqvist.

[11] It is worth noting that, for the unit price index for coffee in Chicago, a regression on the CES index and a constant yields an R-squared of 0.98. A regression of the variable weight index on the fixed weight index and the best price yields an R-squared of 0.96. Thus, although the best fit CES in that one case does not require the additional information contained in the best price, our best price methodology still yields a very good approximation of the variable weight index.

an exact index and hence is another potential benchmark. For these cities and products, a regression where the Tornqvist index is the dependent variable and the geometric mean and the best price are the controls, uncovers a the best price coefficient that is consistently positive and statistically different from zero. This suggests that the BLS's preferred geometric mean index would better approximate the Tornqvist if the best price measure were also considered.

C. Inflation Measurement

Finally we assess what the patterns that we have documented will imply for the measurement of inflation. To do this, we first use our scanner data to broaden our sample of products and to aggregate prices nationally. We use all categories available from our data vendor where the product size variation is small, products are fairly homogenous, and sufficient data exists in our database. This results in twenty one additional grocery categories plus coffee and peanut butter. It is certainly clear, from our analysis using the city-level data discussed above, that the *level* of aggregated prices will differ depending on whether prices are aggregated using time-varying expenditure weights (unit values), or with fixed arithmetic weights (as are used for price index construction in many countries including Japan), or with geometric weights (as is used by the BLS in the United States).

Nonetheless, our price level results do not necessarily have important implications for tracking inflation in unit values. If the mismeasurement in levels is constant over time, inflation may be well-measured with BLS procedures. Previous research, however, suggests several reasons to expect meaningful time variation in the frequency or depth of sales. Given the strong purchase response to sales, this would lead to time variation in unit value inflation and best price inflation that may not be fully captured in a BLS-style geometric mean inflation index. This time variation in the frequency or depth of sales could occur either due to cyclical factors as in Kryvtsov and Vincent (2014) or due to long run shifts in selling technologies as in Basker (2013) and Nakamura (1998), or due to consumer substitution patterns differing in high and low inflation environments as in Handbury et al. (2013). Indeed, Nakamura and Steinsson (2008) document that the fraction of BLS price quotes that are sales increased dramatically in the processed food, unprocessed food, household furnishings, and apparel categories over the 1988 to 2005 period. For the processed food and apparel categories, sales frequencies have doubled. Furthermore, the average size of sales has also increased substantially over the period in three of the four aforementioned categories. This continues a trend documented by Pashigian (1988) in the frequency and depth of sales that begins in the 1960s. If any of these forces are at work, inflation in unit values may be better approximated by a methodology such as ours that measures discounts explicitly, rather than a geometric mean methodology.

To construct national indices, we draw prices from our national scanner dataset, approximating the sampling methodology used by the BLS. As described above, we take 240 price quotes each month for each product category. The procedure for sampling stores is designed to mimic the BLS procedure as closely as is possible with our data, and is described in the Data Appendix. As is done for the CPI, the week of the month in which a price is collected in each store is randomized. As we did for coffee and peanut butter, we focus on the most popular sized products. We identify a set of candidate UPCs in each store based on historical expenditure weights, and then sample a particular UPC in each store using these weights. The Data Appendix describes the way we approximate the BLS procedures for rotating new stores into the sample.

We also construct best price indices. Our model suggests that because of the option to stockpile, the best price in a given week is a conservative measure of the price savings available to a determined bargain hunter. However, the BLS only collects prices in a given store in one week. So, the construction of a multi-week best price would require additional enumerator store trips and thus would be extremely expensive for the BLS relative to the current methodology. In order to create a practical proposal, we depart from the model's prediction that emphasizes a multi-week best price and collect instead the best price that is available during the week when the quote for the regular index would be gathered. The one week best price is selected from the candidate UPCs that were identified in that store as being potentially eligible for inclusion in the regular index.

Finally, we compute a national unit value index. We use this as our benchmark and assess the extent it could be approximated by the other indices, as it utilizes quantity data unavailable to the BLS. To compute this, we sample stores and weeks identically to the method used to calculate the fixed weight, geometric mean and best price indices. For each sampled store, for the entire set of candidate UPCs from which we sampled for the price indices, we measure total expenditures for the week and divide by total volume in ounces.

To examine all 23 categories more systematically, we calculate twelve-month inflation rates as the log change in unit values. Recall that, in our model, the unit value is identically equal to a weighted average of the fixed weight index and the best price. To examine the fit of our model, we estimate a specification based on (14) above. Specifically, we estimate the following equation using nonlinear least squares:

$$(15)\quad ln(unitvalue_t) - ln(unitvalue_{t-1}) = \gamma + ln(\alpha fixedweightindex_t + (1-\alpha)bestprice_t) - ln(\alpha fixedweightindex_{t-1} + (1-\alpha)bestprice_{t-1}) + \epsilon_t$$

where ϵ is a postulated residual. The coefficient α is not constrained in our estimation procedure. If our model is correctly specified, our estimate of α should be between zero and one and can be interpreted as the share of Loyal consumers in the marketplace. The coefficient for the constant term, γ should be zero and the explanatory power of the specification should be high. The results for all 23 products are given in Table 6. The alpha coefficients range from 0.2 to 0.7 and average 0.4. Thus, the implied coefficient on the best price index averages 0.6. The coefficient for the fixed weight index is, in all cases, both statistically significantly different from zero and statistically significantly different from one. That is, $1-\alpha$, the best price coefficient is significantly different from zero in all cases. The constant terms are in general quite small and average 0.0002, although they are statistically significantly negative in some cases and positive in others. Finally, the overall explanatory power of the specifications is high. The minimum R-squared value is 0.53, but more than half are greater than 0.85.[12]

While the results in Table 6 suggest that our model performs fairly well and that the best price is an important component to approximating unit value inflation in combination with the fixed weight, it is possible that the BLS-style geometric mean index approximates better unit values. To examine this, we repeat the specification in (15), replacing the fixed weight index with the geometric mean index. The results are presented in Table 7 for all 23 products. As was the case in the store-level levels specification for coffee and peanut butter, the results for the geometric mean specifications are actually quite similar to the fixed weight specifications. While the differences in the explanatory power of the two specifications are similar, it is interesting to note that, for 22 of the 23 products, the fixed weight specification has a lower mean squared error and higher R-squared than does the corresponding geometric mean specification. Overall, in approximating the unit value, the inclu-

[12]We obtain similar results with alternative specifications. For example, we also examine Newey-West linear regressions of the log difference in unit values on the log differences in both best price and fixed weight indices and find results consistent with those reported here.

sion of a best price measure offers additional explanatory power over the geometric mean type index used by the BLS.

As discussed above, others have advocated the use of either unit values or the Tornqvist as an alternative to the BLS's geometric mean formulation at the lowest item stratum. Clearly, such proposals have been thus far unsuccessful due in part to difficulties in obtaining real time quantity data. Our proposed demand formulation fits the data fairly well and has the advantage of requiring only a second price to be obtained at each visited store from the enumerator. Thus, it is relatively simple to construct.

V. Conclusion

We provide a simple model of consumer heterogeneity and show how that heterogeneity can motivate temporary price discounts by retailers. Margins that vary dramatically over time, even when consumer preferences are stable, are a natural outcome of our model. By the very nature of second degree price discrimination, the seller optimally sets prices to encourage deal-driven brand switching only by those consumers who have the least disutility from switching brands. Indeed, in our stark model, only those with no disutility from switching brands actually switch them. Our empirical findings suggest that many consumers have a low disutility of switching brands. We show that the share of all goods in our sample that are sold at sale prices is much larger than the number of product-weeks in which sales occur.

We demonstrate that, from our model, the unit value emerges as the exact index; changes in the unit value index are associated with changes in consumer utility as long as storage costs are negligible. However, it is well understood that statistical agencies have largely been reluctant or unable to adopt price index methodologies that involve gathering real-time quantity data as is required to create a unit value for all goods purchased or a Tornqvist index. We show that our model can be used as a structural model of prices paid. In particular, even without high frequency quantity data, unit values can be approximated as a weighted average of the fixed weight index and the "best price". If approximating the unit value index is the goal, the best price is a crucial input.

This raises two questions that we begin to address here. First, what weights should a statistical agency use if attempting to construct a weighted average of the best price and the fixed weight index? Our parameter estimates vary by city and product. Thus, we suggest that weights could be approximated using short time series of quantity data not obtained in real time, essentially as we have done.

Perhaps one criticism of our approach is that approximating the weights in the way we have proposed would be difficult or unacceptable and that it would be difficult to implement a weighting system that varies across products and cities. However, the limitation that the "best fit" model would vary across cities and products is, in fact, not unique to our formulation. In our store-level section, we demonstrated that the "best fit" CES demand system also varied across cities and products. This implies that the geometric mean formulation used by the BLS is a much better approximation for some products than for others. Time and cost limitations could determine the extent to which a "one size fits all" best price parameter could be used, just as the geometric mean methodology employs a "one size fits all" CES demand parameter of one.

Second, is this approach truly practical for a statistical agency? We concede that obtaining real-time quantity data has been deemed impractical, and obtaining some measure of the full distribution of prices over time (to say, estimate quantities based on a CES or discrete choice demand model) may be impractical. However, our approach only requires the enumerator to record a second price on each trip to the store. This strikes us as not very costly, while providing significantly more information about a price which is relevant to a very large number of consumers.

Finally, our evidence suggests that the disconnect between unit values and BLS-style indices remains when inflation is computed. Inflation in unit values will depend on shopping and discounting strategies, while a BLS-style inflation indicator will not fully capture these effects. Hence, the gap in measured inflation rates for these two constructs will vary whenever the depth or frequency of discounts changes. Prior research suggests that business cycle conditions and the level of inflation impact the depth and frequency of sales. Our results suggest that a best price indicator helps explain unit value inflation in a parsimonious way.

Clearly, the importance of strategic consumer responses to temporary sales is of paramount importance in some sectors, and of more limited importance in others. However, as Varian notes in his 1999 *Handbook of Industrial Organization* survey of price discrimination, sellers almost always *want* to engage in price discrimination and price discrimination schemes involve substantial computational costs. Both the consolidation of the retailing sector over the last decades and the rapid decline in IT costs suggest that data-driven price discrimination schemes are likely to become more, rather than less important in the future. Thus, if macroeconomists are to successfully model price-setting, and statistical agencies are to successfully measure prices and inflation, confronting price discrimination appears to be an inevitable challenge.

REFERENCES

Acquisti, Alessandro and Varian, Hal R. (2005). 'Conditioning prices on purchase history', *Marketing Science* 24(3), 367–381.

Aguiar, Mark and Hurst, Erik. (2007). 'Life-Cycle Prices and Production', *The American Economic Review* 97(5), 1533–1559.

Basker, Emek. (2013), 'Change at the Checkout: Tracing the Impact of a Process Innovation', Working Papers 1302, Department of Economics, University of Missouri.

Basker, Emek, Klimek, Shawn D. and Pham, Van H. (2010), 'Supersize It: The Growth of Retail Chains and the Rise of the Big Box Retail Format', Working Papers 0809, Department of Economics, University of Missouri.

Bradley, Ralph, Cook, Bill, Leaver, Sylvia G. and Moulton, Brent R. (1997), "An Overview of Research on Potential Uses of Scanner Data in the US CPI", Technical report, Ottawa Group International Working Group on Price Indices Third Meeting.

Bronnenberg, Bart J., Kruger, Michael W. and Mela, Carl F. (2008). 'Database paper-The IRI marketing data set', *Marketing Science* 27(4), 745–748.

Burke, Jeremy M., Taylor, Curtis R. and Wagman, Liad. (2012). 'Information acquisition in competitive markets: An application to the US mortgage market', *American Economic Journal: Microeconomics* 4(4), 65–106.

Caves, Douglas W., Christensen, Laurits R. and Diewert, Walter Erwin. (1982). 'Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers', *Economic Journal* 92(365), 73–86.

Chevalier, Judith A., Kashyap, Anil K. and Rossi, Peter E. (2003). 'Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data', *American Economic Review* 93(1), 15–37.

Davies, George R. (1924). 'The problem of a standard index number formula', *Journal of the American Statistical Association* 19(146), 180–188.

Davies, George R. (1932). 'Index Numbers in Mathematical Economics', *Journal of the American Statistical Association* 27(177A), 58–64.

Diewert, W. Erwin. (1995), 'Axiomatic and Economic Approaches to Elementary Price Indexes', NBER Working Papers 5104, National Bureau of Economic Research, Inc.

Eichenbaum, Martin, Jaimovich, Nir and Rebelo, Sergio. (2011). 'Reference Prices, Costs, and Nominal Rigidities', *The American Economic Review*

101(1), 234–262.

Ellickson, Paul B. and Misra, Sanjog. (2008). 'Supermarket pricing strategies', *Marketing Science* 27(5), 811–828.

Ellison, Glenn. (2005). 'A Model of Add-on Pricing', *The Quarterly Journal of Economics* 120(2), 585–637.

Feenstra, Robert C. and Shapiro, Matthew D. (2003*a*), 'High-Frequency Substitution and the Measurement of Price Indexes', *in* 'Scanner Data and Price Indexes', NBER Chapters, National Bureau of Economic Research, Inc, pp. 123–150.

Feenstra, Robert C. and Shapiro, Matthew D. (2003*b*), 'High-Frequency Substitution and the Measurement of Price Indexes', *in* 'Scanner Data and Price Indexes', NBER Chapters, National Bureau of Economic Research, Inc, pp. 123–150.

Guimaraes, Bernardo and Sheedy, Kevin D. (2011). 'Sales and monetary policy', *The American Economic Review* 101(2), 844–876.

Handbury, Jessie, Weinstein, David E. and Watanabe, Tsutomu. (2013). 'How Much Do Official Price Indexes Tell Us About Inflation?', *NBER Working Paper* 19504(w19504).

Holmes, Thomas J. (2001). 'Bar Codes Lead to Frequent Deliveries and Superstores', *RAND Journal of Economics* 32(4), 708–25.

Hosken, Daniel and Reiffen, David. (2007). 'Pricing behavior of multiproduct retailers', *BE Journal of Theoretical Economics* 7.

Ivancic, Lorraine, Diewert, W. Erwin and Fox, Kevin J. (2009), 'Using a Constant Elasticity of Substitution Index to estimate a Cost of Living Index: from Theory to Practice', Technical Report 2010 ECON 15, Australian School of Business, University of New South Wales.

Kehoe, Patrick J. and Midrigan, Virgiliu. (2012). 'Prices are sticky after all', *NBER Working Paper* 16364.

Klenow, Peter J. and Willis, Jonathan L. (2007). 'Sticky information and sticky prices', *Journal of Monetary Economics* 54, 79–99.

Kryvtsov, Oleksiy and Vincent, Nicolas. (2014), 'On the Importance of Sales for Aggregate Price Flexibility', Working papers, Bank of Canada.

Lal, Rajiv and Matutes, Carmen. (1994). 'Retail Pricing and Advertising Strategies', *The Journal of Business* 67(3), 345–370.

Nakamura, Alice, Diewert W. Erwin Greenlees John S. Nakamura

Leonard I. and Reinsdorf, Marshall B. (2014), Sourcing Substitution and Related Price Index Biases, *in* 'Measuring Globalization: Better Trade Statistics for Better Policy', WE Upjohn Institute for Employment Research, pp. 21–88.

Nakamura, Emi and Steinsson, Jon. (2008). 'Five Facts about Prices: A Reevaluation of Menu Cost Models', *The Quarterly Journal of Economics* 123(4), 1415–1464.

Nakamura, Leonard I. (1998). 'The retail revolution and food-price mismeasurement', *Business Review* (May), 3–14.

Nakamura, Leonard I. (1999). 'The measurement of retail output and the retail revolution', *Canadian Journal of Economics* 32(2), 408–425.

Pashigian, B. Peter. (1988). 'Demand Uncertainty and Sales: A Study of Fashion and Markdown Pricing', *American Economic Review* 78(4), 936–953.

Pesendorfer, Martin. (2002). 'Retail sales: A study of pricing behavior in supermarkets', *The Journal of Business* 75(1), 33–66.

Reinsdorf, Marshall B. (1999). 'Using Scanner Data to Construct CP1 Basic Component Indexes', *Journal of Business & Economic Statistics* 17(2), 152–160.

Salop, S. and Stiglitz, Joseph E. (1982). 'The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents', *American Economic Review* 72(5), 1121–30.

Shapiro, Matthew D. and Wilcox, David W. (1997), 'Alternative Strategies for Aggregating Prices in the CPI', NBER Working Papers 5980, National Bureau of Economic Research.

Sobel, Joel. (1984). 'The timing of sales', *The Review of Economic Studies* 51(3), 353–368.

Taylor, Alan M. and Taylor, Mark P. (2004). 'The Purchasing Power Parity Debate', *Journal of Economic Perspectives* 18(4), 135–158.

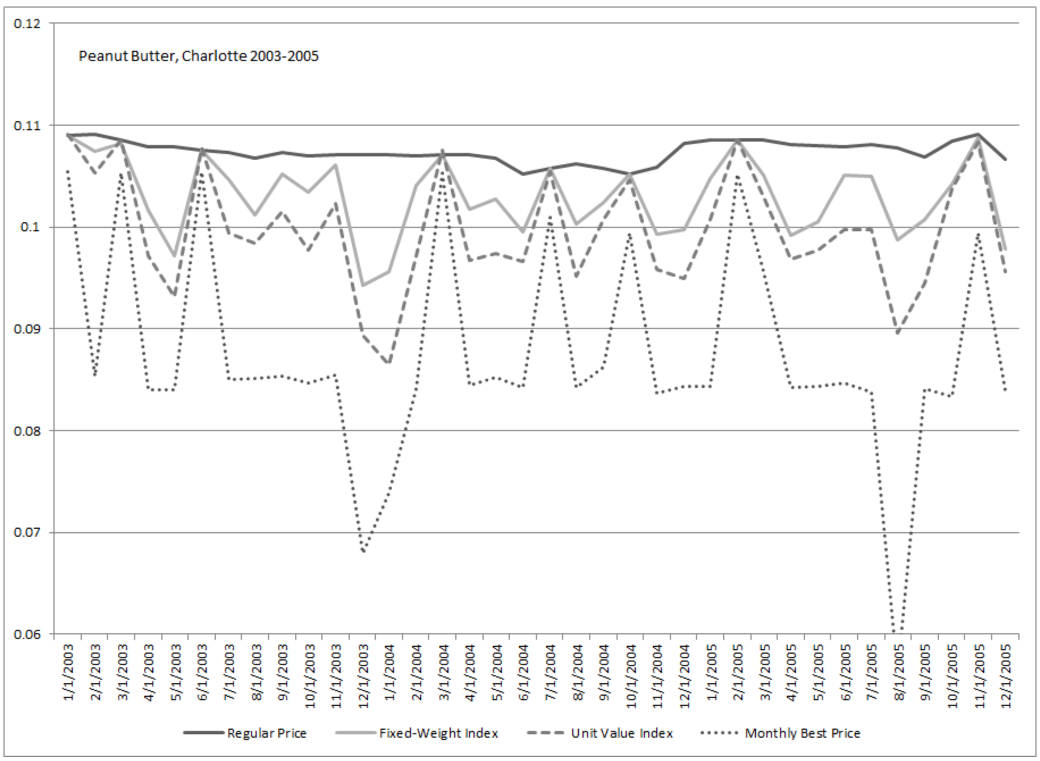
Triplett, Jack E. (2003), 'Using Scanner Data in Consumer Price Indexes. Some Neglected Conceptual Considerations', *in* 'Scanner data and price indexes', University of Chicago Press, pp. 151–162.

Varian, Hal Ronald. (1980). 'A Model of Sales', *American Economic Review* 70(4), 651–59.

Walsh, Correa Moylan. (1921), *The problem of estimation*, King & Son.

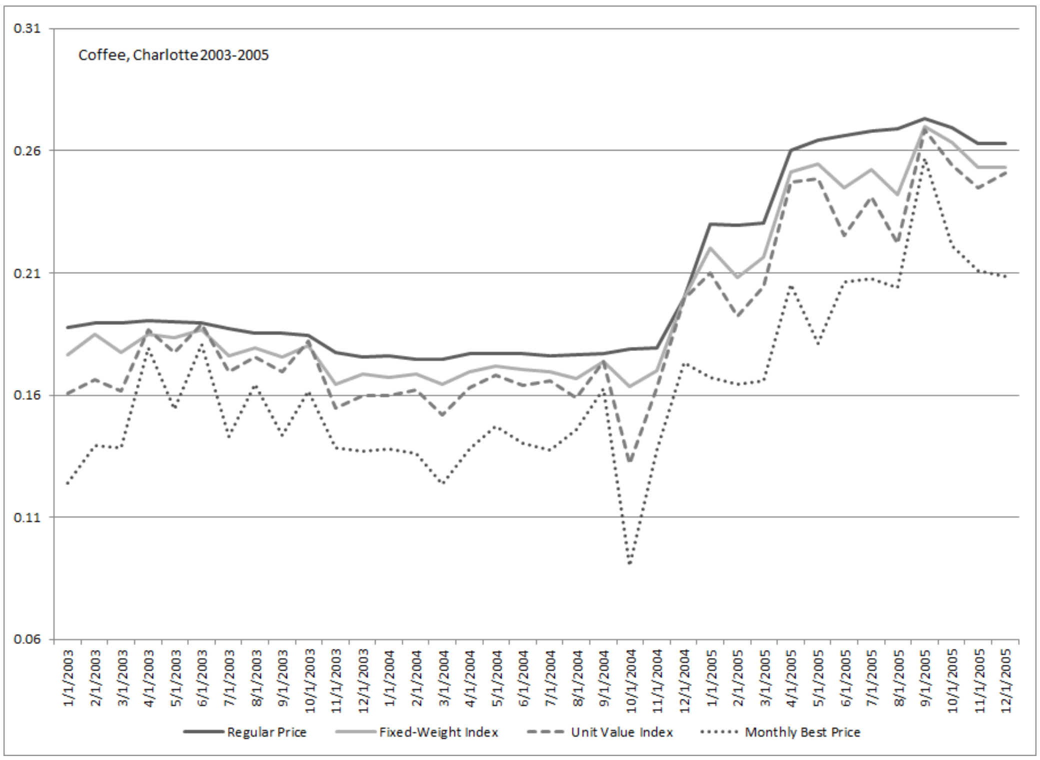
Wong, Arlene and Nevo, Aviv. (2014), 'The Elasticity of Substitution between Time and Market Goods: Evidence from the Great Recession', Working paper, Department of Economics, Northwestern University.

Figure 1. Monthly Peanut Butter Regular Prices, Fixed Weight Index, Variable Weight Index, and Best Prices.



Note: The figure shows a fixed weight index of regular prices (defined in the text and produced using our modification of the Kehoe and Midrigan (2012) algorithm for detecting sales), a fixed weight index of posted prices, unit values, and the monthly best price.

Figure 2. Monthly Coffee Regular Prices, Fixed Weight Index, Variable Weight Index, and Best Prices.



Note: The figure shows a fixed weight index of regular prices (defined in the text and produced using our modification of the Kehoe and Midrigan (2012) algorithm for detecting sales), a fixed weight index of posted prices, unit values, and the monthly best price.

Table 1—Share of Ounces Sold and Share of Weeks at Regular and Sale Prices: Totals for Sample Cities

Peanut butter	Ounces sold		Weeks		
	Regular price	Sale price	Regular price	Sale price	Avg Disc
Charlotte	60.03%	39.97%	75.91%	24.09%	17%
Chicago	33.92%	66.08%	59.17%	40.83%	21%
Hartford	50.08%	49.92%	92.45%	7.55%	27%
Houston	63.49%	36.51%	74.57%	25.43%	12%
Knoxville	65.24%	34.76%	73.13%	26.87%	11%
Los Angeles	49.49%	50.51%	65.83%	34.17%	13%
New York	37.49%	62.51%	78.63%	21.37%	21%
St Louis	34.88%	65.12%	67.73%	32.27%	26%
West Tx-New Mexico	46.26%	53.74%	68.60%	31.40%	19%
Average	48.99%	51.01%	72.89%	27.11%	19%
Coffee					
	Regular price	Sale price	Regular price	Sale price	Avg Disc
Charlotte	31.51%	68.49%	54.23%	45.77%	9%
Chicago	43.27%	56.73%	52.01%	47.99%	13%
Hartford	18.56%	81.44%	49.01%	50.99%	12%
Houston	42.89%	57.11%	57.83%	42.17%	6%
Knoxville	44.59%	55.41%	56.10%	43.90%	7%
Los Angeles	41.48%	58.52%	50.42%	49.58%	14%
New York	13.16%	86.84%	43.79%	56.21%	16%
St Louis	31.88%	68.12%	52.71%	47.29%	11%
West Tx-New Mexico	40.32%	59.68%	50.98%	49.02%	9%
Average	34.18%	65.82%	51.90%	48.10%	11%

Note: Totals are given for the store identified in the Data Appendix for each city. The sale and regular prices are calculated as described in the text based on our modification of the Kehoe and Midrigan (2012) algorithm for detecting sales. The average discount is the simple average over all sales or any items on sale in a given week.

Table 2—Ounces Sold and Weeks at Regular and Sale Prices: Charlotte

	Ounces sold		Weeks	
Peanut Butter	Regular price	Sale price	Regular price	Sale price
Jif	66.10%	33.90%	77.66%	22.34%
Peter Pan	54.13%	45.87%	71.43%	28.57%
Skippy	52.75%	47.25%	79.07%	20.93%
TOTAL	60.03%	39.97%	75.91%	24.09%
Coffee				
	Regular price	Sale price	Regular price	Sale price
Folgers	34.18%	65.82%	47.50%	52.50%
JFG	72.36%	27.64%	85.45%	14.55%
Maxwell House	25.34%	74.66%	36.65%	63.35%
TOTAL	31.51%	68.49%	54.23%	45.77%

Totals are given for the sample store in Charlotte. The sale and regular prices are calculated as desc

Table 3—Structural Estimates of Price Coefficients: Explaining Store-Level Unit Values with Fixed Weight Price Indices and Best Prices

		Peanut Butter Coefficients		Coffee Coefficients
Charlotte	Fixed weight index	0.804	Fixed weight index	0.737
		(0.024)		(0.038)
	Best Price	0.234	Best Price	0.292
		(0.022)		(0.040)
	constant	-0.0038	constant	-0.0040
		(0.002)		(0.004)
Chicago	Fixed weight index	0.542	Fixed weight index	0.648
		(0.032)		(0.031)
	Best Price	0.548	Best Price	0.386
		(0.039)		(0.042)
	constant	-0.006	constant	0.0058
		(0.003)		(0.006)
Hartford	Fixed weight index	0.484	Fixed weight index	0.437
		(0.044)		(0.035)
	Best Price	0.587	Best Price	0.667
		(0.029)		(0.038)
	constant	-0.005	constant	-0.0162
		(0.004)		(0.005)
Houston	Fixed weight index	0.646	Fixed weight index	0.831
		(0.045)		(0.017)
	Best Price	0.319	Best Price	0.206
		(0.030)		(0.017)
	constant	0.0065	constant	-0.007
		(0.004)		(0.002)
Knoxville	Fixed weight index	0.664	Fixed weight index	0.678
		(0.037)		(0.028)
	Best Price	0.289	Best Price	0.306
		(0.032)		(0.031)
	constant	0.007	constant	0.0086
		(0.003)		(0.002)
Los Angeles	Fixed weight index	0.687	Fixed weight index	0.716
		(0.047)		(0.038)
	Best Price	0.316	Best Price	0.291
		(0.032)		(0.033)
	constant	0.0042	constant	0.0053
		(0.004)		(0.008)
New York	Fixed weight index	0.414	Fixed weight index	0.348
		(0.037)		(0.043)
	Best Price	0.590	Best Price	0.697
		(0.040)		(0.047)
	constant	0.005	constant	0.0009
		(0.004)		(0.008)
StLouis	Fixed weight index	0.808	Fixed weight index	0.646
		(0.073)		(0.023)
	Best Price	0.403	Best Price	0.375
		(0.045)		(0.020)
	constant	-0.0258	constant	-0.002
		(0.007)		(0.003)
West Tx-New Mexico	Fixed weight index	0.669	Fixed weight index	0.915
		(0.067)		(0.031)
	Best Price	0.401	Best Price	0.183
		(0.044)		(0.026)
	constant	-0.007	constant	-0.0257
		(0.007)		(0.006)

Note: For each city and category we run a single regression motivated by Equation (12) in the text. The dependent variable is the unit value for the dominant brands in the sampled store. The brands for each store are listed in Data Appendix Table A.1. The independent variables are the fixed weight index for those brands, the monthly best price amongst those brands and a constant. Standard errors are in parentheses.

Table 4—Explaining Store-Level Unit Values with BLS-style Geometric Mean Price Indices and Best Price

		Peanut Butter Coefficients		Coffee Coefficients
Charlotte	Geomean Index	0.827	Geomean Index	0.743
		(0.024)		(0.039)
	Best Price	0.209	Best Price	0.284
		(0.022)		(0.041)
	constant	-0.0037	constant	-0.003
		(0.002)		(0.003)
Chicago	Geomean Index	0.593	Geomean Index	0.694
		(0.033)		(0.031)
	Best Price	0.493	Best Price	0.336
		(0.039)		(0.042)
	constant	-0.007	constant	0.005
		(0.003)		(0.005)
Hartford	Geomean Index	0.503	Geomean Index	0.453
		(0.045)		(0.036)
	Best Price	0.571	Best Price	0.649
		(0.030)		(0.039)
	constant	-0.0053	constant	-0.0162
		(0.004)		(0.005)
Houston	Geomean Index	0.683	Geomean Index	0.863
		(0.044)		(0.017)
	Best Price	0.290	Best Price	0.173
		(0.030)		(0.017)
	constant	0.005	constant	-0.007
		(0.004)		(0.002)
Knoxville	Geomean Index	0.689	Geomean Index	0.699
		(0.038)		(0.028)
	Best Price	0.270	Best Price	0.285
		(0.032)		(0.032)
	constant	0.006	constant	0.0086
		(0.003)		(0.002)
Los Angeles	Geomean Index	0.732	Geomean Index	0.756
		(0.046)		(0.038)
	Best Price	0.276	Best Price	0.248
		(0.032)		(0.033)
	constant	0.0029	constant	0.0051
		(0.004)		(0.007)
New York	Geomean Index	0.441	Geomean Index	0.373
		(0.038)		(0.044)
	Best Price	0.567	Best Price	0.668
		(0.039)		(0.048)
	constant	0.004	constant	0.0005
		(0.004)		(0.008)
StLouis	Geomean Index	0.825	Geomean Index	0.672
		(0.070)		(0.023)
	Best Price	0.373	Best Price	0.346
		(0.045)		(0.020)
	constant	-0.024	constant	-0.0016
		(0.006)		(0.003)
West TX- New Mexico	Geomean Index	0.726	Geomean Index	0.937
		(0.066)		(0.030)
	Best Price	0.353	Best Price	0.146
		(0.044)		(0.026)
	constant	-0.009	constant	-0.022
		(0.007)		(0.006)

Note: For each city and category we run a single regression. We replace the fixed weight price index that is suggested by Equation (12) with a BLS-style geometric mean price index. The dependent variable is the unit value for the dominant brands in that store. The brands for each store are listed in Data Appendix Table A.1. The independent variables are the geometric mean index for the brands under consideration in that store, the monthly best price for those brands and a constant. Standard errors are in parentheses.

Table 5—Explaining Store Level Unit Values with the Best Fit CES Index and the Best Price

		Peanut Butter Coefficients		Coffee Coefficients
Charlotte	CES Index 4.5	0.893	CES Index 2	0.748
		(0.027)		(0.041)
	Best Price	0.136	Best Price	0.276
		(0.025)		(0.043)
	constant	-0.0027	constant	-0.0021
		(0.002)		(0.004)
Chicago	CES Index 8	0.899	CES Index 7	0.98
		(0.040)		(0.031)
	Best Price	0.167	Best Price	0.026
		(0.044)		(0.039)
	constant	-0.006	constant	0.0052
		(0.003)		(0.004)
Hartford	CES Index 10	0.624	CES Index 10	0.562
		(0.052)		(0.042)
	Best Price	0.456	Best Price	0.525
		(0.036)		(0.045)
	constant	-0.0066	constant	-0.013
		(0.004)		(0.005)
Houston	CES Index 8.5	0.852	CES Index 5	0.998
		(0.049)		(0.019)
	Best Price	0.123	Best Price	0.032
		(0.034)		(0.019)
	constant	0.0053	constant	-0.0066
		(0.003)		(0.002)
Knoxville	CES Index 8	0.818	CES Index 8.5	0.873
		(0.044)		(0.035)
	Best Price	0.171	Best Price	0.118
		(0.036)		(0.039)
	constant	0.0031	constant	0.0068
		(0.003)		(0.002)
Los Angeles	CES Index 6.5	0.85	CES Index 4.5	0.844
		(0.050)		(0.041)
	Best Price	0.15	Best Price	0.128
		(0.037)		(0.036)
	constant	0.0036	constant	0.014
		(0.004)		(0.007)
New York	CES Index 9.5	0.692	CES Index 10	0.484
		(0.053)		(0.047)
	Best Price	0.377	Best Price	0.523
		(0.047)		(0.053)
	constant	-0.0057	constant	0.009
		(0.004)		(0.006)
StLouis	CES Index 10	0.778	CES Index 4.5	0.755
		(0.063)		(0.031)
	Best Price	0.252	Best Price	0.239
		(0.051)		(0.029)
	constant	-0.009	constant	0.0075
		(0.005)		(0.003)
West TX-New Mex	CES Index 7	0.925	CES Index 3.5	0.993
		(0.066)		(0.033)
	Best Price	0.105	Best Price	0.0544
		(0.051)		(0.029)
	constant	-0.0027	constant	-0.0158
		(0.005)		(0.006)

Note: For each city and category we run a single regression. We replace the fixed weight price index that is suggested by equation (12) with a constant elasticity of substitution (CES) price index. The CES substitution parameter for each store-product is separately chosen by a grid search to match the unit value index as closely as possible. The preferred substution parameter is shown in the table. The dependent variable is the unit value for the dominant brands in that store. The brands for each store are listed in Data Appendix Table A.1. The independent variables are the CES index for the brands under consideration in that store, the monthly best price for those brands and a constant.

Table 6—National specifications

VARIABLES	1 beer	2 coffee	3 deodorant	4 detergent	5 facial_tissue	6 frankfurters	7 frozen_pizza	8 ketchup
constant	-0.0032	0.0014	0.0002	0.0023	0.0068	-0.0061	-0.0020	-0.0003
	(0.0007)	(0.0021)	(0.0011)	(0.0021)	(0.0024)	(0.0017)	(0.0010)	(0.0017)
fixed weight	0.5537	0.2275	0.4877	0.3096	0.1979	0.2657	0.3177	0.3164
	(0.0451)	(0.0679)	(0.0371)	(0.0507)	(0.0362)	(0.0316)	(0.0400)	(0.0463)
R-squared	0.76	0.95	0.73	0.99	0.82	0.89	0.96	0.78
Observations	117	117	117	117	117	117	117	117

VARIABLES	9 margarine	10 mayo	11 milk	12 mustard	13 paper_towel	14 peanut	15 razors	16 shampoo
constant	0.0021	-0.0021	0.0015	0.0013	-0.0001	-0.0002	-0.0028	-0.0026
	(0.0011)	(0.0018)	(0.0008)	(0.0012)	(0.0022)	(0.0009)	(0.0014)	(0.0018)
fixed weight	0.3553	0.4523	0.5175	0.4619	0.2324	0.4991	0.6956	0.4500
	(0.0257)	(0.0557)	(0.0321)	(0.0392)	(0.0554)	(0.0320)	(0.0399)	(0.0373)
R-squared	0.98	0.92	0.99	0.93	0.85	0.98	0.84	0.82
Observations	117	117	117	117	117	117	117	117

VARIABLES	17 soft_drinks	18 soup	19 spaghetti	20 sugarsub	21 toothbrush	22 toilet_tissue	23 toothpaste
constant	0.0008	0.0004	-0.0010	0.0042	0.0038	-0.0009	0.0016
	(0.0015)	(0.0008)	(0.0009)	(0.0040)	(0.0013)	(0.0020)	(0.0012)
fixed weight	0.2943	0.3619	0.4268	0.6228	0.4472	0.1556	0.5169
	(0.0375)	(0.0276)	(0.0334)	(0.1343)	(0.0416)	(0.0573)	(0.0379)
R-squared	0.80	0.97	0.96	0.53	0.71	0.91	0.89
Observations	117	117	117	117	117	117	117

Note: For each category we estimate (15) using non-linear least squares. The coefficient on the fixed weight index, α, and the constant, γ, are reported along with their standard errors, which are shown beneath the coefficients in parentheses.

Table 7—National specifications - Geometric mean

VARIABLES	(2) beer	(5) coffee	(3) deodorant	(4) detergent	(6) facial_tissue	(7) frankfurters	(8) frozen_pizza	(9) ketchup
constant	-0.0035	0.0007	0.0004	0.0039	0.0068	-0.0063	-0.0021	0.0002
	(0.0008)	(0.0021)	(0.0013)	(0.0023)	(0.0024)	(0.0018)	(0.0011)	(0.0017)
geometric mean	0.5097	0.1569	0.4485	0.1831	0.2026	0.2648	0.3267	0.3307
	(0.0481)	(0.0786)	(0.0399)	(0.0429)	(0.0378)	(0.0335)	(0.0453)	(0.0487)
R-squared	0.72	0.95	0.66	0.98	0.81	0.88	0.95	0.78
Observations	117	117	117	117	117	117	117	117

VARIABLES	(10) margarine	(11) mayo	(12) milk	(13) mustard	(14) paper_towel	(1) peanut	(15) razors	(16) shampoo
constant	0.0026	-0.0020	0.0019	0.0020	0.0001	-0.0009	-0.0027	-0.0028
	(0.0011)	(0.0019)	(0.0010)	(0.0012)	(0.0023)	(0.0009)	(0.0015)	(0.0019)
geometric mean	0.3575	0.4265	0.5906	0.4991	0.1970	0.5379	0.6944	0.4503
	(0.0269)	(0.0606)	(0.0436)	(0.0408)	(0.0558)	(0.0352)	(0.0418)	(0.0411)
R-squared	0.98	0.91	0.98	0.94	0.84	0.98	0.82	0.80
Observations	117	117	117	117	117	117	117	117

VARIABLES	(17) soft_drinks	(18) soup	(19) spaghetti	(20) sugarsub	(21) toothbrush	(22) toilet_tissue	(23) toothpaste
constant	0.0005	0.0007	-0.0011	0.0043	0.0049	-0.0009	0.0020
	(0.0015)	(0.0008)	(0.0010)	(0.0041)	(0.0014)	(0.0021)	(0.0012)
geometric mean	0.3108	0.3934	0.4750	0.5952	0.4336	0.1193	0.5210
	(0.0419)	(0.0312)	(0.0396)	(0.1389)	(0.0436)	(0.0596)	(0.0406)
R-squared	0.79	0.97	0.96	0.51	0.68	0.91	0.87
Observations	117	117	117	117	117	117	117

Note: For each category we estimate an alternative version of (15)using non-linear least squares. Here the fixed weight index is replaced by the BLS-style geometric mean index. The coefficient on the geometric mean index, α, and the constant, γ, are reported along with their standard errors, which are shown beneath the coefficients in parentheses.

The U.S. census bureau separates the country into four regions which are in turn made up of nine divisions. For the store-level analysis we chose one store from each of the nine divisions; the national analysis relies on a much larger sample that is described below.

A1. Store Level Analysis

The particular stores were chosen as follows. First, we picked a random city in each division and then within that city identified the chain with the most number of stores. We then picked a random store within the chain and verified that it had no more than 15% of the weeks missing for the three main national peanut butter brands (Skippy, Jif and Peter Pan). We also checked that dominant selling UPCs were 16 to 18 ounce jars and that the main national brands were among the top sellers. If the most popular chain did not satisfy these conditions, then we selected the second largest chain in that city. We started this project before the most recent release of the data were available, so the screens were imposed using the data from 2001 to 2007. We used these same stores to study coffee, focusing on package sizes of 11 to 13 ounces.

To select UPCs for inclusion our analysis we start by identifying the top ten brands in each store in terms of average yearly sales. Importantly, most brands have multiple UPCs that are priced identically, such as "master blend" and "original roast" types of coffee. We aggregate all the UPCs within a brand where the correlation of the log prices price per ounce is greater than 0.85 (and the level of prices is within 15%). We do this iteratively to assemble all of the versions of a brand that are essentially the same. Once we have aggregated as many UPCs as is possible, we compute price for composite UPC by dividing the total dollar amount sales of all UPCs in the aggregate by the total ounces.

Having identified the top 10 UPCs in this fashion, we next eliminate private label and premium brands. We do this for three reasons. First, private label discounting strategies and demand is usually different than for branded items (at least for peanut butter and coffee). One way to see this is to recognize that the normal private label price is often lower than the sale price for branded UPCs, and yet many consumers do not switch. Second, the premium products (e.g. organic peanut butter or fair trade coffee) are such that even when they go on sale, they remain more expensive than the standard leading brands. So although there are undoubtedly some consumers that prefer generic or premium products, the willingness to switch between these products and the regular leading brands is undoubtedly more complicated than is posited by our model.

Finally, as a practical matter we are interested in exploring the importance of a best price for consumer behavior. The best price in many stores would almost always just be the private label price and the premium price would likely never be the best price. So by limiting the consideration to UPCs which have similar average prices we are capturing the kind of substitution that is described by the model. Hence, we implement this by pruning the set of candidate UPCs so that their average price per ounce is no more than 25% above or below the price for a reference price for peanut butter and coffee; the reference price for peanut butter in a given store is the average price of the national brands present in all the stores (Skippy, Jif and Peter Pan), while the reference price for coffee is the average price for the two national brands that were always present (Maxwell House and Folgers).[13]

The final step in our data construction is to exclude any UPCs which have substantial periods of missing data. We require a UPC to have been present in at least 6 years and to have non-missing observations for at least 60% of the total weeks in the sample. For the 9 stores in our store-level analysis this process leaves with us with between 3 and 5 brands per store that are used in computing the best price. The exact stores and brands are shown in Table A1.

Summary statistics for the price indices used city by city are found in Table A2. Using the rules described above to decide which UPCs qualify for consideration in each store, the price indices are defined as follows.

Fixed Weight: In each city, for each product, we construct a weighted average of the UPC prices. The weights in the first quarter of the sample are equal to the first quarter's expenditure share of each UPC. In each subsequent quarter, these weights are adjusted so that the weights are 15/16 of the prior quarter's weights and 1/16 of the current quarter's expenditure weight. This reflects the BLS's procedure of rotating sampling units (a combination of a store and product) over a four year cycle. The fixed weight is simply the weighted arithmetic mean of the prices.

Geometric Mean: The geometric mean is constructed the same way as the fixed

[13]To decide which UPCs are excluded, we compute UPC specific price deviations from the reference price in each store in each week and then compute the average value of the deviation. If that average is above 25 percent in absolute value the UPC is dropped.

Table A1—Brands used by market

Store (Market)	Peanut Butter Brands	Coffee Brands
250517 Charlotte, SC	JIF PETERPAN SKIPPY	FOLGERS JFG MAXWELLHOUSE
262433 Chicago, IL	JIF PETERPAN SKIPPY	EIGHTOCLOCK FOLGERS HILLSBROTHERS MAXWELLHOUSE
534239 Hartford, CT	JIF LEAVITTTEDDIE PETERPAN REESES SKIPPY	CHOCKFULLONUTS FOLGERS MAXWELLHOUSE
230491 Houston, TX	JIF PETERPAN SKIPPY	FOLGERS MAXWELLHOUSE SEAPORT
224312 Knoxville, TN	JIF PETERPAN SKIPPY	FOLGERS JFG MAXWELLHOUSE
286394 Los Angeles, CA	JIF PETERPAN SKIPPY	DONFRANCISCO FOLGERS MAXWELLHOUSE MELITTA YUBAN
279568 New York, NY	JIF PETERPAN REESES SKIPPY	CHOCKFULLONUTS FOLGERS MAXWELLHOUSE
232633 Saint Louis, MO	JIF PETERPAN SKIPPY	FOLGERS MAXWELLHOUSE WHITECASTLE
200439 West Texas/New Mexico	JIF PETERPAN SKIPPY	EIGHTOCLOCK FOLGERS MAXWELLHOUSE MJB

weight, but in the last step, a geometric mean of the weighted prices is calculated rather than an arithmetic mean.

Best Price One Month: We calculate the minimum price per ounce among all of the eligible UPCs over the entire month. For this calculation, each week of the year is assigned in its entirety to a calendar month. This is done because the data from the vendor are aggregated to the weekly level.

Best Price One Week: We calculate the minimum price per ounce among all of the eligible UPCs over each week.

Unit Value: We calculate the total spending on all the UPCs that qualify for consideration for each store and divide by total ounces.

A2. National Analysis

We complement our detailed findings for coffee and peanut butter by constructing national indices for more categories. Each category of the IRI Marketing Data Set is further divided (by IRI) into smaller categories. For example, the category "condiments" is divided into two subcategories, mustard and ketchup. We begin with these subcategories. We select 23 for our analysis. The main criteria for selection was that the category contained well-defined sizes of relatively homogeneous products and that the product set remained somewhat stable over the sample period. For example, we excluded diapers because the pricing is a function of both package sizes (24 diapers, 48 diapers, etc.) and sizes (Newborn, 3-6 months, up to 5T). Further, some of the more complex categories that we excluded experienced rapid product change. For example, in razor cartridges, 2 blade cartridges were the norm at the beginning of our sample, and had been supplanted by 3, 4, and 5 blade cartridges by the end of our sample. We also excluded products for which regulation and taxation could be a complicating factor (e.g. cigarettes). The remaining 23 categories are listed in Table A3.

For the calculation of national inflation estimates we followed the BLS sampling procedures to the extent possible. The BLS does not provide detail about product selection by category; their procedure is supposed to select a very representative item in each store. For each of our 23 product categories, we consider the full span of sizes that are amongst the sizes represented by the 10 highest overall revenue UPCs nationally in the first and last quarter of the data (2001q1 and 2011q4,

Table A2—Summary Statistics- City Data

		Peanut Butter	Coffee
Charlotte	Unit Value Price	0.116	0.248
	Fixed Weight Price	0.119	0.257
	Monthly Best Price	0.101	0.214
	Geometric Mean Price	0.118	0.256
	Total Ounces Sold	8073	3431
	Observations	129	129
Chicago	Unit Value Price	0.140	0.315
	Fixed Weight Price	0.151	0.328
	Monthly Best Price	0.118	0.25
	Geometric Mean Price	0.150	0.325
	Total Ounces Sold	4277	1221
	Observations	129	129
Hartford	Unit Value Price	0.126	0.224
	Fixed Weight Price	0.140	0.266
	Monthly Best Price	0.108	0.186
	Geometric Mean Price	0.138	0.264
	Total Ounces Sold	12898	10522
	Observations	129	129
Houston	Unit Value Price	0.118	0.274
	Fixed Weight Price	0.121	0.277
	Monthly Best Price	0.104	0.245
	Geometric Mean Price	0.121	0.276
	Total Ounces Sold	2414	2538
	Observations	127	127
Knoxville	Unit Value Price	0.118	0.248
	Fixed Weight Price	0.120	0.253
	Monthly Best Price	0.108	0.220
	Geometric Mean Price	0.120	0.252
	Total Ounces Sold	4501	2800
	Observations	129	129
Los Angeles	Unit Value Price	0.162	0.325
	Fixed Weight Price	0.165	0.341
	Monthly Best Price	0.141	0.258
	Geometric Mean Price	0.164	0.338
	Total Ounces Sold	4576	6339
	Observations	129	129
New York	Unit Value Price	0.123	0.221
	Fixed Weight Price	0.140	0.279
	Monthly Best Price	0.101	0.177
	Geometric Mean Price	0.139	0.275
	Total Ounces Sold	9218	15538
	Observations	129	129
St Louis	Unit Value Price	0.117	0.275
	Fixed Weight Price	0.129	0.288
	Monthly Best Price	0.097	0.239
	Geometric Mean Price	0.128	0.286
	Total Ounces Sold	9233	3339
	Observations	129	129
West Tx	Unit Value Price	0.138	0.314
	Fixed Weight Price	0.148	0.321
	Monthly Best Price	0.113	0.252
	Geometric Mean Price	0.147	0.319
	Total Ounces Sold	2692	1391
	Observations	121	121

respectively). We include all of these product sizes in our sampling procedure unless removing the 8th, 9th, or 10th most popular item from the group reduces the distance from the smallest to second smallest item or the largest to second largest item by more than 10 percent. This replicates the judgmental decision we made in deciding how to pick the package sizes for coffee and peanut butter by essentially dropping any UPCs with unusual sizes if their market share is low.

For example, the overall most popular products for liquid laundry detergent in 2001 and 2011 range from 50 to 200 ounces. However, the 200 ounce product is the ninth most popular product, and the next largest size represented in the list is 150 ounces. So, we define the category as containing sizes from 50 to 150 ounces. Prices are, of course, computed in per ounce measures for aggregation.

We gather all the UPCs that fit our criteria description in each store and calculate the total amount spent on these items in each month divided by the total ounces sold in that month. We call this the benchmark price per ounce for that store in that month. To exclude premium products, we keep all the UPCs which have a price that is plus or minus 25% of the benchmark price. Having trimmed the data in this fashion, we are left with a data set with the properties described in Table A3.

The IRI coverage does not match the population distribution of the U.S. so we do not want to just sample randomly from these stores. Accordingly, we divide the US into the four regions used by the BLS: The Northeast, Midwest, South, and West. We then sampled from each of these regions to get a distribution of stores that would mimic the BLS sampling weights for these regions. For each product for each month, we sampled 48 prices from the Northeast region, 48 prices from the Midwest region, 80 prices from the South and 64 prices from the West.

C. Choosing Stores

The stores in the national sample are initially chosen randomly using the total expenditure in that store for each category (relative to total expenditure for that category in the region) to determine the probability that the store is selected. At the time a store enters that sample, we randomly pick a week during the month at which price quotes from that store will be collected. If the chosen store has missing data it is replaced, drawing again proportionally to expenditure shares. Starting with the next quarter, we begin our sample rotation, whereby 1/16 of the stores will be replaced each quarter. (The initial order in which stores are replaced is random). To replace a store that is rotating out of the sample we draw a new one using expenditure weights from the prior quarter. We believe this procedure approximates the strategy that the BLS pursues in selecting outlets to sample.

D. Choosing UPCs

Based on total revenue for each UPC, we find the top 10 UPCs per store in the first quarter and use those while the store is in the sample. From the top 10 UPCs, we sample one per store. The probability of being chosen is proportional to each UPC's fraction of the spending relative to total spending for all of the 10 UPCs. If the chosen UPC is not available during a month, we choose another UPC from the top 10 for that period. When a new store rotates into the sample, its set of top 10 UPCs is identified using the expenditure shares from the prior quarter. A new UPC for that store will be selected and that UPC will be sampled for as long as the store is in the sample. If the selected UPC is missing then another from the top 10 will be randomly selected. This will mean that over time as stores change the list of UPCs is evolving to track recent purchase patterns.

E. Indices

A dataset containing all the sampled stores and UPCs comes out of this procedure. Each observation consists of information relating to a given week, month, and store. This information consists of the unit value (dollars paid per ounce), region, store, and whether the UPC is part of the top ten or a replacement. These annual inflation variables are summarized in Table A4.

Geometric Mean: This is our approximation of how the BLS would calculate an index for a given product. Each sampled store is sampled for one week of the month. We use one UPC per store and take the geometric mean across stores of the sampled prices per ounce for the month. The sampling procedure that governed the selection of stores and UPCs already accounts for the popularity of stores and UPCs, so the equally weighted geometric mean is what we report.

Fixed Weight: This differs from the geometric mean only because we take an arithmetic average of the UPCs rather than a geometric one.

Best Price One Week: Stores are sampled in one week of the the month as for the Geometric Mean. We then find the minimum price per ounce among the top 10 UPCs in the sampled store for the week. The index level is the arithmetic average of store best prices over the month.

Sampled Unit Value: We calculate the total spending on the top 10 UPCs divided by the total ounces for each store. We then calculate the arithmetic mean across stores assuming equal weights.

When an inflation is reported it is computed as the logarithmic change of an index. Note that the annual inflation measures are in many cases quite volatile. For coffee and peanut butter, the changes in prices correlate quite substantially with changes in the prices of the underlying agricultural commodities which are quite volatile. Prices in other categories may reflect technical change issues that are not captured in our index methodology (a problem that the BLS also confronts). For example, the highest average inflation levels reported below are for laundry detergent. It may be possible to clean a load of laundry with fewer ounces of detergent at the end of the sample period than at the beginning and that prices per unit cleaning power deviate from prices per ounce; our methodology does not capture this transition.

Table I.4—Summary Statistics

VARIABLES	(2) beer	(5) coffee	(3) deodorant	(4) detergent	(6) facial_tissue	(7) frankfurters	(8) frozen_pizza	(9) ketchup
fixed weight	0.0204	0.0694	-0.0001	0.0742	0.0244	0.0211	0.0105	0.0216
	(0.0186)	(0.1005)	(0.0243)	(0.2380)	(0.0640)	(0.0480)	(0.0520)	(0.0301)
geometric mean	0.0207	0.0716	-0.0006	0.0732	0.0242	0.0216	0.0109	0.0204
	(0.0199)	(0.1052)	(0.0272)	(0.2372)	(0.0633)	(0.0518)	(0.0546)	(0.0314)
unit value	0.0188	0.0744	0.0005	0.0724	0.0314	0.0218	0.0099	0.0189
	(0.0156)	(0.1117)	(0.0231)	(0.2483)	(0.0616)	(0.0558)	(0.0550)	(0.0399)
best price	0.0248	0.0745	0.0012	0.0640	0.0241	0.0322	0.0131	0.0176
	(0.0210)	(0.1190)	(0.0346)	(0.2297)	(0.0534)	(0.0661)	(0.0620)	(0.0444)
Observations	117	117	117	117	117	117	117	117

VARIABLES	(10) margarine	(11) mayo	(12) milk	(13) mustard	(14) paper_towel	(1) peanut	(15) razors	(16) shampoo
fixed weight	0.0588	0.0489	0.0363	0.0387	0.0400	0.0288	0.0337	-0.0016
	(0.0845)	(0.0670)	(0.0795)	(0.0460)	(0.0639)	(0.0592)	(0.0320)	(0.0333)
geometric mean	0.0578	0.0491	0.0359	0.0370	0.0396	0.0284	0.0337	-0.0011
	(0.0852)	(0.0702)	(0.0857)	(0.0479)	(0.0672)	(0.0591)	(0.0333)	(0.0374)
unit value	0.0540	0.0454	0.0377	0.0365	0.0386	0.0270	0.0313	0.0015
	(0.0836)	(0.0703)	(0.0846)	(0.0512)	(0.0635)	(0.0646)	(0.0398)	(0.0456)
best price	0.0461	0.0460	0.0363	0.0308	0.0382	0.0252	0.0357	0.0107
	(0.0811)	(0.0725)	(0.0933)	(0.0501)	(0.0638)	(0.0683)	(0.0556)	(0.0651)
Observations	117	117	117	117	117	117	117	117

VARIABLES	(17) soft_drinks	(18) soup	(19) spaghetti	(20) sugarsub	(21) toothbrush	(22) toilet_tissue	(23) toothpaste
fixed weight	0.0278	-0.0015	0.0165	0.0418	0.0002	0.0570	0.0033
	(0.0448)	(0.0437)	(0.0516)	(0.0469)	(0.0245)	(0.0794)	(0.0391)
geometric mean	0.0287	-0.0023	0.0164	0.0421	-0.0014	0.0589	0.0026
	(0.0446)	(0.0459)	(0.0520)	(0.0441)	(0.0276)	(0.0844)	(0.0407)
unit value	0.0282	-0.0030	0.0127	0.0429	-0.0009	0.0521	0.0015
	(0.0368)	(0.0476)	(0.0524)	(0.0612)	(0.0256)	(0.0801)	(0.0363)
best price	0.0273	-0.0046	0.0109	0.0299	-0.0098	0.0524	-0.0045
	(0.0396)	(0.0524)	(0.0481)	(0.0494)	(0.0368)	(0.0854)	(0.0420)
Observations	117	117	117	117	117	117	117