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

Jeffrey R. Campbell† Benjamin Eden‡

October, 2006

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 heterogeneity in price durations both between and within store-product cells. 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.

Twenty-one percent of the prices in these data change in an average week. If the probability of a nominal adjustment remains the same as a price ages, then this implies that the average price lasts about five weeks. However, a randomly chosen price remains unchanged for more than nine 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 in part 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 not very important quantitatively in our sample, because most price changes occur with prices 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 most extreme prices are relatively young (less than a month old).

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 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 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. 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

The balanced panel so created has $83,394$ prices. The division of revenues by units sold yielded $21,281$ prices which cannot be expressed in whole cents (e.g. \$1.3529). Such a fractional price could arise either from technical mistakes in price setting or from time aggregation. Technical errors occur when the price displayed in the store differs from that in the computer. For example, suppose that a store manager decided to change the price on Monday from \$1.29 to \$1.40. The price must be changed both on the computer and on the shelf. Suppose that erroneously it changes only on the computer. Then those customers who notice the lower price on the shelf will complain and receive the item for \$1.29, while others who are less attentive will pay \$1.40. At some point, the store will correct the shelf price, but the earlier mixing can nevertheless result in an average price of \$1.3529.

A fractional price can also arise from time aggregation when a store manager changes an item's price during the middle of a week. To illustrate this possibility, change the previous example to suppose that the manager changes the price from \$1.29 to \$1.40 successfully on a Wednesday. Those customers buying the good on Monday or Tuesday would pay \$1.29 while those buying later would pay \$1.40. Just as before, the average revenue price changes twice in these three weeks, while the daily price in the computer changes only once. The spurious price change arises because the second week's average revenue embodies two prices.

To address these problems we replaced fractional prices by the median price in the individual purchase history data whenever this was possible. In our example, if we observe three consumers purchasing the item for \$1.29, \$1.29, and \$1.40 we change the second week's price to \$1.29. This was done for 15,346 prices out of the 21,365 fractional prices in the data. For the remaining prices that we could not replace we checked whether they are part of a descending or increasing sequence of prices (as in the example). If the fractional price was a part of a monotone sequence we concluded that it is likely to be the result of time aggregation and replaced the fractional price by the following week's price. 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. We changed 2,934 prices in this manner. 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.

Figure 2 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.[2] This price changed 29 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, and remains roughly constant for the rest of the year. The average of other firms' prices approximately equaled \$1.10 throughout these eleven months. The price changed much more often in 1986. After two very modest price increases in January, the price dropped to \$0.92 for four weeks and then returned to \$1.09.[3] Throughout these changes, the average price at other stores fluctuated little. The return to \$1.09 lasted

[2]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.

[3]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.

only seven 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 a dramatic temporary price increase. In 1987, the price returned to a pattern of much less frequent price changes. It ended the sample period at \$1.15.

The Figure shows three price increases that last exactly one week. Those are fractional prices that were not corrected because there were no purchase history data to replace them and they are not within a monotone sequence of prices. However we suspect that these observations might not be real price changes. If these are indeed "mistakes" they would affect our calculations of the weekly frequency which will change from $29/123 = 0.24$ to $23/123 = 0.19$. We thought of smoothing all price increases that last one week only and revert to exactly the previous week's price, but this increases the danger of imposing to much of our prior on the data: After all a manager may decide to increase the price and then realize that this leads to a loss of revenues and as a result revert to the previous price after a week. However, we need to keep in mind the possibility that one-week price increases are not actual price increases and to check for robustness of our results to counting such "mistakes" as actual price changes.

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.

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 3.9 percent of the observations are sale prices. Sugar's frequency of sale prices, 8.3 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 1 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 0.37, 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 11.9 percent for Peanut Butter to 17.4 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 1 suggests, removing date-specific means substantially lowers variation. For example, margarine's standard deviation drops from 13.4 percent to 9.2 percent. Because $(9.2/13/4)^2$ 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.[4] In fact, adding store-specific UPC dummies lowers the standard deviations at most 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 somewhat, but overall Table 2indicates that substantial price dispersion remains even after controlling for heterogeneity across markets, stores, products, and weeks.

Next, 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, 21 percent of prices change in a given week. The second column

[4]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.

reports the frequencies after excluding prices changes to or from a sale price. A little under 1/4 of all price changes end a sale price. 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 (2004).[5] 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 60 percent of prices change in a given month. However, we find instead that approximately 37 percent of prices change when sampled monthly. This suggests that the Calvo assumption of a constant weekly probability of price adjustment does not hold good in our data. In any case, the prices in our data appear to be somewhat more flexible than those in the BLS sample.

Table 4 provides direct and inverse-frequency type estimates of the average price duration. The first two columns report direct measures using average realized durations of prices charged before the sample's last year.[6] 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 9.3 weeks for all products. Across categories, it varies from 5 weeks for Peanut Butter to 9.6 weeks for margarine. The second column reports the average durations of the sample's newly-set prices. This is about half of the average duration for all prices.

The last two columns of Table 4 provides estimates based on the inverse of the price-change frequency. The first of the two is the standard estimate. This is very close to the average duration for newly set prices. The last column computes the inverse of the frequency

[5] 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."

[6] We define a newly-set price as one week old.

for each product/store cell and then averages over cells. Baharad and Eden (2004) use Jensen's inequality to argue that if the frequency of price change is cell specific and there is no within cell heterogeneity then the correct measure is the last column. Since the average of the inverses is less than the direct measure (6.9 rather than 9.3) we conclude that there must be some within cell heterogeneity.

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. 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 3 plots the unconditional hazard function estimated using the observations from all product categories. The chance of a newly-set price changing equals 0.40. As the price ages, this probability drops precipitously. It equals 0.34 for a two-week-old price and 0.21 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 0.08. The unreported hazard functions calculated separately for each product category resemble Figure 3. If transitory measurement errors (like those we suspect in Figure 2) infect our data, then they can make the measured hazard function drop from 1 week (the age of a newly-set price) to 2 weeks even if the true hazard function is flat. Such measurement errors slightly raise the hazard function's level for older prices, but they leave its shape unchanged. That is, imposing measurement error on a process with a constant hazard produces a decline only from 1 to 2 weeks. In fact, the drop from 2 to 3 weeks is larger than that from 1 to 2 weeks; so measurement error alone cannot rationalize the estimated hazard function's shape.

To place Figure 3 into context, it is helpful to recall that the estimated hazard function for a worker leaving unemployment typically decreases: 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 just like the average duration for all prices exceeds in our sample the average duration for newly set prices (see Table 3).

Heterogeneity provides a simple explanation for the decreasing hazard function. We illustrate this here with an example adapted from Darby, Haltiwanger, and Plant (1985). 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. 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 3) and asymptotes to the probability of a rigid price changing. Figure 4 plots the implied raw hazard function from this example when the probability of flexible and rigid prices changing are 0.65 and 0.08 and 60 percent of all newly-set prices are flexible. This simple example reproduces Figure 3 well.

Heterogeneity in price duration is consistent with standard model if it reflects differences between product/store cells. But it is not consistent with standard model if instead there is within-cell heterogeneity: a given store set sometimes a rigid price and sometimes a flexible price for the same product. Table **??** hints that within cell heterogeneity is potentially important in our data, because the average of the inverses was significantly shorter than the directly measured duration. To check whether within cell heterogeneity is important for the decreasing hazard we calculated the hazard within each product/store cell for young prices (with ages less than or equal to 3 weeks) and old prices (with ages greater than 3 weeks). Whenever the hazard for young prices was higher than the hazard for old prices we said that it is declining. Figure 5 displays the fraction of cells that exhibit decreasing hazard function. Overall, 87% of cells have a declining hazard. To testing the null hypothesis of a constant hazard against the alternative of a declining hazard, we construct a standard t-test for each cell. About 27% of cells have slopes that are negative and statistically significant at the 5% level. Only 1.5% of the cells have positive and statistically significant slopes. We also consider the case in which the price setting unit is the store rather than the product/store combination. This will occur for example, if the manager chooses to change the price of all the items in the store at the same time. We find that all stores exhibit statistically-significant decreasing hazard functions.

In summary, these data display a counterintuitive form of time dependence in price setting. Older prices are *less* likely to change than newly-set prices. This does not merely reflect heterogeneity across stores or products in the frequency of price setting. Instead, hazard functions calculated using only one store's or store-product cell's observations typically decrease with the price's age.

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 stores' nominal adjustments. In this section, we examine this possibility. We begin with an indicator of the benefit of a nominal adjustment, 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.[7]

Figure 5 plots the observed frequency of price changes as a function of the relative price's

[7] 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.

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. 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 5 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, 0.21; and the dashed line plots the sample's distribution of relative prices.

There are four notable features of Figure 5. First, the minimum frequency substantially exceeds zero. For both samples, it approximately equals 0.12. 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 62 percent when the price is 35 to 40 percent below average and 53 percent when it is 35 to 40 percent 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.[8]

In light of the negative association of a price's age with the probability that it changes,

[8]Category-specific versions of Figure 5 all display the same features, but the estimated frequencies are considerably noisier.

Figure 6 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 6 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 16.5 percent and 8.8 percent for young and old prices.

The finding that young prices are more dispersed than old prices is not consistent with standard models which often assume that all store that change their price change it to the same price. To examine this surprising result further we plot in Figure 7 the fraction of young prices as a function of the relative price. Because we are concerned that some extreme relative prices might be errors like those we suspect in Figure 2, we exclude one-week old prices from the analysis. The horizontal line plots the overall fraction of young prices, 28%, for reference. We see that the minimum of 21% occurs when the relative price is close to zero and the fraction of young prices increases with the absolute value of the relative price. Eighty percent of the prices between 25 and 30 percent below average are young, and 40 percent of the corresponding above-average prices are young. The results are similar if we include one-week-old prices in the analysis. In a standard model, a fixed nominal price becomes an extreme as inflation erodes it. In these data, stores set extreme relative prices in the recent past.

5 Forecasting Price Changes

Figure 6 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. 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.[9] 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 6 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 5 suggests, the regression function is convex in the relative price. The coefficients multiplying linear and squared terms in lagged units sold are both negative. The coefficient multiplying the squared inverse price age is negative but less than half the magnitude of that multiplying

[9]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.

the corresponding linear term. Thus, the regression function strictly decreases in the price's age.

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 Ketchup, 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 5 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, Tissue, and Tuna and at the ten percent level for Peanut Butter. The estimates for Margarine are unsurprisingly similar to those from pooling all products. That is, exceptionally large sales tend to reduce the probability of a price change. The intuition from Golosov and Lucas (2003) is the opposite: A firm selling a large number of units facing a given menu cost adjusts price more frequently than a rival selling fewer units. The other two categories do not qualitatively resemble these estimates, 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.

–

Finally, consider the coefficients multiplying the inverse of the price's age. The coefficients on the linear term are all positive and (with the exception of Ketchup) statistically significant at the one percent level. Those multiplying the squared term are negative (again with the exception of Ketchup) and statistically significant. The two coefficients are jointly statistically significant at the 1 percent level for all categories. Apparently, the failure of the individual coefficients to be statistically significant for Ketchup reflects an inability to identify the rate of decline. Figure 8 plot the estimated hazard function using p = a + bx +

cx2, where x is the inverse of the age, the coefficients b and c are taken from Table – and a is chosen so that the hazard for age = 1 will be 0.4 as in Figure –. As we can see the estimated hazard for all products, margarine, Ketchup and Tissue are decreasing in age. The estimates hazard for Sugar, Peanut Butter and Tuna increases slightly when moving from age = 1 to age = 2 and then decline with age. –

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 7 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 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

[UNDER REVISION]

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 *UPC*s	Number of Observed Prices	Frequency of Sale Prices	Annualized Rate of Price Change
All Products	94	83394	3.9	-4.1
Ketchup	3	1353	5	-2.8
Margarine	54	57687	2.7	-4
Peanut Butter	3	861	2.8	7.5
Sugar	8	3321	8.3	-6.7
Tissue	8	5412	6.4	-4.4
Tuna	18	14760	6.8	-4.7

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	14.4	14	9.7	8.6	8
Ketchup	14.4	13.2	8.4	8	7.6
Margarine	13.7	13.4	9.2	8.1	7.5
Peanut Butter	11.9	11.9	5.7	5.1	5
Sugar	15	14.9	6.7	6.7	5.8
Tissue	13.8	12.9	7	6.6	6.2
Tuna	17.4	16.4	12.7	11.5	10.6

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	21	16	37	26
Ketchup	22	16	46	20
Margarine	20	16	35	28
Peanut Butter	24	20	44	31
Sugar	23	12	42	23
Tissue	21	14	44	24
Tuna	24	16	38	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 5 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	9.3	4.7	4.8	6.9
Ketchup	7.5	4.4	4.6	5.6
Margarine	9.6	5.1	5	7.3
Peanut Butter	5	3.8	4.3	4.4
Sugar	9.2	4.2	4.3	5.9
Tissue	7.7	4.6	4.7	6.8
Tuna	8.9	3.9	4.2	5.7

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: Slopes of Within-Cell Hazard Function[i]

	All Observations, Percent			Excluding Sales, Percent		
Category	< 0	Significantly < 0	Significantly > 0	< 0	Significantly < 0	Significantly > 0
			Within-Store/Product Cells			
All Products	87.02	26.70	1.47	80.09	21.68	1.77
Ketchup	72.73	27.27	0.00	63.64	18.18	0.00
Margarine	85.93	26.23	1.92	82.52	21.32	2.13
Peanut Butter	71.43	42.86	0.00	71.43	42.86	0.00
Sugar	88.89	11.11	0.00	66.67	0.00	0.00
Tissue	90.91	11.36	0.00	65.91	9.09	0.00
Tuna	91.67	36.67	0.83	80.83	31.67	1.67
			Within-Store Cells			
All Products	100.00	100.00	0.00	100.00	100.00	0.00
Ketchup	70.00	30.00	0.00	60.00	20.00	0.00
Margarine	100.00	100.00	0.00	100.00	100.00	0.00
Peanut Butter	100.00	100.00	0.00	100.00	100.00	0.00
Sugar	94.44	33.33	0.00	66.67	0.00	0.00
Tissue	100.00	31.82	0.00	72.73	18.18	4.55
Tuna	96.15	73.08	3.85	88.46	61.54	3.85

Note: (i) The table's first column gives the percentage of cells with decreasing sample hazard functions, the second and third give the percentages with sample hazard functions with statistically significant negative and positive slopes. The significance level used was 5%. The remaining columns mimic the first three for the sample excluding sale prices. The first panel defines "cell" by product and store, and the second panel defines "cell" by store only.

Table 6: 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	17.00⋆⋆⋆	52.80⋆⋆⋆	-0.50⋆	-1.10⋆⋆⋆	52.90⋆⋆⋆	-25.60⋆⋆⋆	33.30⋆⋆⋆	15.3
	(1.80)	(5.10)	(0.30)	(0.20)	(2.00)	(1.80)	(1.00)	
Ketchup	-13.10	87.70	-1.70	0.50	9.50	6.50	30.20⋆⋆⋆	24.4
	(14.90)	(69.40)	(2.20)	(1.50)	(17.10)	(15.30)	(7.30)	
Margarine	17.00⋆⋆⋆	43.60⋆⋆⋆	-1.00⋆⋆⋆	-0.50⋆	40.30⋆⋆⋆	-11.80⋆⋆⋆	32.90⋆⋆⋆	16.1
	(2.50)	(6.10)	(0.30)	(0.30)	(2.30)	(2.20)	(1.40)	
Peanut Butter	-45.90⋆⋆	415.50⋆⋆⋆	-5.80	-3.90	107.40⋆⋆⋆	-76.50⋆⋆⋆	21.60⋆	36.4
	(21.30)	(69.50)	(3.90)	(2.60)	(22.20)	(19.90)	(13.00)	
Sugar	46.70⋆⋆⋆	146.20⋆⋆⋆	2.60⋆	-0.90	41.70⋆⋆⋆	-28.70⋆⋆⋆	40.00⋆⋆⋆	34.3
	(6.50)	(12.80)	(1.40)	(1.10)	(10.20)	(8.90)	(4.70)	
Tissue	-8.80	141.50⋆⋆⋆	-1.20	-1.80⋆⋆⋆	25.70⋆⋆⋆	-13.50⋆	35.30⋆⋆⋆	21.5
	(6.40)	(20.30)	(0.90)	(0.50)	(7.90)	(7.00)	(3.10)	
Tuna	9.90⋆⋆⋆	46.60⋆⋆⋆	-0.50	-1.60⋆⋆⋆	98.40⋆⋆⋆	-70.20⋆⋆⋆	30.90⋆⋆⋆	18.9
	(3.20)	(6.80)	(0.50)	(0.30)	(5.10)	(4.50)	(1.80)	

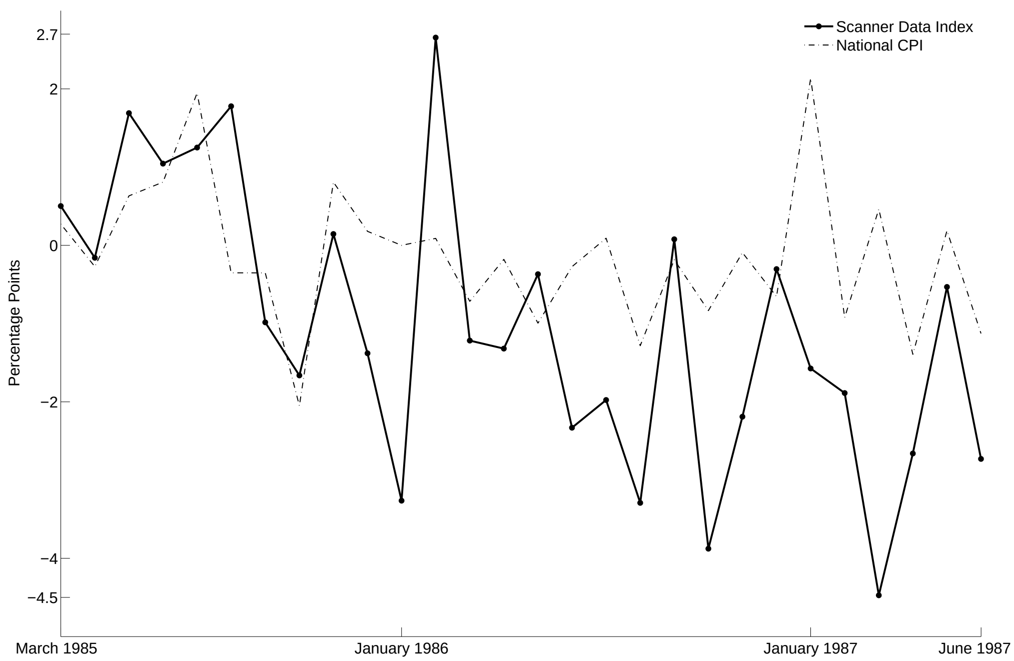
Note: (i) Each column reports estimated coefficients (in percentage points) multiplying the indicated variable. Heteroskedasticity-consistent standard errors are below each coefficient in parentheses. The superscripts ⋆, ⋆⋆, and ⋆ ⋆ ⋆ indicate statistical significance at the 10, 5, and 1 percent levels.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

Table 7: 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	41.5	39.4	39.2	39.2	38.3
Ketchup	42	39.1	39	39	38.7
Margarine	40.8	38.7	38.5	38.5	37.5
Peanut Butter	42.1	38.5	37.8	37.7	36.6
Sugar	43.3	37.9	36.2	36.2	36
Tissue	41	37.2	37	37	36.9
Tuna	43.6	40.6	40.4	40.4	39.5

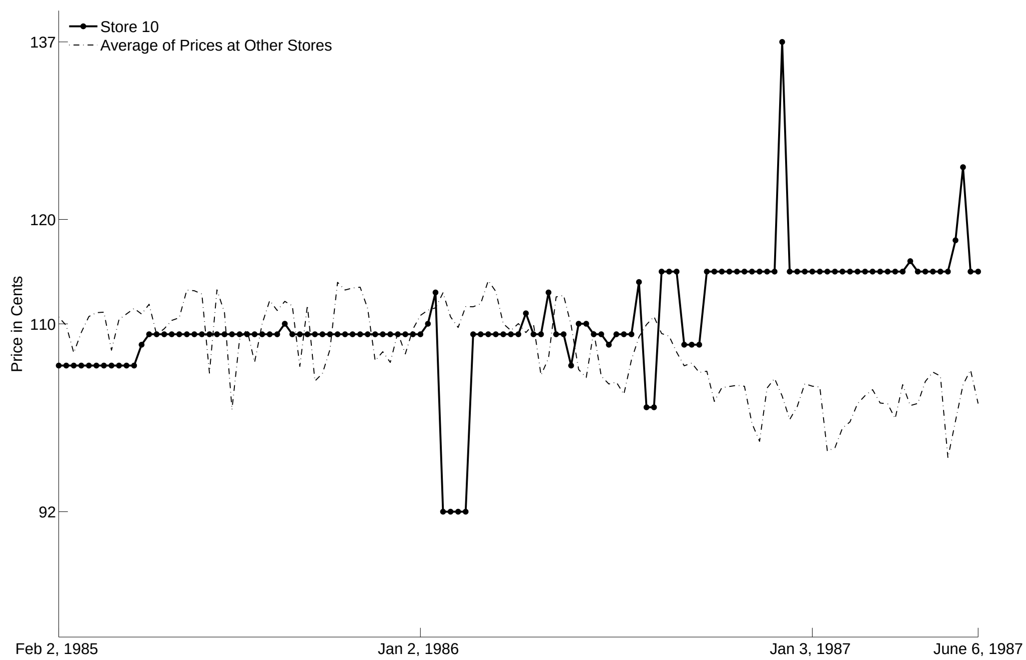
Notes: (i) The table's entries are in-sample root mean-squared errors (in percentage points) from forecasts of $I\{p_{i,t} \neq p_{i,t-1}\}$ based on linear-in-probabilities models that include the specified set of regressors. The maximum possible value for these is 50. (ii) The linear-in-probabilities 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: Inflation Rates for Margarine(i)



Note: (i) Annualized monthly inflation rates.

Figure 2: 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 3: Sample Hazard Function for Price Changes

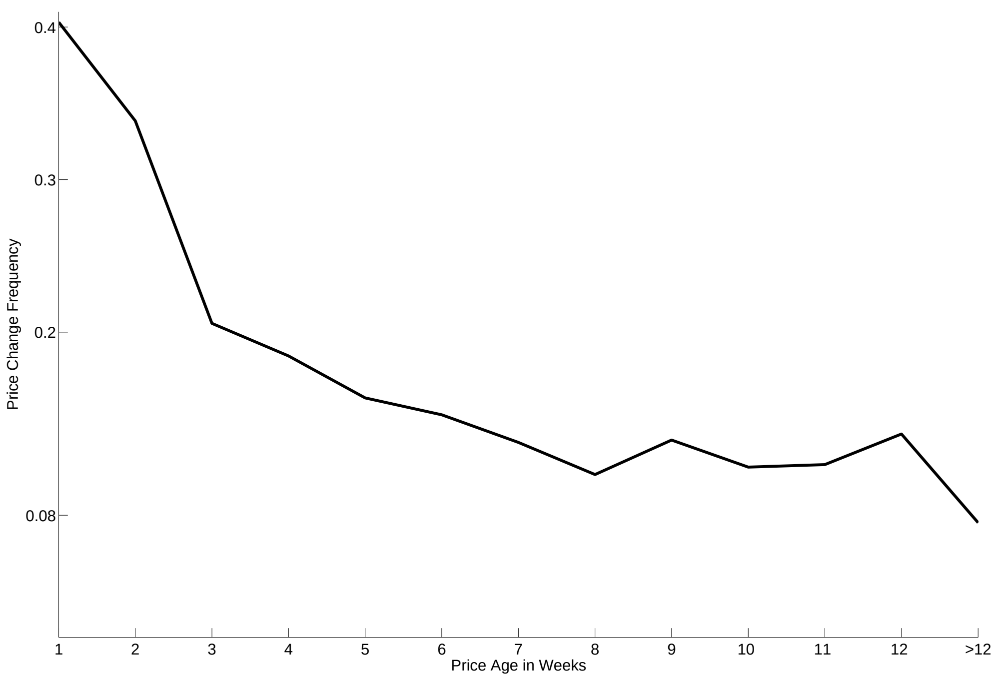
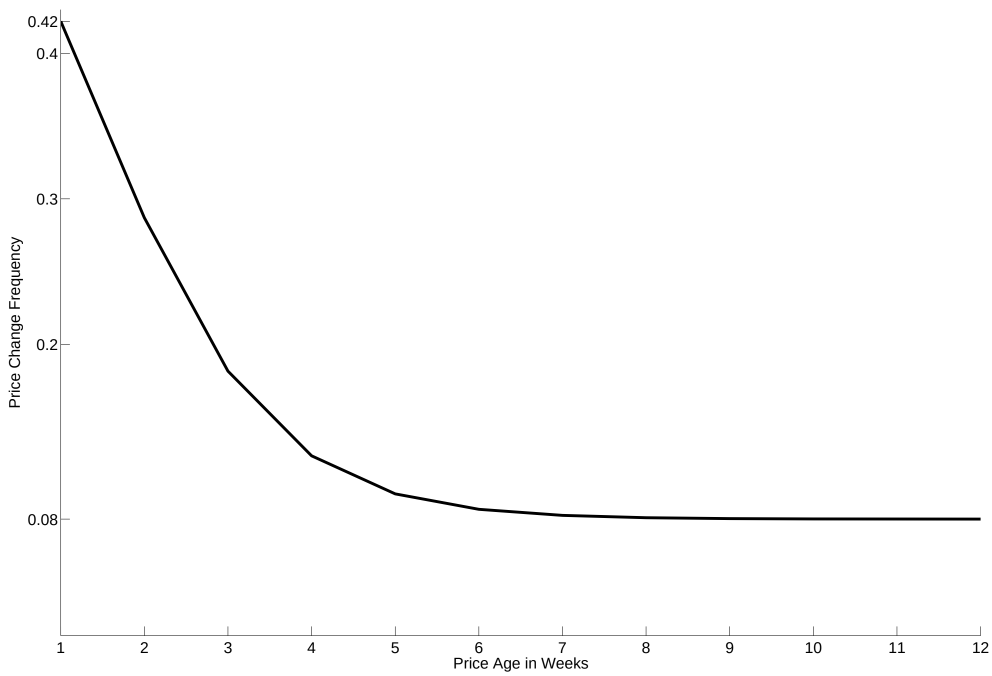
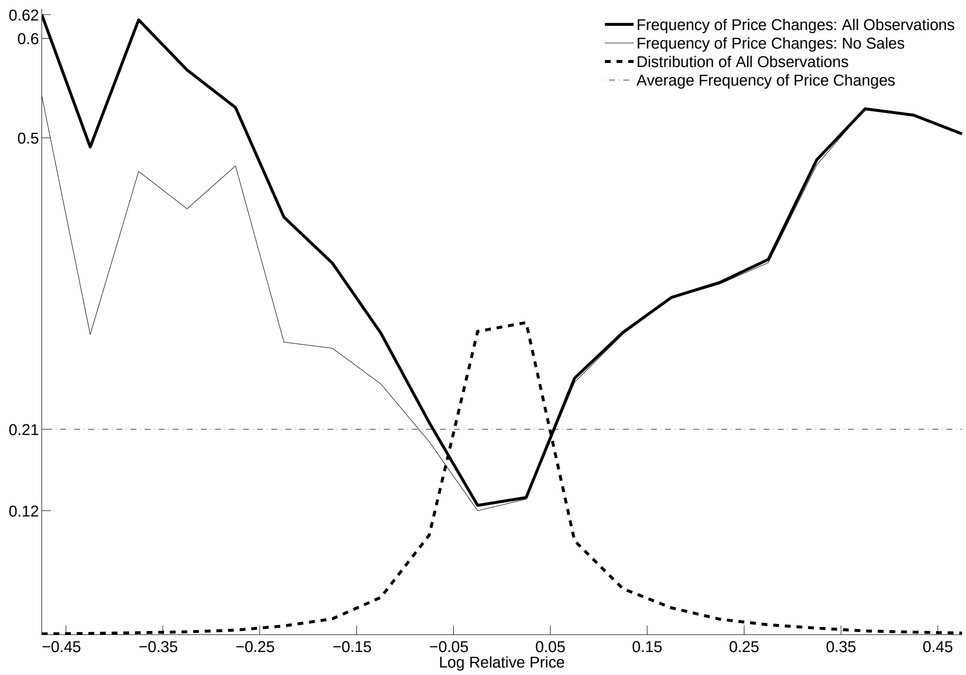


Figure 4: Hazard for Price Changes with Rigid and Flexible Prices(i)



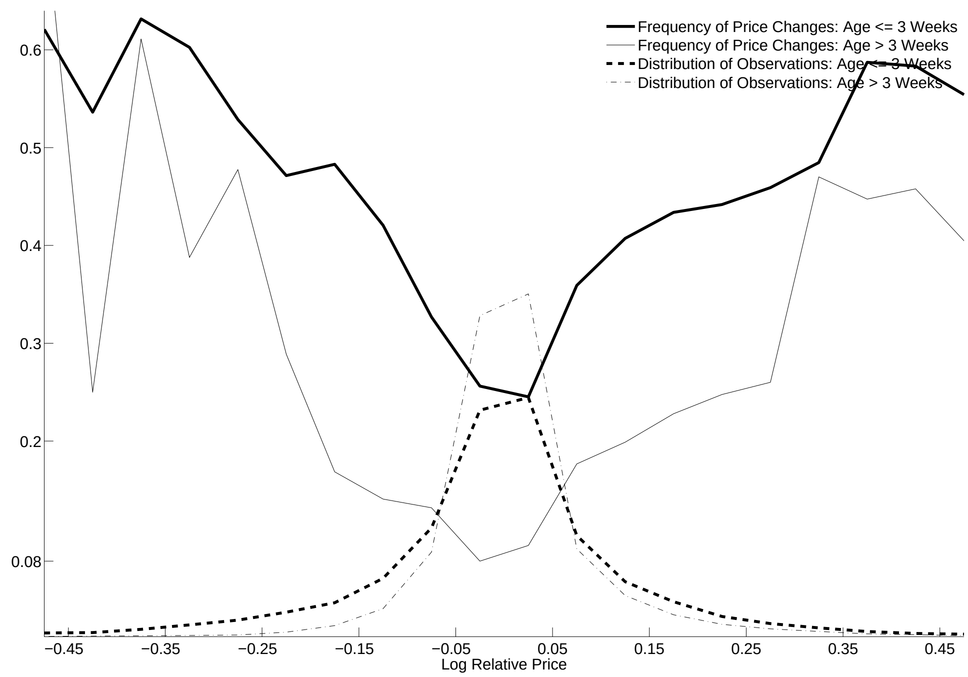
Note: (i) Hazard function from a simple model with flexible and rigid prices. See the text for further details.

Figure 5: Hazard for Price Changes as a Function of the Relative Price(i)



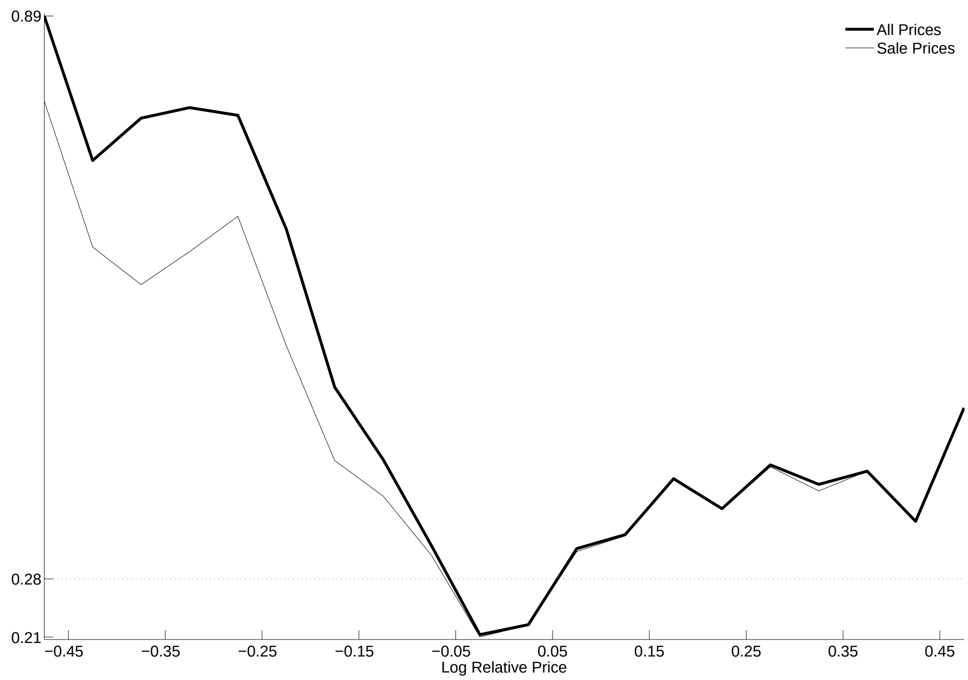
Note: (i) The relative price equals the store's price in the previous week divided by the average of all other prices for the same UPC charged in the current week. See the text for further details.

Figure 6: 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.

Figure 7: The Fraction of Young Prices by Relative Price[(i)]



Note: (i) Young prices are those with ages less than or equal to three weeks. The plotted fractions exclude prices one week old from both the numerator and denominator.

Figure 8: Implied Hazard Functions from the Linear-in-Probabilities Estimation[i]