Ying Zhao
Hong Kong University of Science and Technology

Price Dispersion in the Grocery Market*

This paper provides an exploratory study of price dispersion in the grocery market and checks for the consistency of the evidence on this price dispersion with the existing theories of price dispersion due to costly consumer search, competition, and consumer heterogeneity in the context of the grocery market. Three dimensions of price dispersion are studied: price dispersion for a certain brand across stores, price dispersion within a category in a store across brands, and price dispersion over time for a certain brand. The results show that the standardized price dispersion is positively correlated with consumer search costs, competition, and consumer heterogeneity.

I. Introduction

The proverbial "law of one price" is virtually never empirically valid. Homogeneous goods are often sold at widely different prices by rival firms, even in environments that seem particularly conducive to economic competition. And prices for the same goods often vary over time.

The grocery market seems to be one of the prime examples of price dispersion in which there exists price uncertainty. As Blattberg and Neslin (1990, 344) have argued, "promotions are a way of life for retailers. Every week when the consumer goes to the grocery store, there are as many as 100 items on sale." We often observe price variations across stores, over time for a certain brand, and across brands within a category in a store. But surprisingly, price dispersion in the grocery market, for the most part, has not been

* The author would like to thank J. Miguel Villas-Boas as well as the editor, Pradeep Chintagunta, the anonymous referee, Severin Borenstein, Ganesh Iyer, Aviv Nevo, Russ Winer, and Florian Zettelmeyer for their helpful comments. The author would also like to thank David Bell for providing the data set for this research. The research reported in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, Hong Kong, China (Project no. DAG01/02.BM30). Contact the author at mkyzhao@ust.hk.

[*Journal of Business*, 2006, vol. 79, no. 3]

0021-9398/2006/7903-0006$10.00

studied,[1] even though a significant amount of the marketing literature has been dedicated to the grocery market.

The purpose of this paper is to provide an exploratory study of the various degrees of price dispersion observed in supermarkets and to check for its consistency with the existing theories of price dispersion due to costly consumer search, competition, and consumer heterogeneity. Three dimensions of price dispersion are studied in detail: price dispersion for a certain universal product code (UPC) across stores in a certain week, price dispersion within a category in a store across UPCs in a certain week, and price dispersion over time for a certain brand in a store. In this paper, price dispersion is calculated as the coefficient of variation, which is the ratio of the standard deviation of prices over the mean price.

This is the first paper that studies all three dimensions of price dispersion in the context of the grocery market, where anecdotal evidence suggests that price dispersion along all these dimensions is quite important. Note that the purpose of the paper is not to establish a structural model that takes into account the buyers' search processes and the sellers' pricing decision making. Such a model may require too many strong assumptions about the behaviors of buyers and sellers. Rather, this paper presents some general evidence on different sources of price dispersion in the context of the grocery market. These sources are consumer search costs, competition, and consumer heterogeneity. The results show that observed price dispersion is indeed positively correlated with higher consumer search costs, more intense competition, and greater consumer heterogeneity.

The rest of the paper is organized as follows. Section II briefly discusses the hypotheses on price dispersion and search cost, competition, and consumer

1. There are some studies about price variation in the grocery market. For example, Varian (1980) develops a theoretical model and shows that stores may find it in their interest to randomize prices in an attempt to price-discriminate between informed and uninformed consumers. Narasimhan (1984) presents a theoretical model that shows that consumers who use coupons are more price elastic than those who do not use them. Manufacturers can therefore be seen as using coupons as a price discrimination mechanism to reach price-elastic consumers. His model is empirically supported in tests using diary panel data. Narasimhan (1988) develops a model to study the equilibrium pricing strategies of brands engaged in a pricing game and shows that a mixed-strategy equilibrium exists with the firms randomizing over an interval. Raju, Srinivasan, and Lal (1990) analyze the role played by brand loyalty in determining optimal price promotional strategies used by firms in a competitive setting. Villas-Boas (1995) looks at supermarket prices for coffee and saltine crackers to test whether the sample of prices on each product could have possibly come from the theoretically specified density function predicted by Varian's (1980) model of sales. Rao, Arjunji, and Murthi (1995) offer empirical generalizations that competitive promotions are mixed strategies using data from grocery chains and from Information Resources Inc. scanner panels. Lal and Villas-Boas (1998) study retail price promotions and manufacturer trade deals in markets with multiproduct retailers. Pesendorfer (2002) examines temporary price reductions on ketchup products in supermarkets in Springfield, MO, between 1986 and 1988. Chintagunta, Dube, and Singh (2003) study the impact of price discrimination by a large Chicago supermarket chain and show how a store-pricing policy that is constrained to offer consumers at least as much surplus as a uniform chainwide pricing policy still enables the retailer to generate substantial incremental profits.

heterogeneity in the context of the grocery market. Section III describes the data set and constructs measures of search costs, competition, and consumer heterogeneity. Section IV studies price dispersion across stores. Section V models price dispersion across UPCs within a category in a store. Section VI focuses on price dispersion over time for a certain brand in a store. Concluding comments are presented in Section VII.

II. Price Dispersion and Search Cost, Competition, and Consumer Heterogeneity

Price dispersion arising from costly consumer search has been analyzed by a variety of researchers (see, e.g., Stigler 1961; Reinganum 1979; Salop and Stiglitz 1982; Burdett and Judd 1983; Carlson and McAfee 1983; Rob 1985; Stahl 1989; Kuksov 2006). This literature suggests that price dispersion can persist in markets in which there is imperfect information and consumers incur search costs to obtain price information. As consumers incur search costs to get price information, some consumers will engage in price searching and others will make purchases randomly. Therefore, sellers are able to charge different prices, and we may observe price dispersion in the market (e.g., Reinganum 1979). Price dispersion could also result from consumers' heterogeneous search costs if there is a positive mass of consumers with zero search costs (Salop and Stiglitz 1982). The heterogeneous search costs of consumers can also cause price dispersion over time for a certain product in a store (Varian 1980). Therefore, we expect that

HYPOTHESIS 1. There is a positive correlation between search costs and price dispersion for a certain brand across stores, for a certain brand over time, and across brands within a category in a store.

Price dispersion could also be an outcome of competition in the industry. The theory predicts that the price dispersion within a seller should decrease with increased competition if industry elasticities are the more prevalent basis for segmentation (monopoly-type discrimination), and it should increase with increased competition if heterogeneity in cross elasticity is the more common source of discrimination (competitive-type discrimination) (Borenstein 1985; Holmes 1989). There is also some empirical evidence suggesting that price dispersion is an outcome of competition, for example, in the U.S. airline industry(Borenstein and Rose 1994) and in the Irish grocery market (Walsh and Whelan 1999).

Also, as competition across stores increases, stores may choose to be more vertically differentiated from each other to relax the price competition (Tirole 1997; Iyer 1998). If stores were positioned too close to each other, then consumers would choose between them on the basis of price, which would create the incentive to compete on price, and the net result would be lower profit for both stores (Moorthy 1988). Stores may choose to be differentiated in dimensions such as services and product assortment to soften price com-

TABLE 1 Summary of the Dimensions and Some of the Sources of Price Dispersion

	Across Stores	Within a Store	Over Time
Search cost	+ Reinganum (1979)		+ Varian (1980)
Competition across stores	+ Tirole (1997)	+ Borenstein and Rose (1994)	
Consumer heterogeneity	+ Diamond (1987)	+ Shepard (1991)	

NOTE.—A + sign indicates that existing literature suggests a positive correlation between the row and the column variables.

petition. Price dispersion for the same product across stores may increase as store competition increases and stores are more vertically differentiated from each other because of the competition. Consequently, we expect that

HYPOTHESIS 2. There is a positive correlation between competition across stores and price dispersion for a certain brand across stores and across brands within a category in a store.

Stores are able to exercise some price discrimination based on differences among consumers' price elasticities, search costs, preferences, or willingness to pay for quality. A seller's motivation for price discrimination is likely to increase with the variation of attributes in the population that reflect buyers' search costs, price elasticities, or preferences (Diamond 1987; Shepard 1991). Therefore, we expect that

HYPOTHESIS 3. There is a positive correlation between consumer heterogeneity and price dispersion for a certain brand across stores and across brands within a category in a store.

Table 1 provides a summary of the relationship between the three dimensions and the three potential sources of price dispersion that is suggested by the economic literature. In this paper, we focus on testing these hypotheses using data from the grocery market.

III. Data

A. General

The data are taken from a scanner panel across 23 categories and include store price, consumer shopping history, and demographics. The data are drawn from a suburban area of Chicago and cover a two-year period from June 1991 to June 1993. Six supermarkets are tracked in the market, with one of the supermarkets appearing in the data set only in the eighty-fifth week. The 10 brands with the largest market share are chosen from each of the 23 product categories for the analysis. In the following analysis, let us denote i as UPC, j as the product category, k as the store, t as the week, and h as the consumer.

The product categories used in the empirical study and some summary statistics are listed in table 2. In this table, price dispersion is calculated as

TABLE 2 Summary of Price Dispersion (Average over Two Years)

Product	Across Stores (1)	Across UPCs		Over Time (4)
		Brand Effects (2)	Size Effects (3)	
Bacon	.20	.67	.009	.17
Barbecue sauce	.16	.80	.003	.09
Butter	.16	.44	.003	.15
Cat food	.14	.95	.005	.07
Cereal	.09	.49	.001	.11
Cleaners	.13	.61	.025	.05
Coffee	.12	.89	.011	.12
Cookies	.10	.79	.003	.09
Crackers	.14	.76	.004	.13
Detergents	.10	.86	.001	.07
Hot dogs	.18	.51	.002	.16
Ice cream	.16	.87	.002	.11
Peanut butter	.11	.18	.001	.06
Analgesics	.12	.48	.001	.04
Pizza	.11	.81	.002	.09
Salted snacks	.10	.81	.002	.11
Soap	.09	.67	.003	.05
Dryer softener	.08	.71	.001	.06
Soft drinks	.20	.46	.001	.19
Sugar	.08	.63	.015	.05
Tissue	.18	.46	.011	.12
Paper towels	.12	.42	.009	.07
Yogurt	.13	.17	.001	.09

NOTE.—This table reports summary statistics of price dispersion across stores, price dispersion across UPCs within a category, and price dispersion over time in the data set. Price dispersion is calculated as the coefficient of variation. The data set includes six stores and 23 product categories and is collected over two years. Col. 1, price dispersion across stores, is calculated as the price dispersion across stores for a certain UPC at a certain week aggregated over UPCs in the category and the period of two years. Cols. 2 and 3, price dispersion across UPCs, are calculated as the price dispersion across UPCs within a category in a store aggregated over stores and time. Col. 4, price dispersion over time, is calculated as the price dispersion over time for a certain UPC in a store aggregated over stores and UPCs.

the coefficient of variation, which is the ratio of the standard deviation of prices over the mean price. Price dispersion across stores, across UPCs within a category, and over time for a certain UPC are calculated. Note that to calculate price dispersion across brands, I weighted the prices by their market share when computing the mean prices. The table gives the aggregate results.

To calculate price dispersion across UPCs, I use the unit price (dollars/ounce) for each UPC to allow comparability of the prices of UPCs with different sizes. Price dispersion across UPCs is broken down into two parts: price dispersion due to size effects and price dispersion due to brand effects. The data show that the unit price for different UPCs under the same brand name varies substantially, and the unit price decreases with the size of the UPC. This is consistent with the argument that manufacturers encourage buyers to buy larger sizes of a product. In order to break down price dispersion across UPCs into these two parts, I ran a regression of the unit price of each UPC on the size and the square of the size of the UPC. The square root of the explained variance over the mean price could be seen as the price dis-

persion due to size effects. I determine price dispersion due to brand effects by subtracting the price dispersion due to size effects from price dispersion across UPCs.

An interesting fact shown by the table is that price dispersion due to brand effects is the largest, followed by price dispersion across stores, price dispersion over time, and price dispersion due to size effects.

The table also shows that there is substantial variation in the level of price dispersion over the different product categories. Price dispersion across stores is the lowest for dryer softener and sugar, at 8% of the mean product price, and highest in soft drinks and bacon, at 20%. The price dispersion across UPCs due to size effects for cleaners is the highest, at 2.5% of the mean product price, and lowest for detergent, butter, analgesics, dryer softener, and soft drinks, at 0.1%. Price dispersion across UPCs due to brand differentiation is the lowest for yogurt, at 17% of the mean product price, and the highest for cat food, at 95%. For price dispersion over time, soft drinks exhibit the highest variation and analgesics the lowest. The patterns of price dispersion that vary across categories partially reflect the underlying category-specific impact on prices in the category. Some product categories, for example, soft drinks, exhibit the highest price dispersion across stores and over time. Some of the sources of price dispersion are expected to come from consumer search costs, competition, and consumer heterogeneity.

B. Measures of Consumer Search Costs

Search costs are not directly observable. However, one could argue that they are likely to be higher the less frequently a consumer visits a store, buys from a category, or buys a certain brand. The rationale is that when shoppers make purchases more frequently, they are better informed of the price information in the market, and the price information they gained from the previous purchase may still be relevant. In addition, the costs of time for those consumers who make purchases more frequently are smaller. Therefore, search costs for them may be relatively smaller.

Three measures of search costs that can be tested are generated: SFS_j, the frequency of visiting stores for those consumers making purchases in category j; SFC_j, the consumers' purchase frequency of product category j; and SFB_i, the purchase frequency of brand i. I expect that as purchase frequency increases, search costs decrease, leading to a decrease in price dispersion. The rationale is that as search costs increase, stores are able to charge different prices, which leads to an increase in price dispersion across stores for the same product.

C. Measures of Competition

An interesting fact revealed by the data is that a new store enters the market at week 85. With more stores in the market, competition between the stores increases. This allows us to test the relationship between competition and

price dispersion in the market. To capture the competition, let $\text{DSTORE}_t = 0$ in the first 84 weeks and $\text{DSTORE}_t = 1$ otherwise. Thus DSTORE is an indicator of market competition. Competition is more intense when $\text{DSTORE}_t = 1$. It is expected that as there are more stores in the market, competition increases, which leads to an increase in price dispersion.

The Herfindahl index (HFD_{it}) is used to measure industry concentration across stores for brand i. However, price dispersion might have some impact on industry concentration in the same time period. For instance, suppose that there are two stores, A and B, that offer prices, p_A and p_B, for the same product, with $p_A > p_B$, at time t. Then sales for this product in store B will go up. Thus the concentration in this period will go up. To remove the likely impact of price dispersion on industry concentration in the same time period, I use a lagged concentration index. A lag of four weeks is used in the study.[2] As the Herfindahl index increases, the industry concentration increases, competition decreases, and therefore, price dispersion is expected to decrease.

D. Measures of Consumer Heterogeneity

Consumer heterogeneity is measured by the coefficients of variation in particular consumer demographic variables including the coefficient of variation of the education level of female shoppers for those households that made purchases in category j, VFEDU_j, and the coefficient of variation of income for those households that made purchases in category j, VINCOME_j. Female demographics are used in this study because it has been suggested in the literature that the female family member tends to play a bigger role in household grocery shopping. In particular, education level will affect the efficiency of the housewife in organizing her time (Narasimhan 1984) or searching for information. It is expected that as the coefficients of variation in consumer demographics increase, consumer pools vary more, leading stores to discriminate on prices more. As a result of this, price dispersion increases.

IV. Price Dispersion across Stores

The measure of price dispersion for a certain UPC, i, in category j at time t across K stores is computed as follows:

$$\text{DISP}_{it} = \frac{S_{it}}{\bar{P}_{it}} \times 100,$$

where S_{it} is the standard deviation of P_{ikt} (the price of i at store k at time t), and $\bar{P}_{it}$ is the mean price of brand i at time t across the stores.

2. Different lags were used in the model estimation, and the results did not change.

TABLE 3 Regression Results for Price Dispersion across Stores

Variable	Parameter Estimate	Standard Error
Intercept	.127	.859
SFB	−.05	.01
SFC	−.36	.07
SFS	−.91	.94
HFD	−.19	.09
DSTORE	.016	.03
t	.07	.04
CYCLE	.08	.02
VINCOME	.82	.06
VFEDU	.81	.11
AR(1)	−.27	.01
R^2	.17	

NOTE.—This table presents the regression results for price dispersion across stores. The dependent variable is price dispersion across stores for a certain UPC at a certain week. Price dispersion is measured as the standard deviation of the prices for a UPC across stores at a certain week over the mean of these prices. The independent variables SFB, SFC, and SFS are frequency of buying a brand, frequency of making purchase in a category, and frequency of visiting stores; these three variables are used to test the correlation between search costs and price dispersion. The independent variable HFD is the Herfindahl index across stores for a UPC, and the independent variable DSTORE is a store entry dummy variable; these two variables are used to test the correlation between competition and price dispersion. The variable *t* stands for week, which is used to control for seasonal effects. The variable CYCLE represents total product unit sales of a category in the current period as a proportion of total product unit sales for the category in the initial period. It is used to control for changes in the product cycle. The variables VINCOME and VFEDU are, respectively, the coefficients of variation of income across households and the coefficient of variation of the female education level across households; these two variables are used to test the correlation between consumer heterogeneity and price dispersion.

On the basis of the arguments outlined above, the model is as follows:

$$\begin{aligned}\mathrm{DISP}_{it} = {} & \beta_0 + \beta_1 \mathrm{SFS}_{j_i} + \beta_2 \mathrm{SFC}_{j_i} + \beta_3 \mathrm{SFB}_i + \beta_4 \mathrm{HFD}_{it} \\ & + \beta_5 \mathrm{DSTORE}_t + \beta_6 t + \beta_7 \mathrm{CYCLE}_{j_i t} + \gamma X_{j_i} + \epsilon_{it}. \qquad (1)\end{aligned}$$

Besides the frequency and competition variables, I include time, *t*, to control for factors that vary over time in the market. I also control for changes in the product cycle or seasonal effects with the variable $\mathrm{CYCLE}_{j_i t}$. This represents the total product unit sales of category *j* in the current period as a proportion of the total product unit sales for category *j* in the initial period, with higher values indicating boom periods. It is expected that competition increases as the market becomes larger, leading to increases in price dispersion. The coefficients of variation in consumer demographics, VINCOME and VFEDU, are included in vector X_{j_i}, and ϵ_{it} is the error term that is brand and time specific.

Table 3 reports the empirical results. Since the data used in the study are time-series data, a Durbin-Watson test of the error structure suggests an AR(1) error structure. The model is estimated using generalized least squares with an AR(1) error structure. The three frequency measures are used to test the correlation between search costs and price dispersion; it is expected that as the purchase frequency increases, search costs will decrease, and therefore price dispersion will decrease as well. The signs of the frequency measures

are therefore expected to be negative. This is confirmed by the estimation results. The coefficients of the purchase frequency variables are negative, with the purchase frequency of a category and the purchase frequency of a brand being significant.

The two variables HFD and DSTORE are used to test the correlation between competition and price dispersion across stores. It is expected that as competition across stores increases, stores will try to differentiate vertically from each other, and price dispersion across stores will be greater. The coefficient of HFD is negative and significant. This suggests that as the Herfindahl index increases, competition decreases, which leads to decreases in price dispersion across stores. This could be the result of stores choosing to be more vertically differentiated as competition increases.

The store entry variable is not significant. However, it is directionally correct. The positive sign of the store entry variable suggests that when there are more stores in the market, competition across stores increases, and therefore, price dispersion across stores increases.

The coefficients of variation in female household members' education level and income have a positive and significant correlation with price dispersion. The results suggest that as consumer heterogeneity increases, stores have more incentives to differentiate vertically from each other, and price dispersion across stores is more likely to increase.

A. Price Dispersion after Controlling for Category Effects

One potential issue in the previous estimation is that there are too many things changing across categories. There may exist some spurious correlations between purchase frequency and price dispersion across categories. In order to test the true correlation between purchase frequency and price dispersion, I may need to include category dummy variables in the estimation. By doing so, I am able to control for the category-specific impact. But I could not do so in the previous estimation because the measures of frequency of store visits and category purchases are perfectly collinear with the category dummy variables.

In order to include category dummies in the estimation, three time-varying measures of purchase frequency are constructed. These three measures are time-varying measures of frequency of purchasing a brand, of purchasing in a category, and of visiting stores. The measures are constructed on the basis of the following assumption: the more frequently consumers make purchases around a time t, the higher the probability for consumers to make a purchase at time t. I allow the consumer's probability of making a purchase to change over time.

More specifically, the time-varying measures of purchase frequency are constructed as follows: Let the utility of consumer h purchasing brand i at time t be y_{hit}. We have $y_{hit} = \alpha(t) + \beta X_{hit} + \xi_{hit}$, where ξ_{hit} is the standard normal error, X_{hit} is the time between t and consumer h's last purchase of

TABLE 4 Regression Results for Price Dispersion across Stores (with Control for Category Effects)

Variable	Parameter Estimates	Standard Error
Intercept	3.081	.272
SFB_t	−.012	.004
SFC_t	−.003	.001
SFS_t	−3.529	.311
HFD	−.022	.009
DSTORE	.002	.002
t	.077	.019
CYCLE	.018	.002
VINCOME	.214	.018
VFEDU	−.045	.031
AR(1)	.168	.01
R^2	.35	

NOTE.—This table presents the regression results for price dispersion across stores with category effects controlled (category effects are not reported in the table). The dependent variable is price dispersion across stores for a certain UPC at a certain week. Price dispersion is measured as the standard deviation of prices for a UPC across stores in a certain week over the mean of these prices. The independent variables SFB_t, SFC_t, and SFS_t are, respectively, time-specific frequency of buying a brand, making purchase in a category, and visiting stores; these three variables are used to test the correlation between search costs and price dispersion. Other variables are defined in the note to table 3.

brand i, and $\alpha(t) = \alpha_0 + \sum_{m=1}^{8} \alpha_m t^m$. Consumer h chooses to buy brand i at time t if, and only if, $y_{hit} > 0$. A probit model is estimated using the consumer purchase history data, and it yields the estimate $\hat{\alpha}(t)$. This $\hat{\alpha}(t)$ is an indicator of the intensity of consumer purchases. I thus construct a new variable, SFB_{it}, to simulate the shopping frequency of brand i in a category at time t, where $SFB_{it} = 1 + [\hat{\alpha}(t)/1{,}000]$. The time-varying frequencies of buying in a category (SFC_{jt}) and visiting a store (SFS_{jt}) are constructed following the same approach.

This is a strong test of the theory, because now we are looking at the correlation between search costs and price dispersion after controlling for category effects. The result is reported in table 4 (category effects are not reported). The estimated category effects give some evidence that price dispersion depends on the product category. Prices vary more for soft drinks and bacon and relatively less for dryer softener and sugar. The signs of the frequency and competition variables do not change even after I control for category effects. There are some mixed results on the correlation between the consumer heterogeneity variables and price dispersion across stores, with the coefficient of VINCOME being positive and significant and the coefficient of VFEDU being negative but not significant.

B. Alternative Measures of Purchase Frequency

The measures of frequency of visiting stores, buying in a category, and buying a brand are constructed from consumer purchase data. But price dispersion itself might affect consumers' purchase frequency. There could thus be an

TABLE 5 Regression Results for Price Dispersion across Stores (with Exogenous Measures of Purchase Frequency)

Variable	Parameter Estimates	Standard Error
Intercept	.075	.001
SFC	−.106	.005
HFD	−.18	.09
DSTORE	.002	.002
t	.07	.004
CYCLE	.08	.025
VINCOME	.048	.006
VFEDU	.131	.009
AR(1)	−.241	.009
R^2	.19	

NOTE.—This table reports the regression results for price dispersion across stores with exogenous measures of purchase frequency. The regression results come from using survey data about consumer purchase frequency in a category. The dependent variable is price dispersion across stores for a certain UPC at a certain week. Price dispersion is measured as the standard deviation of prices for a certain UPC across stores at a certain week over the mean of these prices. The variables are defined in the note to table 3.

endogeneity problem of the measures of the purchase frequency. To solve this problem, I obtained some exogenous measures of frequency of purchase.

A survey of consumers' purchase frequency for the 23 product categories was conducted in May 2000 at the San Francisco International Airport. Note that the data set used in the study is about the market in the Chicago area from June 1991 to June 1993. The purpose of the San Francisco survey was to attain measures of purchase frequency that are as disassociated from the Chicago market as possible. The new measure of purchase frequency is used in the estimation. The results again show that shopping frequency has a negative and significant correlation with price dispersion across stores and, more strikingly, that the correlations of shopping frequency with price dispersion are even larger when we use the exogenous measures of purchase frequency. The results are reported in table 5.

V. Price Dispersion across UPCs within a Category

The data suggest that prices vary significantly across UPCs within a category. Price dispersion across UPCs within a category in a store could be due to size effects or brand effects since different brands and different sizes for a brand are offered in the grocery market.

In this section, I study the relationship between price dispersion across UPCs (due to size effects and due to brand effects) and search costs, competition between stores, and consumer heterogeneity. Price dispersion is measured as the standard deviation of prices across UPCs over the mean price. Note that I weighted the brand prices by their market shares when calculating the mean prices.

TABLE 6 Regression Results for Price Dispersion across Brands Due to Brand Effects

Variable	Parameter Estimates	Standard Error
Intercept	26.48	2.44
SFC	−.21	.13
SFS	−25.66	2.49
HFD	−6.90	.92
DSTORE	−.06	.05
t	.04	.01
CYCLE	.17	.03
VINCOME	−1.29	.77
VFEDU	2.20	.26
AR(1)	−.68	.006
R^2	.56	

NOTE.—This table presents the results for price dispersion across UPCs due to brand effects. The dependent variable is price dispersion across UPCs within a category in a store at a certain week that is due to brand effects. The variables are defined in the note to table 3.

A. Price Dispersion across UPCs Due to Brand Effects

Price dispersion across UPCs within a category in a store may be due to brand differentiation. The presence of a large number of brands within a product category is indicative of a wide variety of goods that consumers can potentially choose. Consumers have heterogeneous brand preferences. There are a couple of potential sources of price dispersion across UPCs within a category due to brand differentiation. First, as consumer heterogeneity increases, stores are able to charge different prices for different brands; therefore, price dispersion across UPCs within a category will increase. Second, consumers differ in their willingness to switch between different brands. If stores can price-discriminate on the basis of consumers' willingness to switch to alternative brands, then price dispersion across UPCs will increase as competition across stores increases. Third, as consumers have search costs, stores are able to charge different prices for UPCs within a category; therefore, price dispersion across UPCs within a category will increase. By studying price dispersion across UPCs within a category due to brand differentiation, we can also capture some of the manufacturers' competitive pricing behavior, because if there are higher search costs, manufacturers can charge different prices.

The model of price dispersion across UPCs within a category due to brand effects is

$$\begin{aligned}\mathrm{DISP}_{jt} = {} & \beta_0 + \beta_1 \mathrm{SFC}_j + \beta_2 \mathrm{SFS}_j + \beta_3 \mathrm{HFD}_{jt} \\ & + \beta_4 \mathrm{DSTORE}_t + \beta_5 \mathrm{CYCLE}_{jt} + \beta_6 t + \phi X_j + \epsilon_{it}. \end{aligned} \tag{2}$$

All the explanatory variables in the above equation have the same specifications as before.

The model is estimated using generalized least squares allowing for an AR(1) error structure. The model results are reported in table 6.

The coefficients of the frequency of purchasing in the category and the

frequency of visiting stores are −0.21 and −25.66, respectively, with the coefficient of the frequency of visiting stores being significant. This result is consistent with the theory of price discrimination based on consumer search costs. As search costs increase, consumers are less willing to do price searches, and stores are more likely to charge different prices for brands within a category.

The coefficient of HFD is −6.90, which is significant. This suggests that as competition across stores increases, price dispersion across UPCs within a category due to brand effects increases. This result is consistent with Borenstein and Rose's (1994) study on the price dispersion in the U.S. airline industry, where they find that the dispersion in the prices an airline charges to different passengers on the same route increases as competition between airlines increases. As competition across stores increases, differences in consumers' willingness and ability to switch to alternative brands become more important, leading to increases in price dispersion across UPCs within a category due to brand effects. The coefficient for DSTORE is negative and is not significant.

There are some mixed results on the correlation between the consumer heterogeneity variables and a store's price discrimination motivation and consequently price dispersion across UPCs due to brand effects. The coefficient of the variation in female educational level is positive and significant, which suggests that stores have more motivation to exercise price discrimination since the female education level across households varies more. The coefficient of the variation of household income is negative and not significant.

B. Price Dispersion across UPCs Due to Size Effects

Price dispersion across UPCs within a category in a store may also be due to size effects since different sizes of a brand are often offered by a manufacturer, and manufacturers set nonlinear prices for different sizes of a brand.

The model of price dispersion across UPCs due to size effects has the same specification as the model of price dispersion across UPCs due to brand effects. It is estimated using generalized least squares allowing for an AR(1) error structure. The model results are reported in table 7.

The coefficients of the two frequency variables are both negative, and the coefficient for the frequency of purchasing in a category is significant. This is consistent with price discrimination based on search costs. When search costs increase, consumers are less willing to engage in price searching across different sizes within a category. Price dispersion across UPCs within a category due to size effects therefore increases.

The correlation between competition across stores and price dispersion across brands due to size effects is tested using the HFD and DSTORE variables. The coefficient of HFD_{jt} is −0.139 and is significant. The coefficient of DSTORE is directionally correct.

The coefficient of the variation in female educational level is positive and

TABLE 7 Regression Results for Price Dispersion across Brands Due to Size Effects

Variable	Parameter Estimates	Standard Error
Intercept	.012	.078
SFC	−.016	.004
SFS	−.028	.078
HFD	−.139	.034
DSTORE	.001	.002
t	.041	.015
CYCLE	.002	.001
VINCOME	−.059	.045
VFEDU	.050	.008
AR(1)	−.268	.007
R^2	.11	

NOTE.—This table presents the results for price dispersion across UPCs due to size effects. The dependent variable is price dispersion across UPCs within a category in a store at week *t* that is due to size effects. The variables are defined in the note to table 3.

significant, which suggests that stores have more motivation to exercise price discrimination since the female education level across households varies more. The coefficient of the variation of household income is negative and not significant.

VI. Price Dispersion over Time

So far, price dispersion across stores was estimated with a small number of observations in each period. There were only six stores at any given time included in the data set. This small number of stores may not reflect the real situation because there may be some other stores that are competing with these six stores but are not included in the data set. One way to solve this problem is to look at price dispersion over time for the same UPC at the same store. By studying price dispersion over time, we can capture relevant competing stores that are in the market but are not included in the data set if these stores have a pricing pattern similar to that of the stores in the data set.

By studying price dispersion over time, we may also capture price discrimination over time based on different consumer search costs. For instance, some consumers have high search costs and make purchases randomly. Other consumers have low search costs and are able to take advantage of good deals in the store; stores can price-discriminate on the basis of the different search costs for these consumers.

The measure of price dispersion over time, DISP_{it}, is calculated as the standard deviation of the price for brand *i* from week *t* to $t + 7$ over the mean

TABLE 8 Regression Results for Price Dispersion over Time

Variable	Parameter Estimates	Standard Error
Intercept	.126	.039
SFB_t	−.022	.013
SFC_t	−.05	.008
SFS_t	−.198	.034
HFD	−1.435	.052
DSTORE	.068	.020
t	.023	.014
CYCLE	.204	.019
VINCOME	.117	.015
VFEDU	.315	.025
AR(1)	−.49	.002
R^2	.44	

NOTE.—This table presents the results for price dispersion over time. The dependent variable is price dispersion over time for a certain UPC in a store. The variables are defined in the note to table 4.

price for brand i for this time period.[3] It appears that eight weeks is a reasonable cycle for price change. The model is

$$\mathrm{DISP}_{it} = \beta_0 + \beta_1 \mathrm{SFC}_{j_i t} + \beta_2 \mathrm{SFB}_{it} + \beta_3 \mathrm{SFS}_{j_i t} + \beta_4 \mathrm{HFD}_{it} + \beta_5 \mathrm{DSTORE}_t + \beta_6 t + \beta_7 \mathrm{CYCLE}_{j_i t} + \sum_{j=2}^{23} \gamma_j D_{j_i} + \phi X_{j_i} + \epsilon_{it}. \quad (3)$$

Time-varying purchase frequency measures are included in the model to allow for the control of category specifics. The variables DSTORE and HFD are used to test the correlation between competition between stores and price dispersion over time. I also control for changes in the product cycle with the variable CYCLE. Time, t, is used to control for the time-varying effects. Category-specific dummy variables, D_{j_i}, are used to control for the category effects on price dispersion. Coefficients of variation of consumer demographics are included in the vector X_{j_i} to control for consumer heterogeneity.

The model is estimated using generalized least squares allowing for an AR(1) error structure. Results of the model are reported in table 8 (category effects are not reported here).

The coefficients of the time-varying frequencies of buying a brand, buying in a category, and visiting stores are −0.022, −0.05, and −0.198, and they are all significant. The results suggest that as search costs increase, price dispersion over time for a product increases. This is consistent with the theory of search costs and price discrimination. As search costs increase, a store is more likely to randomize its price for the same product to discriminate between those consumers who are willing to search for prices and those who shop randomly.

3. I also construct the DISP_{it} variable using time periods other than eight weeks. This does not affect the results presented below.

The relationship between competition between stores and price dispersion over time is tested by using DSTORE and HFD variables. The coefficient of DSTORE is 0.068 and is significant. This indicates that the increase in competition resulting from the entry of a new store increases price dispersion across time for a certain brand. The coefficient of HFD is -1.435 and is significant. These results show that price dispersion over time for a product in a store increases as competition across stores increases. The results are consistent with Borenstein and Rose's (1994) study on the price dispersion in the U.S. airline industry. Consumers differ in their purchase behavior: some consumers make purchases randomly, whereas others engage in price searching and make purchases when prices are low. Therefore, price competition for those consumer segments that tend to buy at low prices is greater and the prices paid by those consumers are lower. Hence, price dispersion over time increases as competition across stores increases.

The coefficients of the consumer heterogeneity variables are both positive and significant. This suggests that price dispersion over time for a product increases as consumer heterogeneity increases. Again, price dispersion over time depends on categories. For example, given search cost, paper towels have lower price dispersion and cookies have higher price dispersion.

VII. Discussion and Conclusion

This paper provides an exploratory study on price dispersion in the grocery market and empirically investigates some of the potential sources of price dispersion across stores, across UPCs within a category, and over time for a certain brand. This study found support for the positive correlation between search costs and price dispersion on all three dimensions. There is also some evidence of competitive price discrimination across stores, across brands within a category, and over time for a certain brand. Some positive correlation is also found for the heterogeneity in the consumer pool and price dispersion across stores and over time.

Anecdotal evidence suggests that price dispersion is important in the grocery market. We often observe price dispersion across stores, across brands within a category, and over time in the grocery market. There are some studies in the marketing literature that look at the grocery market, but few pertain to price dispersion in the grocery market. Price dispersion itself is price discrimination that is important to understand. The fact that different product categories are sold in the grocery market allows us to check different levels of price dispersion across categories. The rich data set, which includes consumer purchase history and demographics as well as store prices for various brands over a long period of time, allows us to study the three dimensions of price dispersion. By understanding some of the sources of price dispersion in the grocery market, marketing managers may be able to make better decisions on marketing strategies. For example, if the prices for the same product are different between two grocery stores and if the manager at the high-price

store understands that the price difference is due to consumer search costs instead of competition, then he or she does not need to lower the price for that product at his or her store to match the competitor's price.

The analysis presented here makes it possible for us to study some of the sources of price dispersion in a market in which both firms' and consumers' decision-making processes are rather complicated. In this process, consumers incur search costs, and firms need to take into account consumers' search costs when making pricing decisions. Consumers also need to make decisions on where to search and how much information to gather. Moreover, the information consumers have gained from previous shopping experiences may still play a role, so that this is a dynamic process for consumers. A structural model in this context may require too many strong assumptions about firm and consumer behavior.

In conclusion, this is the first paper that empirically studies three dimensions of price dispersion in the grocery market. The study documents the existence of significant price dispersion across stores, over time, and across brands within a category in a store in the grocery market. Normalized price dispersion is greater when search costs are higher, competition is more intense, and consumer heterogeneity is greater.

References

Blattberg, Robert, and Scott Neslin. 1990. *Sales promotion: Concepts, methods, and strategies*. Englewood Cliffs, NJ: Prentice-Hall.

Borenstein, Severin. 1985. Price discrimination in free-entry markets. *Rand Journal of Economics* 16:380–97.

Borenstein, Severin, and Nancy L. Rose. 1994. Competition and price dispersion in the U.S. airline industry. *Journal of Political Economy* 102:653–83.

Burdett, Kenneth, and Kenneth L. Judd. 1983. Equilibrium price dispersion. *Econometrica* 51: 955–70.

Carlson, John, and Preston McAfee. 1983. Discrete equilibrium price dispersion. *Journal of Political Economy* 91:480–93.

Chintagunta, Pradeep, Jean-Pierre Dube, and Dan Vishal Singh. 2003. Balancing profitability and customer welfare in a supermarket chain. *Quantitative Marketing and Economics* 1:111–47.

Diamond, Peter. 1987. Consumer differences and prices in a search model. *Quarterly Journal of Economics* 102:429–36.

Holmes, Thomas. 1989. The effects of third-degree price discrimination in oligopoly. *American Economic Review* 79:244–50.

Iyer, Ganesh. 1998. Coordinating channels under price and nonprice competition. *Marketing Science* 17:338–55.

Kuksov, Dmitri. 2006. Search, common knowledge, and competition. *Journal of Economic Theory*, forthcoming.

Lal, Rajiv, and J. Miguel Villas-Boas. 1998. Price promotions and trade deals with multi-product retailers. *Management Science* 44:935–49.

Moorthy, K. Sridhar. 1988. Product and price competition in a duopoly. *Marketing Science* 7: 141–68.

Narasimhan, Chakravarthi. 1984. A price discrimination theory of coupons. *Marketing Science* 3:128–47.

———. 1988. Competitive promotional strategies. *Journal of Business* 61:427–99.

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

Raju, Jagmohan, V. Srinivasan, and R. Lal. 1990. The effects of brand loyalty on competitive price promotional strategies. *Management Science* 36:276–304.
Rao, Ram, Ramesh Arjunji, and B. P. S Murthi. 1995. Game theory and empirical generalizations concerning competitive promotions. *Marketing Science* 14:G89–G100.
Reinganum, Jennifer. 1979. A simple model of equilibrium price dispersion. *Journal of Political Economy* 87:851–58.
Rob, Rafael. 1985. Equilibrium price distribution. *Review of Economic Studies* 52:487–504.
Salop, Steven, and Joseph Stiglitz. 1982. The theory of sales: A simple model of equilibrium price dispersion with identical agents. *American Economic Review* 72:1121–30.
Shepard, Andrea. 1991. Price discrimination and retail configuration. *Journal of Political Economy* 99:30–53.
Stahl, Dale. 1989. Oligopolistic pricing with sequential consumer search. *American Economic Review* 79:700–712.
Stigler, George. 1961. The economics of information. *Journal of Political Economy* 69:213–25.
Tirole, Jean. 1997. *The theory of industrial organization*. Cambridge, MA: MIT Press.
Varian, H. 1980. A model of sales. *American Economic Review* 70:651–59.
Villas-Boas, J. Miguel. 1995. Models of competitive price promotions: Some empirical evidence from the coffee and saltine crackers markets. *Journal of Economics and Management Strategy* 4:85–107.
Walsh, Patrick, and Ciara Whelan. 1999. Modeling price dispersion as an outcome of competition in the Irish grocery market. *Journal of Industrial Economics* 41:325–43.