Rigid Prices: Evidence from U.S. Scanner Data*

Jeffrey R. Campbell† Benjamin Eden‡

October, 2005

Abstract

This paper uses over two years of weekly scanner data from two small US cities to characterize time and state dependence of grocers' pricing decisions. In these data, the probability of a nominal adjustment *declines* with the time since the last price change. This reflects differences over time in the flexibility of prices charged by a single store for a given good. We also detect state dependence: The probability of a nominal adjustment is highest when a store's price substantially differs from the average of other stores. However, extreme prices typically reflect the selling store's recent nominal adjustments rather than changes in other stores' prices.

JEL Classification: E31, L16, L81.

*The authors gratefully acknowledge research support from the National Science Foundation through grant 013048 to the NBER. Gadi Barlevy, Marco Bassetto, and Mario Crucini made many helpful suggestions. The views expressed in this paper are those of the authors and do not reflect those of the Federal Reserve Bank of Chicago, the Federal Reserve System, or its Board of Governors.

†Federal Reserve Bank of Chicago and NBER.

‡Vanderbilt University and University of Haifa

1 Introduction

This paper measures time and state dependence in grocers' pricing decisions using scanner data. The observations cover most transactions of items in five product categories in two Midwestern cities. These weekly records allow accurate calculation of the age of an item's price and of the corresponding prices paid at other stores.

Thirty percent of the prices in these data change in an average week, and this number falls little if we eliminate price changes that begin or end sales. If the probability of a nominal adjustment remains the same as a price ages, then this implies that the average price lasts three weeks. However, a randomly chosen price remains unchanged for nearly six weeks. The discrepancy between these duration estimates reflects a fact at odds with familiar macroeconomic models of pricing: the frequency of nominal adjustment *declines* with the time since the last price change. This counterintuitive time dependence does not merely reflect heterogeneity in price flexibility across products or stores. Instead, it arises from occasional spells of flexibility punctuating otherwise rigid prices.

Models of state-dependent pricing – such as those of Barro (1972); Sheshinski and Weiss (1977); Caplin and Spulber (1987); Caplin and Leahy (1991); Dotsey, King, and Wolman (1999); and Golosov and Lucas (2003) – imply that the benefit of a nominal adjustment is highest when the price differs substantially from other sellers' prices. Our findings reproduce this qualitatively: Increasing the difference between an item's price and the average price for the same item at other stores substantially raises the probability of a price change. However state-dependent pricing is quantitatively unimportant in our sample, because most price changes occur with prices already close to average. Furthermore, patterns of price adjustment are inconsistent with simple menu-cost models, in which extreme prices arise from the erosion of a fixed nominal price by other sellers' price adjustments. We find that grocers *choose* most extreme prices, which they then quickly abandon.

Many papers that examine price data collected to construct the CPI precede this work. Examples from the Euro zone, Israel, Poland, and the United States include Dhyne et al. (2004), Lach and Tsiddon (1992), Konieczny and Skrzypacz (2005), and Klenow and Kryvtsov (2005). These micro-CPI data record the prices for many more items than do the scanner data we employ. We therefore view this paper as complementing earlier studies by looking at a narrower but richer set of data. Our data are weekly. This is an advantage over the monthly CPI data because stores typically change prices more than once in a single month. Furthermore, our data are more detailed than CPI data. In Bils and Klenow (2004) for example, there is one product category called "margarine". We examine the pricing of 54 different margarine products. This allows us to construct relevant comparisons of prices across stores. Finally, scanners directly measure transaction prices with little human intervention; unlike BLS enumerators.

Dutta, Bergen, and Levy (2002) and Chevalier, Kashyap, and Rossi (2003) examined scanner data of prices at a single Chicago supermarket chain. These observations share the high frequency and accuracy of the data we employ, and in addition they record the supermarket's markup over wholesale cost. The advantage of this paper's data arises from their coverage of multiple sellers. In leveraging this feature, we follow Kashyap (1995). He compared price adjustments of three retailers selling a few identical items. Our work examines 94 grocery products each sold in five or more stores.

The remainder of this paper proceeds as follows. The next section discusses the source of the data and how we use it to detect nominal price changes. It also presents summary statistics and foretells our results with the behavior of the price for a specific item at one store. Section 3 measures time dependence of pricing decisions and presents estimates of a price's average duration. Section 4 studies the dependence of price changes on stores' relative prices. Section 5 unifies the study of time and state dependence by estimating linear regression models of the decision to change a nominal price. These results reinforce the impression that state dependence is quantitatively unimportant, because the relative price improves the model's forecasts only marginally. Section 6 discusses the robustness of our results to different measurement strategies, and Section 7 offers concluding remarks.

2 Data

Our data source is the ERIM scanner data set collected by A.C. Nielsen. The marketing department at the University of Chicago's Graduate School of Business graciously makes these data available on its web site.[1] Nielsen collected these data from two small Midwestern cities - Springfield, Missouri and Sioux Falls, South Dakota - from the fifth week of 1985 through the twenty-third week of 1987. The data come from the checkout scanners of these cities' supermarkets and drug stores. The sample includes observations from 23 stores in Springfield and 19 stores in Sioux Falls. Together, they account for about 80% of the two markets' grocery and drug retail sales. We identify a product with a UPC. These codes differ across different packagings of the same good (e.g. 8 oz. and 16 oz. sizes) and across different varieties of that good (e.g. flavored and unflavored margarine). For each product in six categories - ketchup, margarine, peanut butter, sugar, toilet tissue, and tuna - the data record the revenues from the sales of that product as well as the quantity sold at each store. We measure the price with average revenue per unit sold. The measure of revenues includes the face value of coupons, so changes in customers' coupon redemption do not directly influence these price measures. A.C. Nielsen also issued identification cards to approximately 10 percent of each city's households. These customers presented their cards at stores' checkout counters, and A.C. Nielsen used the resulting observations to construct household-level purchase histories. These allow us to observe the exact transaction prices and locations for goods purchased by these households.

To properly measure price changes, we require uninterrupted time series of individual stores' prices for particular goods. The price observations from households' purchase histories are incomplete for all but the most popular products and stores, because we only observe a store's price for a good in a given week if one of the sample households makes the corresponding purchase. Hence we use the average revenue per unit and employ the household-level records to assess their quality. Some product-store combinations do not

[1]The data can be found at `http://gsbwww.uchicago.edu/research/mkt/Databases/ERIM/ERIM.html`.

have complete sales data, because the store either introduced or discontinued the product during the sample period. Furthermore, in a few cases a store might not sell a particular product to any household during a given week. We use a balanced panel that does not include such incomplete observations. The final criterion for inclusion of a product in our sample is that no fewer than five stores sold it in any week. This allows us to determine whether a store's price is close to other stores' prices for the same good.

2.1 Measuring Price Changes

Time aggregation complicates the measurement of price changes when using average revenue per unit. To illustrate this problem, consider the daily price of a single good at a hypothetical store, plotted in Figure 1. In week 1, the price is \$1.29, and on Wednesday of week 2, the price increases to \$1.40. It stays at this level throughout week 3. If the quantity sold on a given day does not change over time, then the average revenues are \$1.29, \$1.3529, and \$1.40. The average revenue changes twice in these three weeks, while the daily price only changes once. The spurious price change arises because the second week's average revenue embodies two prices.

One symptom of averaging multiple prices in a week is that the price cannot be expressed in whole cents, as in our example. With this in mind, we applied a simple correction to average revenue that removes such spurious price changes. Any average-revenue price that cannot be expressed in whole cents which is either part of a descending or increasing sequence of prices (as in the example) is replaced by the price of the following week. Applying this rule to our example changes the price of \$1.3529 in week 2 to \$1.40; and the number of price changes in the corrected data is accurate. Average-revenue prices that were not in whole cents but were either greater than or less than both the previous and following week's prices were rounded to the nearest whole cent but otherwise left unchanged. If we change our example so that the store lowers the price from \$1.40 to \$1.26 on the beginning of week 3 and maintains this price through week 3, then the three weeks' average prices are \$1.29,

\$1.3529, and \$1.26. We count two price changes after rounding the second week's price down to \$1.35, as actually occurred.

To check the accuracy of the resulting price observations, we compared them with the median transaction price from the household-level records for each store-UPC-week combination. In 75 percent of the available pairs, the two prices matched exactly. For 10 percent of the remaining cases, the median transaction price equals either the previous week's or the following week's average revenue price. We do not expect these two series to match perfectly, because our correction does not eliminate the effects of time aggregation. Accordingly, we view this rate of matching between the two price measures as validating our observations' quality.

2.2 Summary Statistics

Before proceeding to examine price changes, we document some of the scanner data set's most salient characteristics. Table 1 provides summary statistics for the sample we use. Its first two columns report the number of products observed by category and the number of distinct prices recorded. The sample includes observations of 94 products, and most of these are either margarine or tuna. Overall, the data record $83,394$ prices.

The third column of Table 1 reports the fraction of prices that we identify as sale prices. We wish to ensure that this paper's results do not merely reflect firms' switching between sale and "regular" prices, because some authors discount these as variation arising from a simple pricing rule rather than a conscious change in that rule. To identify sales, we look for price declines of 10 percent or more in a given week that the store completely reverses within 2 weeks. All prices between the initial decline and the reversal are sale prices. With this criterion, only 2.8 percent of the observations are sale prices. Sugar's frequency of sale prices, 5.4 percent, exceeds that of any other category. Klenow and Kryvstov (2005) report that the BLS identifies 15 percent of food prices collected to produce the CPI as sale prices. Apparently, the stores in our sample use sale prices relatively infrequently.

Table 1's final column reports the annualized average rates of price change in percentage points. We find this of interest, because inflation expectations impact firms' price choices. The prices in all categories but Peanut Butter declined over the sample period. The corresponding average annual growth rate of the consumer price index for margarine is -1.7 percent. The matching *CPI*'s for the other categories all display price growth, so the deflation in Springfield and Sioux Falls did not typify the national experience.

Aggregate fluctuations in inflation also concern price-setting producers. To illustrate the sort of aggregate fluctuations facing the sample's stores, Figure 2 plots two monthly measures of annualized inflation for margarine over the scanner data's sample period. The first measure uses a geometric average fixed-weight price index conceptually similar to the *CPI*. We built this with the scanner data following the procedure of Richardson (2003). The second measure is the margarine *CPI* itself. The *CPI*-based inflation varies much less than the scanner-based inflation. Their standard deviations are 0.9 and 1.8 percent. Their sample correlation is 0.07, which suggests that location-specific shocks dominate the scanner-data based price index. We could not locate a national-level *CPI* for toilet tissue. The other categories' scanner-based inflation rates also display considerably greater variance than their corresponding *CPI*-based rates.

Next, consider the variability of prices across stores and time. The first column of Table 2 reports residuals' standard deviations from regressions of the price's logarithm against a set of UPC dummy variables. These standard deviations range from 11.4 percent for Peanut Butter to 16.4 percent for Tuna. The table's remaining columns report residual standard deviations from regressions that include progressively richer sets of dummy variables. The regression underlying the second column's results includes two sets of UPC dummy variables, one for each market. This accounting for systematic differences between prices in Springfield and Sioux Falls lowers the standard deviations little. The regression for the third column includes one set of UPC dummy variables for each market and week. As Figure 2 suggests, removing date-specific means substantially lowers variation. For example, margarine's standard deviation drops from 12.8 percent to 8.8 percent. Because $(8.8/12.8)^2$ approximately equals 1/2, the cross-sectional variance of prices at a given date and the time-series variance of the average price across dates roughly equal each other.

Table 2's final two columns further decompose the cross-sectional dispersion of prices. The regression used for the fourth column adds store-specific UPC dummies that are invariant across time to the regression in the third column. We expect this dummy to substantially reduce the standard deviations if stores consistently follow "low-price" or "high-price" strategies.[2] In fact, removing store-specific UPC dummies lowers the standard deviations only about one percentage point. This indicates that there are few *systematic* differences in the prices for a given product across either markets or stores. The final column quantifies the contribution of stores switching between sale prices and regular prices to price dispersion. For this, we added two sets of store-specific UPC dummies to the regression from the third column, one for regular prices and another for sale prices. Accounting for the differences between sale and regular prices lowers the standard deviations little. Overall, Table 2 points to aggregate time-series fluctuations and transitory idiosyncratic shocks that do not reflect sales as the two major sources of price variance.

Finally, we consider the frequency of price changes, which Table 3 reports. The first column gives the weekly average fraction of prices that changed for the whole sample and each of the six categories. Overall, thirty percent of prices change in a given week. The second column reports the frequencies after excluding prices changes to or from a sale price. Given the scarcity of sales in these data, it is unsurprising that sales account for only a minor fraction of price changes. Table 3's final columns compare the price changes in these data with those tabulated by the BLS while constructing the *CPI*. The third column computes the average monthly frequency that is obtained by "visiting" a store during the first week of each month, and the last column reports the BLS estimates as reported in Bils and Klenow

[2]Because the store-specific dummies also vary across UPC's, this regression will account for persistent heterogeneity across stores in the pricing of particular items that does not reflect store-wide pricing strategies.

(2004).[3] If the probability of a price changing did not depend on the price's age, as in Calvo's (1983) model of price adjustment, then the weekly frequencies we observe would imply that 3/4 of prices change in a given month. However, we find instead that approximately 1/2 of prices change when sampled monthly. This suggests that the assumption of a constant probability of price adjustment does not hold good in our data. In any case, the frequency of price changes in these data far exceed the BLS frequencies.

One explanation for the greater frequency of price changes in the data set we use is that the presence of scanners lowers the cost of price adjustment. Because scanners were ubiquitous in the late 1990's when the BLS sample was collected, we discount this explanation. Another possibility lies in the difference between the two methods used to collect prices. The BLS handbook of methods mentions the possibility that sometimes the last month's price may be used instead of the current price. In chapter 17 of the 2004 edition it says:

> The pricing methodology in the commodities and services component of the CPI allows the previous period's price to be available at the time of collection. This dependent pricing methodology is believed to reduce response variance for measuring change, but may cause response bias and lag.

Thus, BLS enumerators missing many small price changes might be contributing to the apparently heightened rate of price change in the scanner data. This is one example of human transcription error, which presumably infects the automatically-collected scanner data relatively little. For these reasons, we believe that the scanner data estimates of the frequency of price change are more reliable than those of the BLS.[4]

[3]Two of the scanner data's categories have identically named BLS item categories, Margarine and Peanut Butter. We matched Ketchup with "Other condiments (excl olives, pickles, and relishes)," Sugar with "Sugar and artificial sweeteners," Tissue with "Cleaning and toilet tissue, paper towels, napkins," and Tuna with "Canned fish or seafood."

[4]Bils and Klenow (2004) exploit the BLS observations to measure which product categories have more flexible prices. The scanner data suggests that the BLS underestimates the *level* of price flexibility, but we have no concrete evidence that they systematically fail to properly determine which product categories have

2.3 One Store's Price for Fleischmann's Margarine

We complete our description of the data with Figure 3, which plots one store's price of a single product (Fleischmann's Margarine) along with the average of all other store's prices for the same item.[5] This price changed 51 times during the 123 week sample, which is typical for the data. It begins at \$1.06, and it rises to \$1.09 in two steps in April of 1985. Between late May and early August, it experiences several price changes of one to three cents, but it returns to \$1.09 and finishes the year at that price. Such a return of a price to a "regular" price occurs often in these data. The average of other firms' prices approximately equaled \$1.10 throughout these eleven months. The price changed much more often in 1986. After three very modest price increases in January, the price dropped to \$0.92 for three weeks, rose to \$1.52 for one week, and then returned to \$1.09.[6] Throughout these changes, the average price at other stores fluctuated little. The return to \$1.09 lasted only six weeks. In April, the price entered a period with very frequent changes that ended only in September. At that point, it approximately settled at \$1.15. The year ended with two dramatic price increases that ultimately proved to be temporary. In 1987, the price returned to a pattern of much less frequent price changes. It ended the sample period at \$1.15.

This price's evolution foretells the key results of this paper. First, many extreme relative prices do not arise from a combination of unchanged nominal prices and changing competitors prices, as is standard in models of state-dependent pricing. Instead, the store chose most of the extreme relative prices that it charged. Second, extreme relative prices tend to be short-lived. The three price peaks of 1986 testify to this. Finally, periods of stable prices

more flexible prices.

[5]Here and throughout the paper, we construct the average of all other stores' prices for the same item by dividing total sales of the item across all other stores by the number of units sold by those stores. As a consequence, this price measure will embody the prices of some stores that do not belong to our balanced panel of price choices.

[6]Because this price decrease lasted more than two weeks, we do not label it as a sale. Indeed, our procedure identifies none of the prices for this particular product at this store as sale prices.

alternate with periods of frequent change.

3 Time Dependence

We now turn to the measurement of time dependence in grocers' decisions to change nominal prices. For this we use the unconditional hazard function, which plots the adjustment frequency as a function of the price's age.[7] In Calvo's (1983) model of stochastic price setting, this function does not vary with the price's age, whereas in Taylor's (1980) model of staggered pricing it equals zero until the interval of price rigidity passes, at which point it jumps to one. Standard models of state dependent pricing clearly imply the unconditional hazard function increases with the price's age for very young prices, because a producer gains nothing from changing a newly-set and hence optimal price.

Figure 4 plots the unconditional hazard function estimated using the observations from all product categories. The chance of a newly-set price changing equals 0.59. As the price ages, this probability drops precipitously. It equals 0.29 for a two-week-old price and 0.22 for a three-week-old price. As the price ages further, the hazard function continues its decline at a more gradual pace. For very old prices, the probability of a price change equals only 0.09. The unreported hazard functions calculated separately for each product category resemble Figure 4.

None of the standard models of nominal price rigidity produce a decreasing hazard function. In the remainder of this section, we uncover its source in two steps. We first quantify the decreasing hazard's economic significance by comparing two estimates of a price's average duration that equal each other when the hazard is constant. Next, we show that heterogeneous flexibility across price setters cannot account for the difference between them. This leads us to conclude that the decreasing hazard reflects true duration dependence.[8]

[7]In the following, we define a newly-set price as one week old.

[8]The regression analysis of price changes in Section 5 reinforces this conclusion by showing that the likelihood of a price changing declines with its age, even after controlling for store and UPC fixed effects.

3.1 Two Estimates of Average Duration

To place the decreasing hazard function of Figure 4 into context, it is helpful to recall that the estimated hazard function for a worker leaving unemployment typically decreases. This reflects the fact that the newly unemployed are more likely to find work than are their long-term counterparts. Because of this, the average remaining duration of unemployment for the stock of currently unemployed workers exceeds the corresponding average for the flow of the newly unemployed. This suggests a simple statistic for summarizing the hazard function's rate of change, the average duration of a randomly sampled price minus the average duration of a newly-set price.[9]

The first two columns of Table 4 report estimates of these two durations from our sample. Both columns report average realized durations of prices charged before the sample's last year. Omitting prices charged late in the sample practically eliminates bias from right-censoring. The first column reports the average duration for all prices charged. This equals 5.9 weeks for all products. Across categories, it varies from 2.8 weeks for Peanut Butter to 6.4 weeks for Tuna. The second column reports the average durations of the sample's newly-set prices. This equals 3.2 weeks for all products, and it varies little across categories. For Peanut Butter, these estimates differ by only a few days, but for the other categories the difference is considerably greater. For Tuna, the first estimate exceeds twice the second. For Margarine, focusing on newly-set prices reduces average duration by almost three weeks. In this sense, the decreasing hazard and its impact on the difference between these estimates are both quantitatively significant.

[9] If the hazard for a nominal price change is constant, as in Calvo's (1983) model, then this difference equals zero with a large sample. In Taylor's (1980) model of staggered price setting, the average duration of newly-set prices, T, exceeds the average duration of the stock of all prices, $(T+1)/2$.

3.2 Heterogeneity

Heterogeneity provides a simple explanation for the decreasing hazard function. We illustrate this here with an example adapted from the study of unemployment duration.[10] Suppose a price is either flexible or rigid. Within each type there is a constant probability of changing the price, but flexible prices change more frequently. Rigid prices have thus a longer average lifetime. The hazard function initially reflects the average probability across the two groups. As a cohort of prices set on a given date ages, the fraction of rigid prices among the survivors increases. The hazard function declines (as in Figure 4) and asymptotes to the probability of a rigid price changing. Figure 5 plots the implied raw hazard function from this example when the probability of flexible and rigid prices changing are 0.65 and 0.09 and 90 percent of all newly-set prices are flexible. This simple example reproduces Figure 4 well.

If the duration dependence in Figure 4 reflects the fact that some stores adjust their prices more frequently than others, then reconciling Figure 4 with simple models of price setting might be straightforward. The same holds true if some products' prices exhibit more flexibility than others. On the other hand, if price adjustment displays insufficient heterogeneity along these dimensions, then the decreasing hazard acquires economic meaning as a description of individual stores' prices.

Baharad and Eden (2004) use a simple implication of Jensen's inequality to determine the importance of heterogeneous price flexibility in their sample of Israeli grocery prices. Suppose that Calvo's (1983) model of nominal price setting describes individual stores' decisions well, but their probabilities of nominal adjustment differ. Let x_i denote the frequency of price changes at store i. Because $1/x$ is convex, Jensen's inequality yields

$$1/\left(\frac{1}{N}\sum_{i=1}^{N} x_i\right) \leq \frac{1}{N}\sum_{i=1}^{N}\frac{1}{x_i}. \tag{1}$$

The left-hand side of (1) is the inverse of the average price change frequency, while the right-hand side is the average of the inverse of store-specific price-change frequencies. Baharad

[10] See Darby, Haltiwanger, and Plant (1985).

and Eden (2004) calculated the sample analogues of both expectations from a sample of 381 products from a moderate inflation period in Israel. They found that the inverse of the average frequency equalled about four months, while the average of inverse frequencies was eight months.

If the decline in the hazard function arises primarily from heterogeneity of durations across products or stores, then the average of inverse frequencies estimate of the duration should be close to the directly measured average reported in Table 4's first column and there should be a large difference between the inverse of the average frequency and the average of inverse frequencies. The last two columns of Table 4 show that this is not the case. They report sample analogues of the expectations in (1) calculated from the scanner data. The third column is the familiar inverse-frequency estimate of average duration. These are very close to the average durations of new prices in the second column.[11] For the final column, we calculated the average price change frequency for each store-product pair, inverted the results, and averaged them. The estimates exceed the standard inverse frequency estimate by very little. For all categories but Peanut Butter, they lie far below the average durations in the first column. We conclude that the decreasing hazard function in Figure 4 reflects true duration dependence rather than heterogeneity across products or stores.

4 State Dependence

Existing models of state-dependent pricing cannot easily generate a decreasing hazard function like that we observe, because the benefit of changing a newly-set price is small and grows as the price ages. Nevertheless, the insights of state-dependent pricing models might yet improve our understanding of producers' nominal adjustments. In this section, we examine this possibility. We begin with an indicator of the benefit of a nominal adjustment,

[11]Because the number of price changes in its denominator equals the number of new prices and the sample length in its numerator approximately equals the sum of new prices' durations, the inverse frequency approximately equals the average duration of newly-set prices in the complete sample by construction.

the price relative to the average of other stores' prices for the same product. In standard state-dependent models, extreme values of this relative price lead to a higher probability of nominal adjustment.[12]

Figure 6 plots the observed frequency of price changes as a function of the relative price's logarithm. The relative price equals the ratio of the store's nominal price in the *previous* week divided by the sales-weighted average of all other stores' prices in the current week. This measures the nominal deviation that a price adjustment in the current week could close. The results we report change little if we instead measure the relative price using the average of all other stores' prices in the previous week. Before estimation, we accounted for some stores systematically following high-price or low-price rules by normalizing the mean of each store-product pair's log relative price to zero. On the horizontal axis, zero indicates a relative price equal to the average for this store-product pair. We divided the interval [-1/2,1/2] into twenty equally sized bins and calculated the price change frequency for each of them. The thick solid curve in Figure 6 gives the frequencies for all store-item-week observations. The thin solid gives the analogous frequencies calculated excluding sale prices. For visual reference, the light horizontal line gives the unconditional frequency of a price change, 0.3; and the dashed line plots the sample's distribution of relative prices.

There are four notable features of Figure 6. First, the minimum frequency substantially exceeds zero. For both samples, it equals 0.19. Thus, even a store with an "average" price might change it. Second, most of the relative price observations are located close to their average values. Together, these two observations strongly suggest that the relative price cannot substantially improve forecasts of the occurrence of nominal adjustment. Third, moving the relative price away from its average substantially increases the probability of a nominal adjustment. The estimated probability of a nominal adjustment is 80 percent when the price is 35 to 40 percent below average and 50 percent when it is 35 to 40 percent

[12]One possible objection to this measure of relative prices is that its denominator includes prices charged at other stores owned by the same firm. We address this below in Section 6.

above average. In this sense, these observations display a basic feature of menu-cost pricing models. Fourth and finally, excluding sales alters the results as expected. The frequencies are imperceptibly different for above-average prices, while they are substantially lower for prices more than 10 percent below average.[13]

In light of the negative association of a price's age with the probability that it changes, Figure 7 plots the frequency of price changes against the mean-adjusted log relative price for samples of young prices – those with ages less than or equal to three weeks – and for old prices – those with ages of four weeks or more. Neither sample includes sales, and both exclude prices from initial left-censored spells. As the results from time-dependent pricing suggest, the adjustment frequencies of new prices exceed those of old prices substantially. Furthermore, extreme relative prices increase the nominal adjustment frequencies of both young and old prices. Figure 7 also plots the estimated relative-price distributions for both young and old prices. Unsurprisingly, both of their modes are near the mean of zero. What is somewhat more surprising is that young prices display *more* dispersion than do old prices. Comparing the two distributions' peaks makes this excess dispersion particularly clear. To quantify this, we calculated the standard deviations of both price distributions. These are 14.3 percent and 8.4 percent for young and old prices. That is, although the data indicate that extreme relative prices last only a short time – as in state-dependent pricing models – producers *choose* these extreme prices rather than arrive at them after the erosion of a fixed nominal price's real value.

5 Forecasting Price Changes

Figure 7 goes some distance towards unifying the consideration of time and state dependence. This section continues in that direction by presenting forecasting models of the decision to change a store's nominal price. The estimated models reinforce most of the findings above.

[13] Category-specific versions of Figure 6 all display the same features, but the estimated frequencies are considerably noisier.

Quantitatively, the price's age contributes much more to the models' forecasts than does the relative price.

All of the models we estimate have the simple linear-in-probabilities form,

$$\Pr[p_{i,t} \neq p_{i,t-1}] = \beta' x_{i,t},$$

where $x_{i,t}$ is a vector of variables known at the time that $p_{i,t}$ is chosen.[14] It includes three sets of dummy variables spanning the sets of stores, products, and calendar dates, a dummy variable indicating whether or not the current price is a sale price, the mean-adjusted relative price, the inverse of the price's age, and their squares. Finally, it contains the mean-adjusted number of units sold by the firm in the previous period as well as its square. We add this to $x_{i,t}$ because Golosov and Lucas (2003) emphasize that firms' with high sales have a greater incentive to change prices in their state-dependent model.

We estimated the linear-in-probabilities model using ordinary least squares separately for each category and for the sample as a whole. Table 5 reports the estimated coefficients for the models' regressors of interest, their heteroskedasticity-corrected standard errors, and each model's R^2. Consider first the model estimated with all products' data. The regressors together explain 24.5 percent of the variation in the decision to change the nominal price. All of the coefficients are statistically significant at the 1 percent level. As Figure 6 suggests, the regression function is convex in the relative price. The coefficients multiplying linear and squared terms in lagged units sold are both positive, indicating that exceptionally large sales tend to lead to price changes, as Golosov and Lucas (2003) suggest. The quadratic in the inverse of the price's age is strictly increasing. That is, the counterintuitive declining hazard rate continues to manifest itself.

Because the sample size varies greatly across the categories, so does the precision of the estimated coefficients. The standard errors for the two categories with the fewest price observations, Ketchup and Peanut Butter, are particularly large. For all categories but Peanut

[14]This includes the logarithmic relative price defined above, which includes the prices all other stores charge at time t. We use this for conformity with the analysis in Section 4.

Butter, a joint exclusion test for the two relative price terms rejects the null hypothesis at the one percent level. The estimated coefficient on the squared term is positive for all of these categories, as the convex hazard in Figure 6 suggests.

The influence of lagged units sold on nominal adjustments varies more across the categories. Joint tests reject its exclusion from the models at the one percent level for Margarine, Sugar, and Tissue and at the ten percent level for Tuna. None of these estimates qualitatively resemble the estimates from pooling all products, because only the coefficients on the squared terms are statistically significant. That is, large deviations in either direction of units sold from its average predict nominal price changes. Overall, there is no pervasive and stable relationship between lagged units sold and nominal adjustments.

In contrast with the other two variables, the price's age lowers the probability of a price change for all of the categories. With two exceptions, the estimated coefficients multiplying the inverse age and its square are positive, and they are jointly statistically significant at the 1 percent level for all categories.[15] The exceptions are the coefficients on the squared inverse age for Peanut Butter and Tuna. These are both negative and less than half of the corresponding linear terms in magnitude, so the estimated hazard functions decline. Thus, the finding that recently set prices are more likely to change than old prices characterizes all of the product categories in this sample even after controlling for the relative price, lagged units sold, temporary sales, and heterogeneity across stores, products, and calendar dates.

A variable's statistical significance indicates that it has some forecasting value, but it does not show that it matters quantitatively. To assess each variable's contribution, Table 6 reports root mean-squared errors (in percentage points) from several specifications of the linear-in-probabilities model. The first column reports the in-sample *rmse*'s from forecasting price changes with only a constant. Unsurprisingly, these practically equal their maximum possible value, 50. The remaining columns report the *rmse*'s from models with progressively

[15] Apparently, the failure of the individual coefficients to be statistically significant for Ketchup reflects an inability to identify the rate of decline.

richer specifications for $x_{i,t}$. The second column corresponds to a model which includes only the sale price indicator and the dummy variables for the store, product, and calendar date. These variables lower the *rmse*'s from 3 to 5 percentage points. The third column gives the results from models that add the two relative price terms, and the fourth column has results from adding the terms in lagged units sold to that specification. Adding these variables lowers the *rmse*'s imperceptibly. For the final column, we added the two terms in the price's inverse age to the fourth column's specification. This yields the same definition of $x_{i,t}$ used in the original regression analysis. The *rmse*'s drop from 1.4 to 2.9 percentage points. In this sense, the price's age is the most quantitatively useful available forecaster of nominal adjustments.

6 Robustness

In this section, we document robustness to changes in measurement strategy. We begin by showing that the results hinge on neither the correction for time aggregation nor the specific definition of a sale price. Firms ultimately control stores' prices, so we also explore the consequences of using a firm's average price for an item (across its stores) as the unit of analysis.

Two measurement decisions permeate the results, the correction for time aggregation based on fractional prices and the definition of a sale. To examine the importance of the time aggregation correction, we recomputed every table and figure with the uncorrected original data.[16] Unsurprisingly, removing the time aggregation correction increases the frequency of price adjustment, from 30 to 35 percent. However, this modification leads to no other substantial changes. Given the scarcity of sale prices as we define them, one might speculate that our definition is too conservative. We investigated this by adopting an alternative which required a price to drop 5 percent (instead of 10 percent) and fully recover within

[16] As before, we first rounded all prices to the nearest whole cent.

four weeks (instead of two weeks). Here also, the *only* substantial change to the results was unsurprising: the frequency of sale prices rose from 2.8 to 7.2 percent.

Chains need not charge identical prices at their stores, but they frequently do so. To the extent that firms follow such a uniform pricing policy, a firm's price choice will impact both the numerator and denominator of our measure of a store's relative price. In this case, the relative prices at large firms will remain closer to one than those at small firms.

To examine whether this measurement issue impacts the results, we wish to repeat our analysis after aggregating the store-level observations into firms. Unfortunately, the ERIM scanner data does not indicate which firms own which stores, so this is not straightforward. Nevertheless some groups of stores seem to belong to a common firm, because their prices *exactly* equal each other more often than not. We use this to assign stores to individual firms with a simple rule. First, we compare the median transaction prices from the household data across all of the sample's stores. Second, we divide stores into firms to maximize the number of firms given the requirement that two stores must belong to a common firm if 50% or more of available price comparisons between them result in an exact match.

The rule divides the 42 sample stores into 9 firms in Sioux Falls and 6 firms in Springfield. The three largest firms in Sioux Falls have 6, 4, and 3 stores. The first and third firms own only groceries, while the second owns only drug stores.[17] The remaining firms in Sioux Falls each have a single store. The rule identifies four multiple-store grocers in Springfield, which have 9, 5, 4, and 3 stores. We cross-checked the results for Springfield using the Southwestern Bell Yellow Pages listings for grocers from 1987, and we gained additional information for some of the listed firms from the 1987 *Chain Store Guide Directory: Supermarkets and Grocery Chains*. With these, we identified four multiple-store grocers serving Springfield: Consumers Market (10 stores), Ramey Super Markets (7 stores), Dillon Food Stores (4 stores), and Smitty's (3 stores). Apparently, either our data is missing three stores from the largest two firms or assigns them to single-store firms. In either case, we believe that this

[17]Our simple rule did *not* use information about stores' types when grouping them into firms.

simple rule has grouped stores into firms with sufficient accuracy for the purpose of assessing robustness.

After assigning the stores to firms, we constructed firm-level average prices, corrected them for time aggregation, and selected a balanced panel of firm-item pairs that were always sold by at least five firms. This results in a somewhat smaller sample of 66 products. In a given week, half of the firm-level prices change. This increase is unsurprising given that a change will occur if the price changes at any store. The switch to firm-level prices changes only one other result meaningfully – the influence of the relative price on the probability of a price change plotted in Figure 6 becomes highly asymmetric. Relative prices substantially below one continue to lead to nominal price changes, but relative prices above one do not. Using firm-level prices does not substantially change the statistical significance of the linear-in-probability models' coefficients, nor does it alter the conclusion that the price's age is the most helpful predictor of nominal adjustment.

7 Conculsion

The ERIM scanner data reveal a counterintuitive time dependence: The longer a nominal price remains unchanged, the *less* likely it is to change. The declining hazard function we estimate with weekly data does not reflect heterogeneity across stores or products. Instead, it follows from a given price's frequency of adjustment changing over time. Sometimes the price appears to be rigid, and at other times it behaves flexibly.

Assessing the broader applicability of this result lies well beyond the scope of this paper, but the evidence summarized by Dhyne et al. (2004) suggests that it does not merely reflect idiosyncracies in the ERIM scanner data. In nearly every Euro zone country, the hazard function for nominal adjustment decreases when measured with monthly *CPI* data. Alvarez, Burriel, and Hernando (2005) attribute this result to heterogeneity across producers in the frequency of price adjustment, and they show that a model with just four groups of price setters fits Spanish *CPI* data well. Heterogeneity undoubtedly contributes substantially to the declining hazard functions measured with national samples of prices, but this paper's results lead us to wonder if the adjustment of Euro zone prices might also display true duration dependence.

The patterns of price changes we document do not arise easily from the theory of optimal price adjustment. This leads us to consider other possibilities. One heuristic description of our results is that they reflect a sort of learning. A producer unsure of the profit-maximizing price experiments with several and eventually settles on one. This price remains in place until the available evidence indicates that it might be substantially improved upon. Rothschild (1974) examines optimal learning of noisily observed demand by a Bayesian monopolist who can choose one of two prices. The monopolist begins by experimenting with both prices, but he eventually settles on one of them permanently. Extending Rothschild's model seems to be a promising approach to account for this paper's results.

References

Álvarez, Luis J., Pablo Burriel, and Ignacio Hernando. 2005. Do Decreasing Hazard Functions for Price Changes Make Any Sense? Documentos de Trabajo 0508, Banco de España.

Baharad, Eyal and Benjamin Eden. 2004. Price rigidity and price dispersion: evidence from micro data. *Review of Economic Dynamics*, **7**(3), 613–641.

Barro, Robert J. 1972. A Theory of Monopolistic Price Adjustment. *Review of Economic Studies*, **34**, 205–272.

Bils, Mark and Peter J. Klenow. 2004. Some Evidence on the Importance of Sticky Prices. *Journal of Political Economy*, **112**(5), 947–985.

Calvo, Guillermo A. 1983. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, **12**(3), 383–398.

Caplin, Andrew S. and John Leahy. 1991. State Dependent Pricing and the Dynamics of Money and Output. *Quarterly Journal of Economics*, **106**, 683–708.

Caplin, Andrew S. and Daniel F. Spulber. 1987. Menu Costs and the Neutrality of Money. *Quarterly Journal of Economics*, **102**, 703–726.

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

Darby, Michael R., John Haltiwanger, and Mark Plant. 1985. Unemployment Rate Dynamics and Persistent Unemployment under Rational Expectations. *American Economic Review*, **75**(4), 614–637.

Dutta, Shantanu, Mark Bergen, and Daniel Levy. 2002. Price Flexibility in Channels of Distribution: Evidence from Scanner Data. *Journal of Economic Dynamics and Control*, **26** 1845–1900.

Dhyne, Emmanuel, Luis J. Álvarez, Hervé Le Bihan, Giovanni Veronese, Daniel Dias, Johannes Hoffman, Nicole Jonker, Patrick Lünnemann, Fabio Rumier, and Jouko Vilmunen. 2004. Price setting in the euro area: Some stylized facts from Individual Consumer Price Data. Unpublished Working Paper, European Central Bank.

Dotsey, Michael, Robert G. King, and Alexander L. Wolman. 1999. State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output. *Quarterly Journal of Economics*, **114**(2), 655-690.

Golosov, Michael and Robert E. Lucas, Jr. 2003. Menu Costs and Philips Curves. NBER Working Paper # 10187, December 2003.

Kashyap, Anil K. 1995. Sticky Prices: New Evidence from Retail Catalogs. *Quarterly Journal of Economics*, **110**(1), 245–274.

Lach, Saul and Daniel Tsiddon. 1992. The Behavior of Prices and Inflation: An Empirical Analysis of Disaggregated Price Data. *Journal of Political Economy*, **100**(2), 349–389.

Klenow, Peter J. and Oleksiy Kryvtsov. 2005. State-Dependent or Time-Dependent Pricing: Does it Matter for Recent U.S. Inflation? NBER Working Paper #11043.

Konieczny, Jerzy D. and Andrzej Skrzpacz. 2005. Inflation and price setting in a Natural Experiment. *Journal of Monetary Economics*, **52**(3), 621–632.

Richardson, David H. 2003. Scanner Indexes for the Consumer Price Index. in *Scanner Data and Price Indexes*, Robert C. Feenstra and Matthew D. Shapiro, eds. University of Chicago Press.

Rothschild, Michael. 1974. A Two-Armed Bandit Theory of Market Pricing. *Journal of Economic Theory*, **9**, pp. 185–202.

Sheshinski, Eitan and Yoram Weiss. 1977. Inflation and Costs of Price Adjustment. *Review of Economic Studies*, **44**, 287–303.

Taylor, John B. 1980. Aggregate Dynamics and Staggered Contracts. *Journal of Political Economy*, **88**(1), 1–23.

Table 1: Summary Statistics

Category	Number of UPCs	Number of Observed Prices	Frequency of Sale Prices	Annualized Rate of Price Change
All Products	94	83394	2.8	-3.8
Ketchup	3	1353	3.9	-2.5
Margarine	54	57687	2	-3.8
Peanut Butter	3	861	1.4	7.8
Sugar	8	3321	5.4	-2.7
Tissue	8	5412	4.1	-4.4
Tuna	18	14760	4.9	-4.6

Table 2: Standard Deviations of Log Prices[(i),(ii)]

				{UPC, Market, Date} ⋃	
Category	{UPC}	{UPC, Market}	{UPC, Market, Date}	{UPC, Store}	{UPC, Store, Sale Price}
All Products	13.7	13.2	9.2	8.2	7.7
Ketchup	14.5	13.1	8.1	7.5	7.2
Margarine	13	12.8	8.8	7.7	7.3
Peanut Butter	11.4	11.4	5.2	4.7	4.6
Sugar	11.8	11.6	4.3	4.3	3.9
Tissue	13.6	12.7	6.9	6.3	6
Tuna	16.4	15.5	11.9	11	10.2

Notes: (i) The table reports residual standard deviations in percentage points from the regressions of the price's logarithm against the given set of regressors. (ii) In the table, {X, Y} indicates a set of dummy variables that span the unique combinations of the values of X and Y.

Table 3: The Frequency of Price Changes(i)

Category	Weekly Observations: All Observations	Weekly Observations: No Sales(ii)	Monthly Observations(iii)	Bils-Klenow(iv)
All Products	30	26	47	26
Ketchup	34	28	54	20
Margarine	30	27	46	28
Peanut Butter	36	34	54	31
Sugar	32	22	43	23
Tissue	30	24	50	24
Tuna	31	23	48	27

Notes (i) The table's entries are frequencies expressed in percentage points. (ii) Weeks during and immediately following sales are excluded from the calculations in this column. (iii) Monthly observations are constructed by using the price of each store-upc pair in the first week of each calendar month. (iv) The entries in this column are the frequencies of price changes reported in Table 1 of Bils and Klenow (2004) for the respective categories. See Footnote 3 in the text for more information regarding the mapping of the ERIM categories into BLS categories. The first frequency in this column is the simple average of those for the six reported categories.

Table 4: Estimates of Average Price Durations[(i)]

	Average Durations[(ii)]		Inverse Frequency Estimates	
Category	All Prices	New Prices	Inverse of Average[(iii)]	Average of Inverses[(iv)]
All Products	5.9	3.2	3.3	3.7
Ketchup	4.8	2.8	2.9	3.2
Margarine	6.1	3.3	3.3	3.8
Peanut Butter	2.8	2.5	2.8	2.9
Sugar	5.4	3.1	3.2	3.5
Tissue	4.7	3.1	3.3	3.8
Tuna	6.4	3	3.2	3.7

Note: (i) Table entries measured in weeks. (ii) These columns report the average duration of prices charged in the sample's first 85 weeks. (iii) This column reports the inverses of the weekly frequencies from the first column of Table 3. (iv) This column reports the average of inverse price frequencies calculated for each store-UPC combination. See the text for further details.

Table 5: Linear-in-Probabilities Estimates[i]

	Relative Price		Lagged Units Sold		Inverse Age			
Category	Original	Squared	Original	Squared	Original	Squared	Sale Indicator	R^2
All Products	16.40***	37.10***	0.80***	1.20***	23.70***	16.50***	36.50***	24.5
	(1.60)	(5.60)	(0.30)	(0.20)	(2.20)	(1.90)	(0.90)	
Ketchup	-34.60***	85.00**	-3.60*	0.20	29.50	5.20	33.40***	27.2
	(13.20)	(37.80)	(2.00)	(1.40)	(19.90)	(17.10)	(6.10)	
Margarine	15.70***	29.40***	0.40	1.40***	17.50***	23.70***	34.10***	25.4
	(2.10)	(7.20)	(0.40)	(0.30)	(2.60)	(2.30)	(1.20)	
Peanut Butter	-18.90	141.00	-3.30	3.00	68.60***	-24.80	12.90	39.8
	(25.60)	(100.80)	(4.50)	(3.60)	(24.00)	(21.00)	(20.00)	
Sugar	55.70***	136.60***	0.10	4.70***	7.20	22.20**	36.60***	36.9
	(6.30)	(12.40)	(1.50)	(1.00)	(11.70)	(10.20)	(4.10)	
Tissue	5.10	75.10***	0.50	2.10***	9.10	20.70***	39.60***	28.2
	(6.80)	(24.00)	(1.00)	(0.50)	(8.90)	(7.70)	(2.90)	
Tuna	4.90	30.70***	0.50	0.60**	44.30***	-4.50	39.50***	27.6
	(3.10)	(5.50)	(0.50)	(0.30)	(5.20)	(4.60)	(1.70)	

Note: (i) Each column reports estimated coefficients (in percentage points) multiplying the indicated variable. Heteroskedasticity-consistent standard errors are below each coefficient in parentheses. The superscripts $\star$, $\star\star$, and $\star\star\star$ indicate statistical significance at the 10, 5, and 1 percent levels. See the text for further details.

Table 6: Root Mean-Squared Errors(i)

		Store, UPC, Date, & Sale Indicators plus Quadratic in			
Category	No Regressors	No Others	Relative Price	& Units Sold(ii)	& Inverse Age(iii)
All Products	46.6	43.4	43.2	43.1	40.6
Ketchup	47.9	45.1	44.9	44.9	43.2
Margarine	46.4	43.3	43.1	43.1	40.2
Peanut Butter	47.9	42.5	42.5	42.5	40.5
Sugar	47.2	41.5	40.2	39.8	38.4
Tissue	46.2	41.5	41.3	41.2	39.8
Tuna	47	42.6	42.4	42.3	40.2

Notes: (i) The table's entries are in-sample root mean-squared errors (in percentage points) from forecasts of $I\{p_{i,t} \neq p_{i,t-1}\}$ based on linear-in-probabilities models that include the specified set of regressors. The maximum possible value for these is 50. (ii) The linear-in-probabilities models underlying these results include quadratic terms in the relative price and units sold. (iii) The linear-in-probabilities models underlying these results include quadratic terms in the relative price, units sold, and the inverse duration. See the text for further details.

Figure 1: Time Aggregation and the Measurement of Price Changes

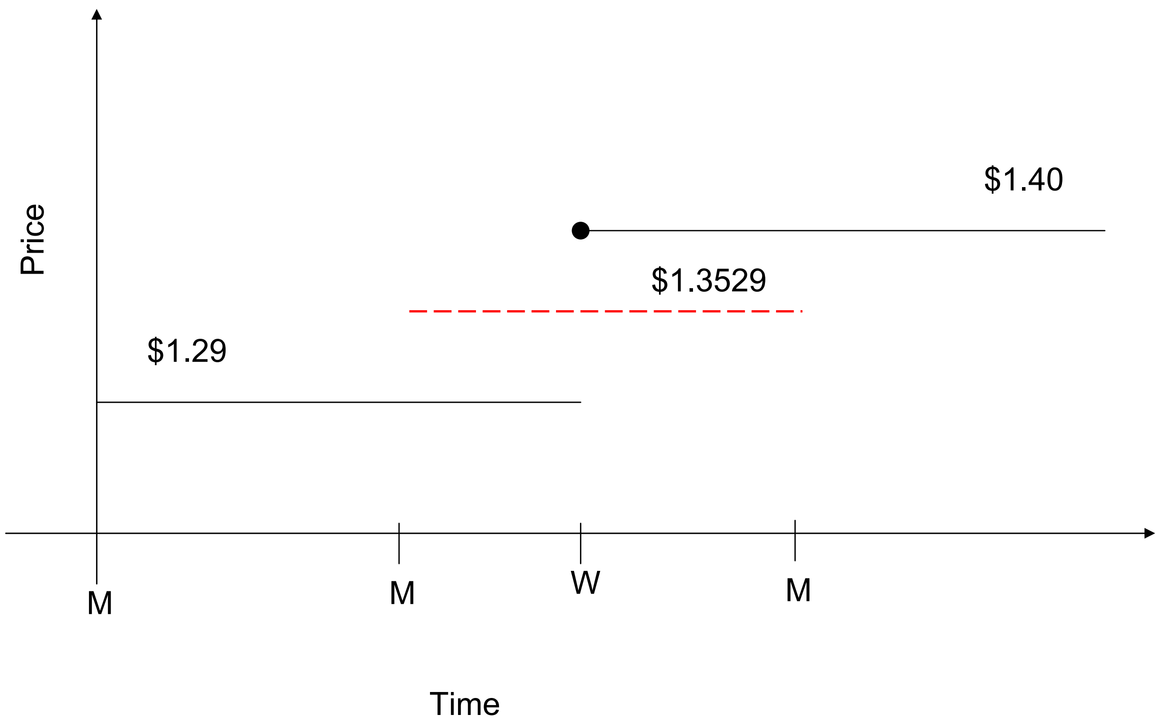
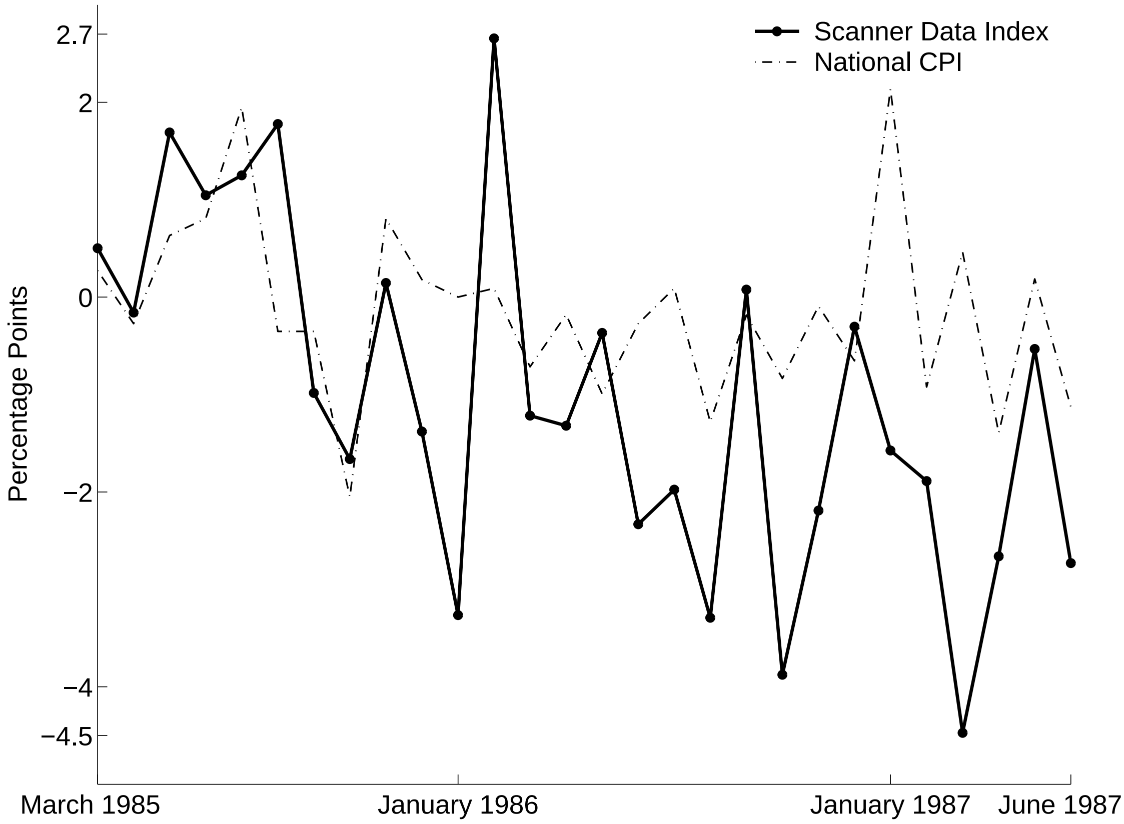
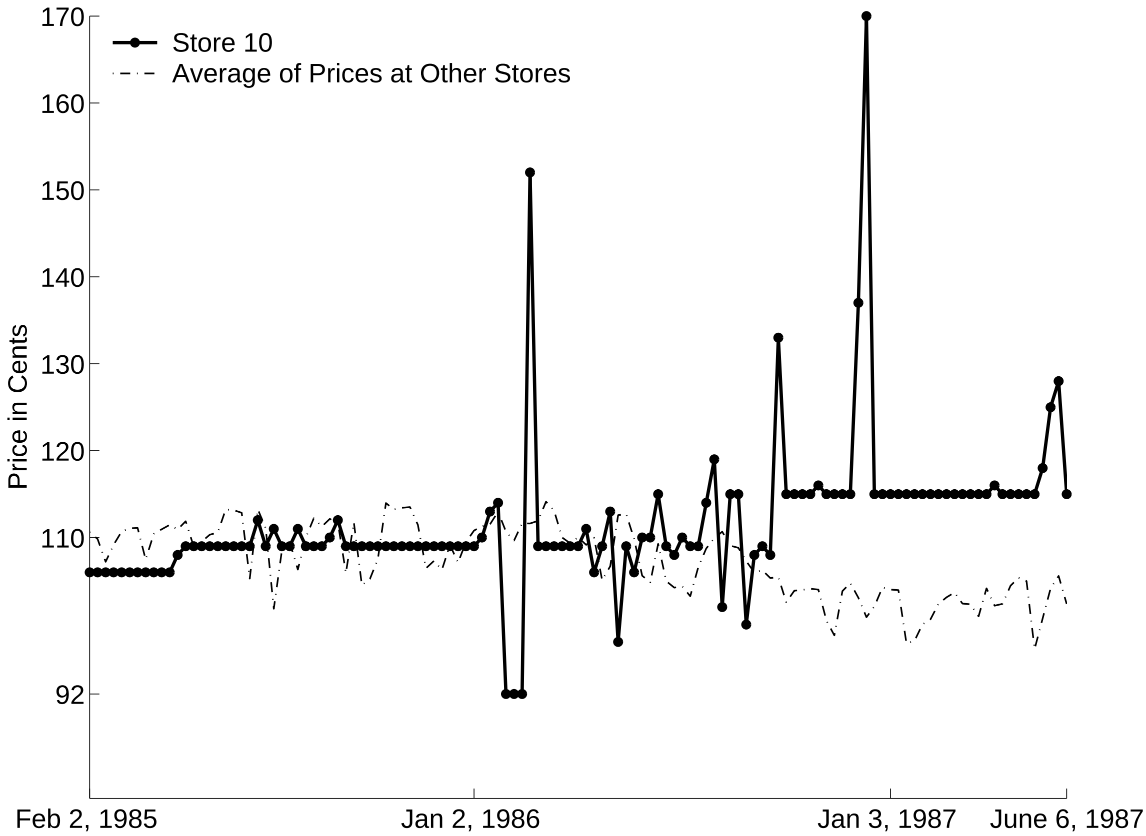


Figure 2: Inflation Rates for Margarine[i]



Note: (i) Annualized monthly inflation rates.

Figure 3: The Price of Fleischmann's Margarine(i)



Note: (i) Weekly observations of the price of Fleischmann's Margarine at a store in Sioux Falls, SD and the average of all other stores' prices for the identical product. Dates are the final days of the given week. See the text for further details.

Figure 4: Sample Hazard Function for Price Changes

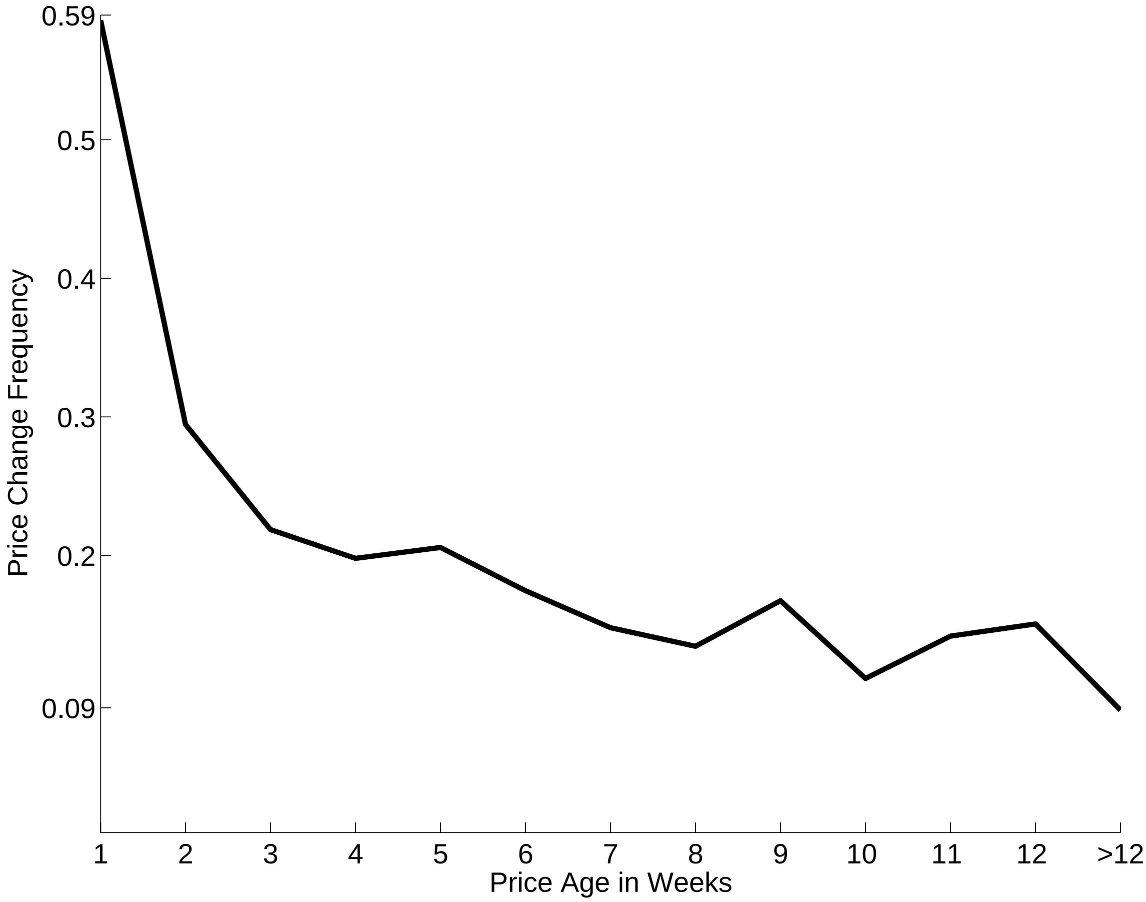
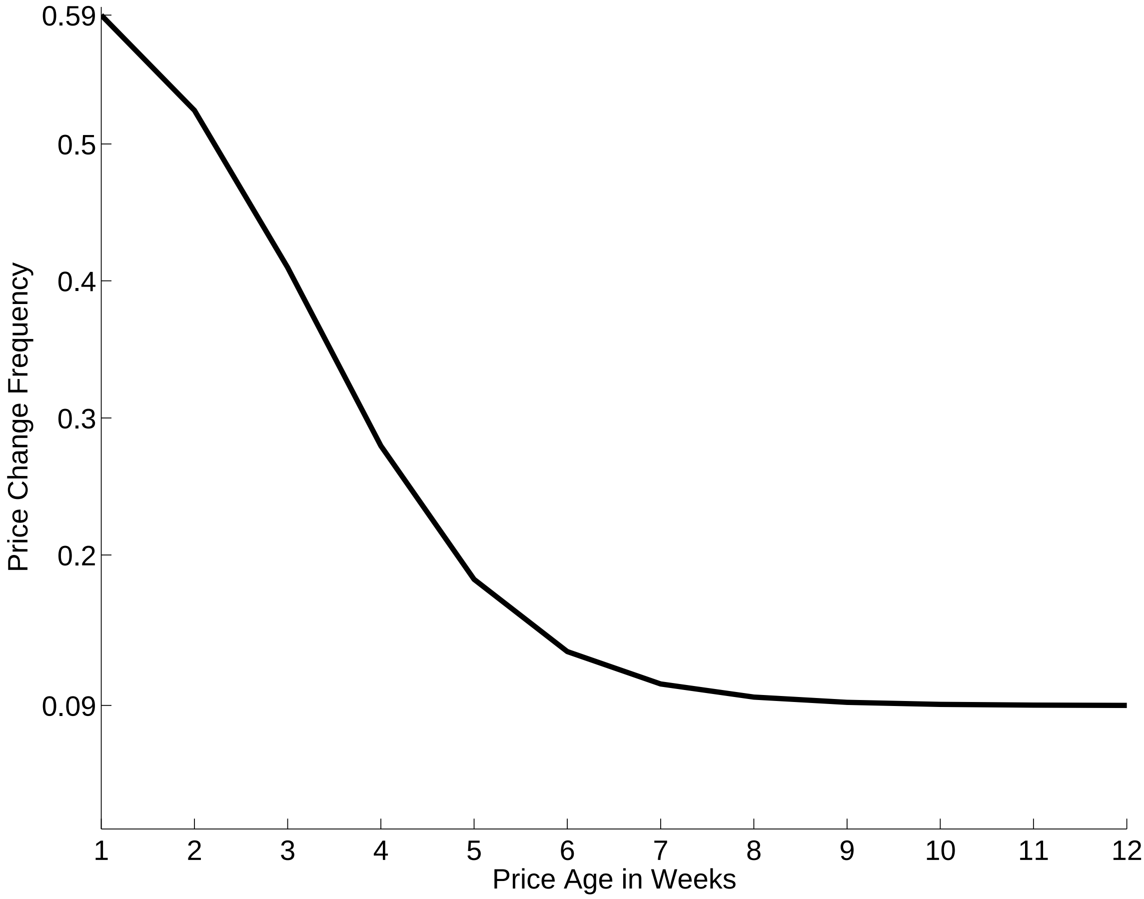
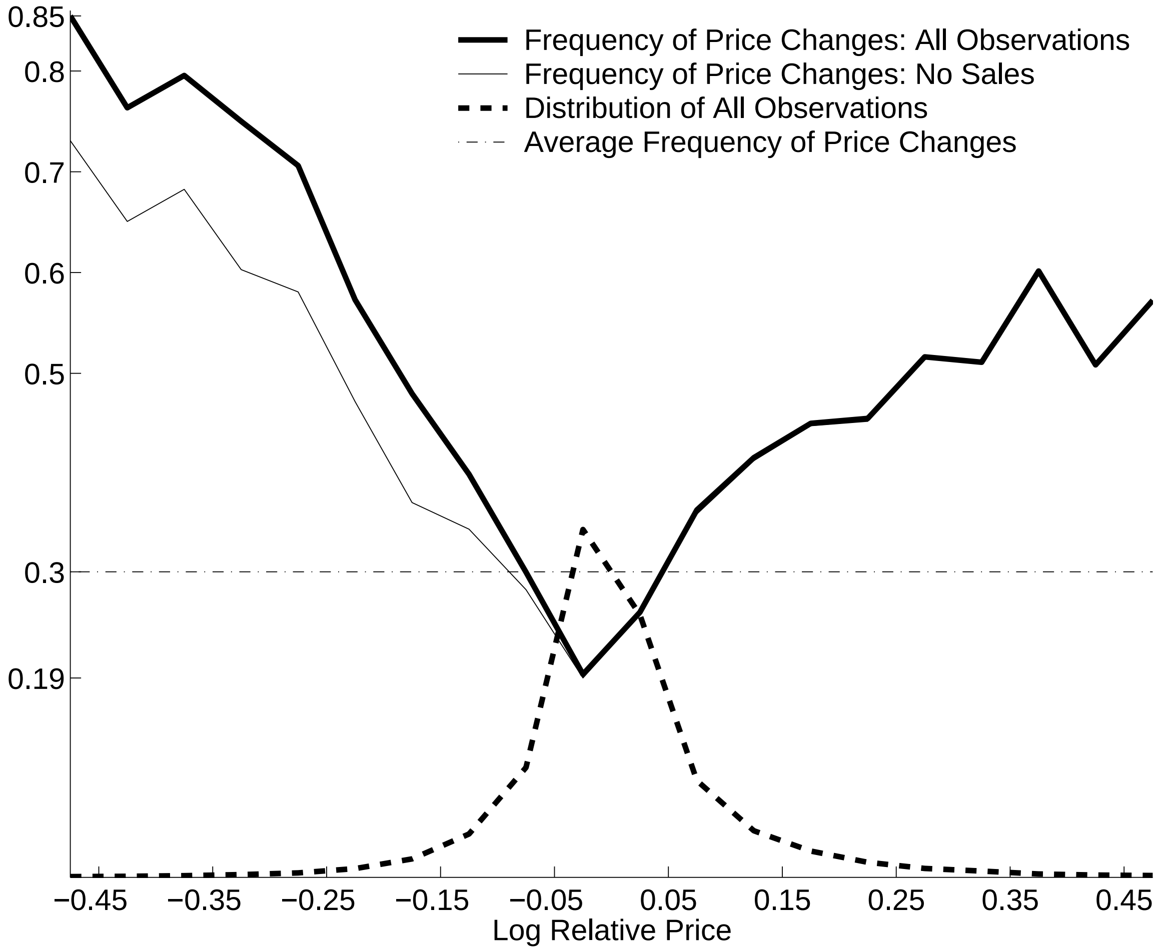


Figure 5: Hazard for Price Changes with Rigid and Flexible Prices[(i)]



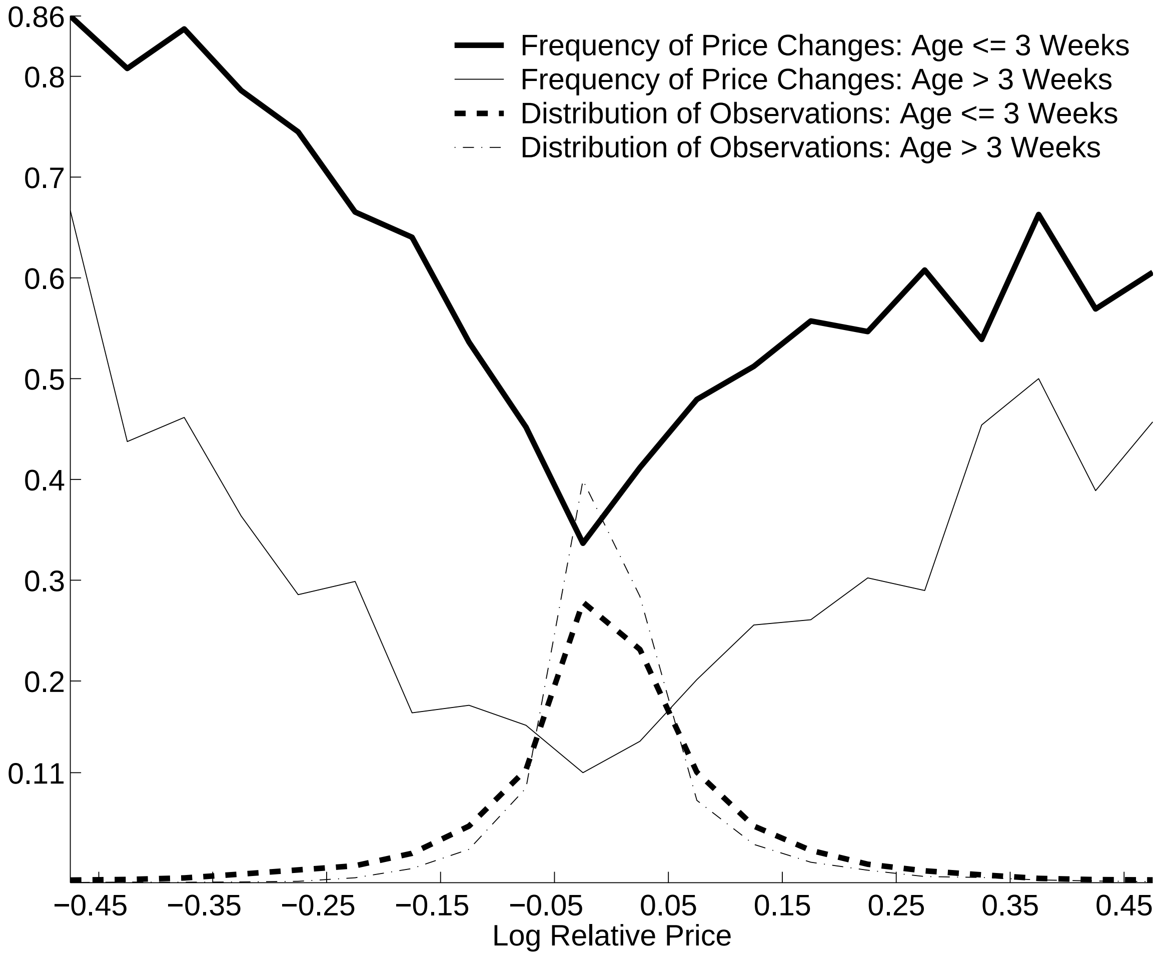
Note: (i) Hazard function from a simple model with flexible and rigid prices. See the text for further details.

Figure 6: Hazard for Price Changes as a Function of the Relative Price[(i)]



Note: (i) The relative price equals the store's price in the previous week divided by the average of all other prices for the same UPC charged in the current week. See the text for further details.

Figure 7: Young and Old Prices' Hazards as Functions of the Relative Price(i)



Note: (i) The relative price equals the store's price in the previous week divided by the average of all other prices for the same UPC charged in the current week. The calculations exclude observations from the initial (left-censored) price spells. See the text for further details.