Menu Costs, Multi-Product Firms, and Aggregate Fluctuations[†]

Virgiliu Midrigan*

Federal Reserve Bank of Minneapolis

August 2006

Abstract

This paper uses scanner price data collected in retail stores to document that (i) although the average magnitude of price changes is large, a substantial number of price changes are small in absolute value; (ii) the distribution of non-zero price changes has fat tails; and (iii) stores tend to adjust prices of goods in narrow product categories simultaneously. I present an extension of the standard menu-cost model to a multi-product setting in which firms face economies of scope in the technology of adjusting prices that can account for these higher-order moments of the data. The model, because of its ability to replicate this additional set of micro-economic facts, can generate aggregate fluctuations much larger than those in standard menu costs economies.

JEL classifications: E31, E32.
Keywords: state-dependent pricing, multi-product firms.

[†]I am indebted to George Alessandria, Bill Dupor, Paul Evans, Patrick Kehoe and Mario Miranda for valuable advice and support, as well as Ariel Burstein, Michael Dotsey, Narayana Kocherlakota, Ellen McGrattan, Martin Schneider, Ivan Werning and seminar participants at a number of venues for comments and suggestions. Any errors are my own.

*Research Department, Federal Reserve Bank of Minneapolis, 90 Hennepin Ave., Minneapolis, MN. Phone: (614) 390-5809. Email: midrigan.1@osu.edu

1. Introduction

New Keynesian Business Cycle models have received widespread attention in the macroeconomics of the last two decades, both as a tool for business cycle accounting, but also as a laboratory that underlies monetary policy discussions. At the heart of these models lies the assumption that individuals goods prices are sticky. In theory nominal price stickiness is typically motivated by menu costs: physical costs of changing price tags, reprinting catalogues, menus and other costs of communicating price changes to consumers. In practice, however, most of these models do not explicitly model the source of nominal price stickiness, but rather postulate that the timing of price changes is exogenous. Although their micro-foundations are not complete[1], these, so-called time-dependent, models continue to be widely studied, partly because of their computational simplicity, and partly because of the conjecture that they are a good reduced-form approximation to models in which price stickiness arises endogenously, from physical adjustment costs[2].

Whether this conjecture is indeed true is still an open question. The predictions of models in which price stickiness arises endogenously, due to menu costs, range from stark monetary neutrality[3] to cases in which the economy is virtually indistinguishable from time-dependent setups[4]. Golosov and Lucas (2003) study the properties of a model with firm-level disturbances capable to match the fact that the average size of price changes is large in the US economy: 10% on average, much larger than what can be explained by aggregate shocks alone. They find that the model produces very little output volatility from monetary shocks. Klenow and Kryvtsov (2004) reach

[1]See Bonomo and Carvalho (2004) and the references therein for models of endogenous time-dependent pricing.

[2]Seminal contributions include Barro (1972) and Sheshinski and Weiss (1977, 1983).

[3]Caplin and Spulber (1987), Caballero and Engel (1993), Golosov and Lucas (2004), Gertler and Leahy (2005).

[4]Klenow and Kryvtsov (2004), another version of the model in Gertler and Leahy (2005). See also Burstein (2003), Dotsey, King and Wolman (1999), Danziger (1999), Caplin and Leahy (1991) for studies that explore the consequences of fixed costs of resetting prices.

an opposite conclusion. They document that there is little evidence of across-firm synchronization in the US price data, contrary to what standard menu cost models predict. A model with time-varying costs of price adjustment that can replicate this feature of the data behaves identically to a time-dependent sticky price model and produces large output variability from monetary disturbances.

This paper revisits the question of whether menu costs of price adjustment can, in fact, generate a monetary transmission mechanism. I start by documenting several salient micro-economic features that characterize firm pricing behavior using a set of scanner price data collected in grocery stores. In addition to the large frequency and magnitude of price changes, documented by Klenow and Kryvtsov (2004), I document three additional features of the data. First, a large number of non-zero price changes are small in absolute value. Second, the distribution of price changes, conditional on adjustment, exhibits excess kurtosis. Finally, prices of goods sold by a particular retailer, especially those in narrow product categories, tend to adjust simultaneously.

The first two facts seem, at a first glance, inconsistent with menu-cost models. Firms that face fixed costs of adjustment only reprice when the losses from not doing so are large, and thus tend to do so by a large amount. As Lach and Tsiddon (2005) argue, however, extensions of the menu-cost model to a multi-product setting in which firms face interactions in the costs of price adjustment of various goods[5] can explain the large number of small price changes. Consider the extreme example of a restaurant whose prices are quoted on a single menu. If a single item on the menu is subject to a idyosincratic disturbance and needs repricing, the restaurant might find it optimal to pay the fixed cost and reprint the menu. Conditional on having payed this fixed cost, changing any other price on the menu is costless: the restaurant will then reprice all its other items, even for products that need small price changes. Indeed, the within-store synchronization observed in the data is indeed consistent

[5]See also Sheshinski and Weiss (1992).

with the presence of cost complementarities in the technology of price adjustment.

I next formulate, calibrate and quantitatively study the properties of a model in which a two-product firm faces a fixed cost of changing its entire menu of prices, but, conditional on paying this cost, zero marginal cost of resetting any given price on the menu. I calibrate the distribution of idiosyncratic marginal cost shocks, the size of the fixed costs of price adjustment, as well as the persistence of the marginal cost processes, by requiring the model to accord with the features of the data enumerated above. I find that the model, because of its ability to replicate this additional set of micro-economic facts, can generate aggregate fluctuations of the same magnitude as in time-dependent economies.

To understand the intuition behind this result, recall that the reason standard menu-cost models with large idiosyncratic uncertainty generate smaller real effects from monetary disturbances than their time-dependent counterparts is the fact that the identity of adjusters in models with menu costs varies endogenously in response to aggregate disturbances. Most firms that adjust in times of, say, a monetary expansion, are firms whose incentive to increase prices arising from the aggregate shock is reinforced by an idiosyncratic disturbance that triggers a desired price change in the same direction. The money shock thus affects the aggregate price level through two channels: by increasing the desired price change of the adjusting firms, but also by changing the mix of adjusters towards firms whose idyosincratic shocks call for larger price increases. This latter selection effect, is absent, by assumption, in time-dependent models, and ensures that the aggregate price level is more responsive to nominal shocks. Its strenght depends however on the mass of firms in the economy whose desired price changes lie in the neighborhood of the adjustment thresholds, a property of the economy that depends on higher-order moments of the distribution idiosyncratic disturbances in the economy.

I find that matching the excess kurtosis of price changes and the large number of small price changes observed in the data requires that the distribution of idyosincratic disturbances be highly leptokurtic. This feature of the calibration, as well as the fact that the adjustment hazard is positive even for goods whose desired price changes are close to zero, because of the cost complementarity in the price adjustment technology, reduces the role of self-selection and therefore the responsiveness of the aggregate price level to monetary shocks.

I proceed as follows. Section 2 discusses the data used in the empirical work, and documents its salient features. Section 3 discusses the model economy. Section 4 quantitatively evaluates its performance. Section 5 concludes. Appendices discuss the non-linear solution techniques used to solve the functional equations that characterize the equilibrium of the model economy and several aspects of the data in more detail.

2. Data

I conduct inference using two sources of publicly available sets of scanner price data, maintained by the Kilts Center for Marketing at the University of Chicago Graduate School of Business[6]. The first dataset was assembled by AC Nielsen and consists of daily observations on the purchasing practices of a panel of households in Sioux Falls (South Dakota) and Springfield (Missouri). I use this household level data to construct a panel of weekly price series spanning more than two years (January 1985 to March 1987), 31 stores and 115 products in six different product categories (ketchup, tuna, margarine, peanut butter, sugar and toilet tissue)[7].

The second source of data is a by-product of a randomized pricing experiment conducted by the Dominick's Finer Foods retail chain in cooperation with the Chicago GSB. Nine years (1989 to 1997) of weekly store level data on the prices of more than 4500 products for 86 stores in the Chicago area are available. The products available in

[6] The data is available online at http://gsbwww.uchicago.edu/kilts/research/index.shtml

[7] The actual number of observations is larger in the original dataset, but I discard stores/goods with a large number of missing observations. The criteria for inclusion in the sample are discussed in the appendix.

this database range from non-perishable foodstuffs (frozen and canned food, cookies, crackers, juices, sodas, beer), to various household supplies (detergents, softeners, bathroom tissue), as well as pharmaceutical and hygienic products.

I discuss, in a data appendix, several aspects regarding the construction of price series. In particular, I time-aggregate weekly data into monthly observations in order to calculate statistics that can be used to evaluate the performance of a model economy in which the length of the period is a month. For Dominick's data, which sets prices on a chain-wide basis, I construct a chain-wide price using the price of the store that has the least number of missing observations for a particular good. Following Golosov and Lucas (2003), I filter out temporary price cuts (sales) that last less than four weeks. I could alternatively incorporate into the model some of the frictions that have been proposed to explain this pattern of retail price variation[8], but this would increase the model's complexity considerably, without producing additional insights. In particular, as I document below, none of the empirical facts I am about to document are an artifact of my decision to purge the data of sales and time-aggregate the data.

A. The Size and Frequency of Price Changes

Figure 1 presents histograms of the distribution of price changes, $\log\left(\frac{p_t}{p_{t-1}}\right)$, conditional on adjustment, for the two sets of data, pooled across all goods/stores/months in each sample[9]. I truncate these distributions, by eliminating the top and bottom 1% of observations, in order to ensure that results are not driven by outliers. Superimposed on each histogram is the density of a normal distribution with the same mean and variance as that of the distribution of price changes. Table 1 reports moments of these distributions, again computed using the truncated sample of observations.

[8] Informational frictions on consumer's side of the market (Varian 1980), demand uncertainty (Lazear 1986), or thick-market explanations (Warren and Barsky, 1995), to name a few.

[9] These histograms and the statistics reported below are unweighted. Weighting goods by average (across time) sales shares within a store produces very similar results.

Several facts emerge in the data.

Fact 1: A large number of price changes are small in absolute value.

Consistent with the evidence presented by Klenow and Kryvtsov (2004), the average size of price changes is large[10]: stores in the AC Nielsen data adjust prices by 10.4% on average, while those in Dominick's sample do so by 7.7%. Notice however, in Figure 1, that a large number of price changes are close to zero. I define, in the data and in the model of the next section, a "small" price change as any price change whose magnitude is less than one-half of the mean of the absolute value of price changes in the data. Roughly 30% of price changes in both datasets are below this cutoff (5.2% and 3.8%, respectively).

Fact 2: The distribution of price changes exhibits excess kurtosis.

Notice, in Figure 1, that the number of price changes in the vicinity of zero is greater than that predicted by a normal distribution, while the tails are somewhat fatter. As Table 1 indicates, the kurtosis of price changes is 3.5 and 5.4, respectively, larger than that of a Gaussian distribution[11].

None of these features of the data are an artifact of my decision to focus on regular price changes. The kurtosis of all (including temporary price cuts) monthly price changes (again excluding the top and bottom 1% of observations) is 4.6 and 3.9 in the AC Nielsen and Dominicks data, respectively. Moreover, 34% and 40% of price changes are less than half the mean of the absolute value of price changes in the sample[12].

Fact 3: Prices in narrow product categories within a store tend to adjust si-

[10]The excessive volatility of individual goods' prices has also been documented for countries other than the US. See Dhyne et. al (2005) for a survey of findings from studies of European micro-price data.

[11]Kurtosis is defined as the ratio of the fourth central moment to the square of the variance. The kurtosis of the normal according to the convention I employ is then equal to 3.

[12]The average size of a price change is 14.3% and 13.3% in the AC Nielsen and Dominick's data respectively if I include price changes arising due to temporary price cuts.

multaneously.

I establish that prices within a store adjust in tandem using a reduced-form discrete-choice specification in which I model each product's price adjustment decision as a function of variables that proxy for marginal cost disturbances, as well as variables that capture the price adjustment decisions of other products sold by a store. To be clear, no attempt is made to identify causality or the source of synchronization here: the exercise below is simply a statistical description of the extent to which price changes within a store are synchronized.

Assume that the good's optimal price is $p^*_{it} = \gamma c_{it} + u_{it}$, where c_{it} collects all observable components of a good's marginal cost. Assuming that the firm sets $p_{it} = p^*_{it}$ every time it adjusts, and that it adjusts whenever $p_{it} - p^*_{it} \notin [s_{it}, S_{it}]$, the firm's price adjustment decision is

$$x_{it} = \begin{cases} 1, & \text{if} \quad \gamma\Delta c_{it} + \Delta u_{it} > S_{it} \\ 0, & \text{if} \quad s_{it} \leqslant \gamma\Delta c_{it} + \Delta u_{it} \leqslant S_{it} \\ -1, & \text{otherwise} \end{cases}$$

where, say, $\gamma\Delta c_{it} = \gamma c_{it} - \gamma c_{i\tau}$ is the growth rate of the product's marginal cost since the previous price adjustment[13], and the sign of x_{it} denotes the direction of the price change, if any. I assume that $\Delta u_{it} \sim N(0, 1)$ as the model's scale and location are unidentified. I use the wholesale price of a particular product (Dominick's) or the average price of a store's competitors (AC Nielsen where wholesale price data is unavailable) as a proxy for the marginal cost of selling a good, in addition to the hourly wage rate in the retail sector and the energy and food CPIs to proxy for economy-wide disturbances to a product's desired price.

[13] Cecchetti (1986) employs a similar (one-sided) model in order to study the price adjustment of magazine prices.

To quantify the extent to which price changes within a store are synchronized, I parameterize the upper and lower thresholds as linear functions of three measures of within-store synchronization: (i) the fraction of all remaining goods within a store whose prices change in a given month; (ii) the proportion of price changes in a particular good-category (data on 29 product categories, ranging from analgesics to toothpastes is available for Dominick's and 6 categories for AC Nielsen); (iii) the proportion of prices of goods produced by the manufacturer of the product in question that experience a price change[14], as well as (iv) the proportion of prices of this particular product that are changed in all remaining stores[15]. All these measures of synchronization are computed based on the adjustment decision of all goods other than i in a given group, and I exclude those observations for which any of these statistics are calculated based on fewer than five observations in a given period.

These measures of within-store synchronization make intuitive sense. Levy, Dutta, Bergen and Venable (1998) use store-level data for five supermarket chains and report the steps undertaken during a price change process. They report that the bulk (60%) of the labor effort used to adjust prices is spent on price tag changes and verification, of which most time (50-60%) goes into finding specific items on shelves. One would thus expect that economies of scope in changing prices are larger for products located in adjacent shelves/aisles. An ideal measure of relevant within-store synchronization would then be the fraction of prices adjusted in a given aisle. In the absence of such data, I use the fraction of price changes within a category group, or produced by a given manufacturer, as these items are usually placed in adjacent locations within the store.

In addition, I allow both the scale and location of the adjustment thresholds

[14]Identified based on the first 5 digits of the upc code. I include this variable only for Dominick's data, as too few goods per manufacturer are available in case of the AC Nielsen data.

[15]Only available in the AC Nielsen data, as Dominick's sets prices on a chain-wide basis. See the Appendix for details.

to differ across months and product-categories, in order to control for heterogeneity across goods/time-periods. I do so by allowing fixed product-category and month effects in the two threshold equations[16].

As Table 2 illustrates, the probability that a particular product experiences a price change does indeed depend on the fraction of other prices within that store that experience adjustment, especially those in particular product category. In case of Dominick's data, an increase in the fraction of remaining prices that change in a given product category from 0 to 1 increases the probability that a given product will experience a price cut by 5% and that of a price increase by 29%, thereby increasing the probability of a price change by 34%. Synchronization is even stronger for goods in a given manufacturer category. An increase in the fraction of price changes of the remaining goods produced by a given manufacturer from 0 to 1 increases the probability that the good in question will also adjust by 55%. The correlation between the price adjustment decisions of various goods is even larger in case of the AC Nielsen data: an increase in the fraction of remaining prices that experience adjustment in a given product category increases the probability that a particular good will adjust as well by 96%.

Table 3 presents an additional set of facts that will be used in order to calibrate the model. It has been widely documented[17] that prices in retail stores adjust frequently. The two datasets I employ here are no exception. Despite the fact that I overestimate the duration of price spells by aggregating weekly data to monthly and eliminating a large number of temporary price cuts, the average price spell lasts 4 months in the AC Nielsen data and 5.2 months in the case of Dominick's prices[18].

Let $\hat{p}_t$ be the price (in logs) of a good in period t, expressed in deviations

[16]Note that the incidental parameters problem does not arise here as the number of observations within a group is large: 6000 on average.

[17]Kackmeister (2005), Dutta, Bergen and Levy (2002).

[18]Average duration falls to 2.7 months and 2.4 months, respectively, if one takes into account price changes associated with temporary price cuts.

from a time trend. If marginal cost shocks are transitory, one would expect two observations of the firm's price, $\hat{p}_t$, sufficiently distant in time, to lie close to each other. In contrast, if shocks are highly persistent, the firm's price wonders away from the mean, and differs considerably from the price the firm has set in the past. Let

$$D_k = \frac{\frac{1}{\mathcal{N}(G_k)} \sum_{t \in G_k}^{T} |\hat{p}_t - \hat{p}_{t-k}|}{\frac{1}{\mathcal{N}(G_a)} \sum_{t \in Ga}^{T} |\hat{p}_t - \hat{p}_{t-1}|}$$

be the mean absolute difference in a good's (detrended) price in periods that are k-months apart, relative to the average absolute value of non-zero price changes, where G_k is the set of time-periods for which prices were recorded in t and $t-k$, G_a the set of periods in which the product has experienced a non-zero price change, and $\mathcal{N}$ the number of elements of a given set. Note the similarity of these statistics, which I call deviance ratios, to the variance ratios popularized by Cochrane (1988) in non-parametric tests of non-stationarity. These deviance ratios are larger, the more persistently $\hat{p}_t$ moves in a given direction, and although they have no structural interpretation, they can be used, in conjuction with the model to be presented below, to infer the persistence of marginal cost shocks. The last rows of Table 3 present the average values of this statistic in the data, at 12- and 24-month horizons. These ratios are close to 1, suggesting that shocks are not too persistent[19].

An alternative measure of persistence is the probability that the next price change will have the same sign as the current one. As Table 3 reports, these probabilities[20] are low (32% and 41%, respectively), suggesting that shocks that trigger

[19]Given the short span of AC Nielsen's price series, deviance ratios are only reported for 12-month horizons in this data.

[20]Calculated as an equally weighted average of the probabilities of two consecutive positive/negative price changes. Equal weights (as opposed to weights based on the long-run probability of a price increase/decrease) are used in order to account for the upward trend in prices in

price changes are transitory and price changes tend to be reversed. Although useful in providing information about the persistence of idiosyncratic shocks, this alternative statistic will only be used as an "over-identifying" check on the model, as it is sensitive to the definition of sales employed to purge data of temporary price cuts.

B. Ex-ante heterogeneity?

As shown by Caballero and Engel (1993), Caplin and Spulber's (1987) neutrality result survives in a world with heterogeneity in menu costs, demand elasticities, etc. across firms, heterogeneity which can, in principle, give rise to the fat tails and large number of small price changes depicted in Figure 1. I ask whether ex-ante heterogeneity is indeed responsible for the features of the data documented above by using variance decompositions in which I gauge the importance of month, product[21], and store-specific effects in explaining the variability of the magnitude and frequency of price changes reported above. Specifically, I estimate

$$y_{it}^{s} = c + d_i + d_s + d_t + e_{it}^{s},$$

where d_i, d_s, d_t are good, store, and month-specific effects and y_{it}^{s} is the size of price changes, $|\Delta \log (p_{ist})|$, or the duration of price spells that end in a given period. As Table 4 indicates, month or store-specific heterogeneity accounts for less than 10% of the variation of the frequency and size of price changes in the data. Good-specific effects are somewhat more volatile, but nevertheless responsible for less than 16% of the variation in the sample.

Figure 2 depicts the higher-order moments discussed above, the kurtosis, and number of price changes in the vicinity of zero for 29 narrow product categories in

Dominick's data. Use of long-run weights results in a probability of two consecutive price changes in the same direction equal to 46% in Dominick's data, but this number overstates the persistence of the price series as 65% of price changes in Dominick's data are positive.

[21]The number of non-zero price changes for a given product is small in Dominick's data in which I have collapsed the prices of the different stores into a single, chain-wide price. I therefore estimate product-category×manufacturer, as opposed to individual good effects for this dataset.

the Dominick's sample. Although substantial differences in the size of price changes across these product categories are evident in the data, all of these products are characterized by a substantial number of small price changes and excess kurtosis, suggesting once again that systematic differences in the size of price changes across goods are not the driving force behind the distributional features documented above.

C. Relationship with other evidence

I have documented three features of the distribution of price changes in grocery stores that will prove important in the calibration of the model economy of the next section: (i) a large number of price changes are small, and (ii) the distribution of price changes is leptokurtic, and (iii) prices within a store tend to adjust in tandem. None of these features of the data I study are unique to grocery stores.

Klenow and Kryvtsov (2004) report that 40% of price changes are less than 5% in absolute value in their dataset of BLS-collected price data covering all goods and services used in the construction of the CPI, a dataset in which prices change by 9.5% on average. Kashyap (1995) uses a dataset of prices for products sold in retail catalogues and also documents that many price changes are small: 44% of price changes in his dataset are less than 5% in absolute value. The kurtosis of price changes, conditional on adjustment, is 15.7 in the data and falls to 6.2 if one excludes the top and bottom 1% of observations[22]. Kackmeister (2005) presents a histogram of the distribution of price changes in a dataset of prices in retail stores: one-third of price changes are less than 10% in absolute value in an environment where the average magnitude of price changes is 20%. Finally, Lach and Tsiddon (1996) provide evidence that stores synchronize price adjustments of various products using a dataset of prices collected in Israel[23]. They find that the variability (across stores) of the fraction of products whose prices change in a given period is larger than what

[22] These numbers are based on my own calculations using the data published in Kashyap (1995).
[23] See also Fisher and Konieczny (2000).

would be expected if price adjustment decisions were independent across goods.

D. Discussion

The large number of small price changes documented above is inconsistent with the predictions of simple menu-cost models. Several extensions to the standard model have been suggested in order to render it consistent with the data. One might assume time-varying adjustment costs[24]. Alternatively, as Kashyap (1995) has suggested, one might allow fluctuations in the degree of market power possessed by firms, arising from variation in consumer search costs over time[25]. In this paper I will explore an alternative route, one that can simultaneously explain both the large number of small price changes observed in the data, but also the within-store synchronization of price changes. As Lach and Tsiddon (2005) have argued, an extension of the state-dependent model to a multi-product environment in which firms face large average costs of adjusting a menu of prices, but a small marginal cost of changing any given price on the menu, can also generate a large number of small price changes. Cost complementarities in price adjustment are not crucial for this paper's key results, as will be made clear below. I nevertheless assume them in the model of the next section because they provide a simple and tractable parametric extension of the menu-cost model capable of replicating the salient features of the micro-price data documented above.

3. The Model

A. Model Economy

Throughout, let s_t denote the event realized at time t, $s^t = \{s_0, s_1, ..., s_t\}$ the history of events up to this period and $\pi(s^t)$ the probability of a particular history as of time 0. The economy is populated by a continuum of consumers and a continuum of monopolistically competitive firms, both of mass 1. Consumers are identical, while

[24]See Caballero and Engel (1999) and Dotsey, King and Wolman (1999).
[25]See also Benabou (1992).

firms (indexed by z) differ according to their productivity level. Each firm sells two products, indexed by $i = 1, 2$. I first discuss the problem of the representative consumer, that of the firm, and then define an equilibrium for this economy.

Consumers

Consumers' preferences are defined over leisure and a continuum of imperfectly substitutable goods. The consumer sells part of her time endowment to the labor market and invests her wealth in one-period shares in firms. In equilibrium, identical consumers own equal shares of all the economy's firms. The representative consumer's problem is to choose, given prices, how to allocate her income across the different goods available for consumption and how much to work:

$$\max_{\{c^1(z;s^t),c^2(z;s^t)\},n(s^t),b} \sum_{t=0}^{\infty} \beta^t \sum_{s^t} \pi(s^t) U(c(s^t), n(s^t)),$$

subject to

$$\int_0^1 \left[p^1(z,s^t)c^1(z,s^t) + p^2(z,s^t)c^2(z,s^t)\right] dz = w(s^t)n(s^t) + \Pi(s^t),$$

where

$$c(s^t) = \left(\int_0^1 \left(\tfrac{1}{2} c^1(z,s^t)^{\frac{\theta-1}{\theta}} + \tfrac{1}{2} c^2(z,s^t)^{\frac{\theta-1}{\theta}} \right) dz \right)^{\frac{\theta}{\theta-1}}$$

is an aggregator over the different varieties of goods that the household consumers, $n(s^t)$ is the supply of labor, $w(s^t)$ the nominal wage rate, $\Pi(s^t)$ the profits the consumer receives from her ownership of firms, $p^1(z,s^t)$ and $p^2(z,s^t)$ are the prices of each good and θ is the elasticity of substitution across goods. Notice that I have assumed that the elasticity of substitution across goods sold by a single firm is equal to the elasticity of substitution across goods sold by different firms.

Firms

Firms produce output using a technology linear in labor:

$$y^i(z,s^t) = a^i(z,s^t) l^i(z,s^t), \; i = 1, 2,$$

where the firm's technology, $a^i(z, s^t)$, evolves according to

$$\log a^i(z, s^t) = \rho_a \log a^i(z, s^{t-1}) + \varepsilon^i(z, s^t),\ i = 1, 2,$$

and $\varepsilon(z, s^t) \in [\varepsilon_{\min}, \varepsilon_{\max}]$ is a random variable, uncorrelated across firms, goods and time-periods[26]. Firms operate along their consumers' demand schedules, derived as solutions to the consumer's problem discussed above:

$$c^i(z, s^t) = \left(\frac{p^i(z,s^t)}{P(s^t)}\right)^{-\theta} c(s^t),$$

where $P(s^t)$ is the price index in this economy, defined as a consumption-weighted average of the prices in this economy:

$$P(s^t) = \left(\int_0^1 \left[\tfrac{1}{2}p_t^1(z, s^t)^{1-\theta} + \tfrac{1}{2}p_t^2(z, s^t)^{1-\theta}\right] dz\right)^{\frac{1}{1-\theta}}.$$

I assume that firms face fixed menu costs of resetting prices. Any time at least one (or both) of the two prices change, the firm must hire ξ additional units of labor. Let $q(s^t) = \beta^t \frac{U_c(c(s^t),n(s^t))}{U_c(c(s^0),n(s^0))}$, where U_c is the marginal utility of consumption, denote the t-period stochastic discount factor. The firm's problem is to maximize

$$\sum_{t=0}^{\infty} \sum_{s^t} \pi(s^t) q(s^t) \Pi(z, s^t),$$

where

[26]Clearly, the assumption that firms sell two goods each, that elasticities of substitution are identical across and within firms and that cost shocks across the two goods are uncorrelated are easily falsifiable and the quantitative exercise here is merely illustrative of the forces at work in a menu-cost model. Increasing the number of goods within a firm will render the model closer to a model with constant adjustment hazards a la Calvo as each good becomes atomistic, and the probability that a particular good's price adjusts is independent of its desired price change. Conversely, as the correlation of shocks across goods increases to unity, or as products sold by the same firm become more substitutable, the model behaves similarly to a single-product model. The calibrated parameter values (especially the distribution of idyosincratic shocks) that allow the model to match the distributional features of the data documented in the previous sections are therefore sensitive to these assumptions I make and can only be interpreted in the context of this particular parametrization of the model. The aggregate properties of the model depend however on the distribution of price changes conditional on adjustment, not that of underlying shocks, as will be shown below, and this statistic is calibrated to match the properties of the micro-data.

$$\Pi(z,s^t) = \sum_{i=1,2} \left(\frac{p^i(z,s^t)}{P(s^t)}\right)^{-\theta} \left(\frac{p^i(z,s^t)}{P(s^t)} - \frac{w(s^t)}{a^i(z,s^t)P(s^t)}\right) c(s^t) -$$
$$-\xi \frac{w(s^t)}{P(s^t)} \mathcal{I}_{p^1(z,s^t)\neq p^1(z,s^{t-1}) \text{ or } p^2(z,s^t)\neq p^2(z,s^{t-1})},$$

and $\mathcal{I}$ is an indicator function. The last term of this expression is the increase in the firm's wage bill if it decides to adjust any of its two prices.

B. Equilibrium

I introduce money by assuming that nominal spending must be equal to the money stock[27]:

$$\int_0^1 \sum_{i=1,2} p^i(z,s^t) c^i(z,s^t) dz = M(s^t)$$

The money supply growth rate $\mu(s^t) = \frac{M(s^t)}{M(s^{t-1})}$ evolves over time according to an AR(1) process:

$$\log \mu(s^t) = \rho_\mu \log \mu(s^{t-1}) + \eta(s^t),$$

where η is an iid $N(0,\sigma_\eta^2)$ disturbance. The equilibrium is a collection of prices and allocations: $p^i(z,s^t), w(s^t), P(s^t), c^i(z,s^t), c(s^t), n(s^t), l^i(z,s^t), y^i(z,s^t)$ such that, taking prices as given, consumer and firm allocations, as well as firm prices solve the consumer and firm problems, respectively, and the labor, goods, and money markets clear.

C. Computing the Equilibrium

I normalize all nominal variables by the money stock in the economy, e.g., $\tilde{P}(s^t) = \frac{P(s^t)}{M(s^t)}$, in order to render the state-space of this problem bounded. Let $\tilde{p}^i_{-1}(z,s^t) = \frac{\tilde{p}^i(z,s^{t-1})}{M(s^t)} \in \mathcal{P}$ be a firm's (normalized) last period's price and $\mathcal{A} = [\frac{\varepsilon_{\min}}{1-\rho}, \frac{\varepsilon_{\max}}{1-\rho}]$ the support of the distribution of technology levels in the economy. The aggregate state of this economy is an infinite-dimensional object, consisting of the growth rate of

[27] See Rotemberg (1987) for a transaction technology that gives rise to this particular specification of money demand.

money: $\mu(s^t)$, but also of the endogenously varying joint distribution of last period's firm prices and technology levels. Let $\boldsymbol{\phi}$: $\mathcal{P}^2 \times \mathcal{A}^2 \to [0,1]$ denote this distribution and $\boldsymbol{\Gamma}$ it's law of motion: $\boldsymbol{\phi}' = \boldsymbol{\Gamma}(g, \boldsymbol{\phi})$. Finally, let $\mathbf{a} = (a^1, a^2)$ be a vector of a firm's technology levels and $\mathbf{p}_{-1} = \left(\tilde{p}^1_{-1}, \tilde{p}^2_{-1}\right)$ collect the firm's last period's nominal prices.

Let $V^a(\mathbf{a};\mu, \boldsymbol{\phi})$ and $V^n(\mathbf{p}_{-1}, \mathbf{a}; \mu, \boldsymbol{\phi})$ denote a firm's value of adjusting and not adjusting its nominal prices, as a function of its last period's prices and current technology, as well as the aggregate state of the economy. These two functions satisfy the following system of functional equations:

$$V^a(\mathbf{a};\mu, \boldsymbol{\phi}) = \max_{\mathbf{p}} \left(\sum_{i=1,2} \left(\frac{\tilde{p}^i}{\tilde{P}} - \frac{\tilde{w}}{a^i \tilde{P}} \right) \left(\frac{\tilde{p}^i}{\tilde{P}} \right)^{-\theta} c - \xi \frac{\tilde{w}}{\tilde{P}} + \beta \int \frac{U_c\prime}{U_c} V(\mathbf{p}'_{-1}, \mathbf{a}'; \mu\prime, \boldsymbol{\phi}\prime) dF(\varepsilon^1, \varepsilon^2, \eta) \right)$$

$$V^n(\mathbf{p}_{-1}, \mathbf{a}; \mu, \boldsymbol{\phi}) = \sum_{i=1,2} \left(\frac{\tilde{p}^i_{-1}}{\tilde{P}} - \frac{\tilde{w}}{a^i \tilde{P}} \right) \left(\frac{\tilde{p}^i_{-1}}{\tilde{P}} \right)^{-\theta} c + \beta \int \frac{U_c\prime}{U_c} V(\mathbf{p}'_{-1}, \mathbf{a}'; \mu\prime, \boldsymbol{\phi}\prime) dF(\varepsilon^1, \varepsilon^2, \eta),$$

where $V = \max(V^a, V^n)$ is the firm's value function and $\mathbf{p}$ is a vector of nominal prices the firm chooses every time it adjusts. The laws of motion for the state variables are:

$$\boldsymbol{\phi}\prime = \boldsymbol{\Gamma}(\mu, \boldsymbol{\phi}), \alpha^{i\prime} = a^{i\rho_a} \exp\left(\varepsilon^i\right), \mu\prime = \mu^{\rho_\mu} \exp(\eta)$$
$$\tilde{p}^{i\prime}_{-1} = \begin{cases} \frac{\tilde{p}^i}{\mu} \text{ if adjust} \\ \frac{\tilde{p}^i_{-1}}{\mu} \text{ otherwise} \end{cases}$$

The unknowns in this problem are the following functions: $V^a(), V^n(), c(), \tilde{w}(), \tilde{P}(), \boldsymbol{\Gamma}()$. To solve this system of functional equations, I (i) allow aggregate variables to depend only on a finite number of the moments of $\boldsymbol{\phi}$, as opposed to the entire distribution, following a suggestion by Krusell and Smith (1997); (ii) replace the unknown functions with a linear combination of orthogonal polynomials; and (iii) solve for the unknown coefficients on these polynomials by requiring that the system of six functional equations (the Bellman equations, as well as the equilibrium conditions) be exactly satisfied at a finite number of nodes along the state-space. A technical appendix discusses the solution method in more detail.

4. Quantitative Results

A. Calibration and Parametrization

I parameterize the utility function as

$$U(c, n) = \log(c) - \psi n.$$

This specification follows Hansen (1985) by assuming indivisible labor decisions implemented with lotteries. I set the length of the period to one month, and therefore choose a discount factor $\beta = .997$. I choose ψ to ensure that in the absence of aggregate shocks households supply 1/3 of their time to the labor markets. To calibrate the process characterizing the growth rate of the money supply, I estimate an AR(1) process for the growth rate of M1 for the US economy for 1985-1997, the years for which the micro-price data used to calibrate the model is available. I choose $\theta = 3$, a number in the range of estimates of demand elasticities available in the retail industry[28]. Table 6 summarizes the choice of parameter values I assign the model.

The rest of the parameters are calibrated: ξ – the size of the fixed costs incurred by the firm when it changes its menu of prices, ρ_a – the parameter that governs the persistence of marginal cost shocks, as well as the distribution of technology shocks. I choose these parameters in order to match the salient properties of the micro-price data discussed in Section 2. I target an average duration of price spells of 4.5 months, an average value for the two datasets; an average size of price changes of 9%; a standard deviation of price changes of 12%; and a kurtosis of 4.5. I also require that the model generates 30% 'small' (less than one-half of the mean) price changes, as well as a 24-month 'deviance ratio' of 1.02. Additional "over-identifying" checks will be used to gauge the persistence of marginal cost shocks in the data. Table 5 reports the choice of moments used to calibrate the model economy.

[28]Nevo (2001), Barsky et. al. (2000), Chevalier, Kashyap and Rossi (2003)

To calibrate the distribution of idiosyncratic technology shocks, I assume that shocks ε_t are drawn from the following parametric family of distributions:

$$\varepsilon_t = \begin{cases} -b_t\varepsilon^{\max}, & \text{with} \quad p = \frac{1}{2} \\ b_t\varepsilon^{\max}, & \text{with } 1 - p = \frac{1}{2} \end{cases}$$

where b_t is a random variable drawn from a Beta distribution with parameters α_1 and α_2. The distribution of technology shocks is thus symmetric around zero, and flexible enough to enable the model to reproduce the distributional features of the data.

B. Results

Benchmark Model

I solve for the unknown parameters: $\rho_a, \xi, \alpha_1, \alpha_2, \varepsilon^{\max}$, by minimizing the sum of squared log-deviations of the model-generated moments from the six targets in Table 5. The last rows of Table 6 (the column labeled Benchmark model) reports the calibrated parameter values.

Marginal cost shocks are fairly transitory: $\rho_a = 0.5$. Firms pay a menu cost equal to 1.2% of their steady-state labor bill (0.8% of revenue) every time they undergo a new price change, a number close to that reported by Levy et. al. (1997) in a study of the price adjustment costs of five large supermarkets. The distribution of technology shocks is highly leptokurtic, with a kurtosis in excess of 20 and a variance of 2.7×10^{-3} ($\alpha_1 = 0.05$, $\alpha_2 = 1.30$, $\varepsilon^{\max} = 0.4$).

Before proceeding to analyze the model's performance, I briefly discuss the consequences of the assumption that the firm faces a fixed cost of changing an entire menu of prices, as opposed to a given price on the menu. Figure 3 plots a firm's adjustment region, in the $(\tilde{p}^1_{-1}, \tilde{p}^2_{-1})$ space (prices are expressed as log-deviations from the optimum) for several values of the firm's productivity parameter. Because of the cost-complementarity in price adjustment, a firm's adjust decision depends on the deviation of its two prices from their respective optima. Small price changes will therefore arise in equilibrium whenever at least one of the firm's two prices are hit by a sufficiently large shocks. The figure also performs a comparative statics exercise. I illustrate in this figure the size of the inaction regions for different values of the productivity level of the firm. Similar to what Golosov and Lucas (2003) find, firms are more willing to adjust their prices in periods when their technology is higher.

I next evaluate the model's performance quantitatively. Note, in Table 5, that the model is successful at matching the salient properties of the microeconomic data documented in Section 2, with all model-based moments close to their targets. In particular, the kurtosis of the distribution of price changes is 4.3, and 33% of price changes are less than 4.4% in absolute value ($\frac{1}{2}$ the mean of absolute value of non-zero price changes). The model also does well in matching other, 'over-identifying' restrictions used to infer the accuracy of the estimate of technology shock persistence. The 12-month deviance ratio is equal to 0.88, a value equal to the average of that in the two datasets of scanner prices. The fraction of two consecutive price changes in the same direction is greater, however, in the model (0.52), than in the data (0.37), suggesting that I over-estimate the persistence of shocks, but this statistic is sensitive to the definition of 'sales' employed.

Given the model's ability to match microeconomic features of the data, I next turn to its aggregate implications. Table 7 (Benchmark, SDP column) reports the volatility and persistence of HP-filtered output in simulations of the model. For comparison, I also report results from a Calvo-type time-dependent model, identical in all respects to the original model, in which firms adjust with constant probability λ, a parameter chosen to match the duration of price spells in the data. The multi-product, menu-cost setup generates business cycle fluctuations from monetary disturbances almost as large as those in the time-dependent model: the standard deviation of output is equal to 0.61% (0.75% in the Calvo setup). Business cycles are equally persistent in the two models: the autocorrelation of output is equal to 0.94.

Standard State-Dependent Pricing Model (Golosov-Lucas (2003))

I next compare the results above to those one obtains in a standard menu cost economy with single-product firms. I abstain from the multi-product features discussed earlier and assume that single-product firms face fixed costs of adjusting their prices. I assume, in addition, that the firm's technology shocks, ε_t, are drawn from a Gaussian distribution with mean 0 and variance σ^2. Three parameters must be calibrated in this model: ξ, ρ, σ^2. I set ρ equal to 0.5, as in the benchmark model[29]. The other two parameters are jointly chosen so that the frequency of price changes, and the mean absolute value of non-zero price changes are equal to those in the data.

As the third column (Golosov-Lucas '03) of Table 5 indicates, the standard menu cost model fails to accord with the micro-price data along two dimensions: it generates no price changes that are less than 4.5% in absolute value and produces a kurtosis of price changes much smaller than that in the data (1.3). I plot, in Figure 4, the distribution of price changes, conditional on adjustment, implied by the standard model, as well as the multi-product model calibrated above. In contrast to the multi-product model, which produces a unimodal, leptokurtic distribution of price changes, similar to that observed in the data, the standard model generates a bi-modal distribution, with no price changes close to zero.

The standard model also generates output fluctuations that are 5 times less volatile than those in a time-dependent model (the standard deviation of HP-filtered output is only 0.15%, compared to 0.73% in a Calvo model). Business cycles are also less persistent (an autocorrelation of only 0.75). These results accord with those of Golosov and Lucas (2003), Caplin and Spulber (1987), and Gertler and Leahy (2005), who find that standard state-dependent models generate, despite nominal rigidities at the firm level, small (if any) business cycle fluctuations from monetary disturbances.

[29]The results we are about to report are insensitive to how persistent technology shocks are.

C. Leptokurtic Shocks or Multi-Product Firms?

I next isolate the role of the two departures from the standard menu-cost model I have introduced and evaluate the contribution of each in generating the difference in results. I first solve a model with economies of scale in the technology of price adjustment in which idiosyncratic shocks are drawn from a Gaussian distribution. The only parameter I calibrate is the size of the menu cost, chosen to ensure the same frequency of price changes as in the Benchmark model. The rest of the parameters are set equal to their calibrated values in the Benchmark model. Notice in the 4-th column of Table 5 that the model fails to match the kurtosis of the distribution of price changes in the data, although it does generate a number of small price changes (21%)[30]. A departure from the assumption of Gaussian shocks is therefore crucial in reproducing the kurtosis of the distribution of technology shocks in the data, at least in this version of the model in which the cost of price adjustment is only spread across two products a firm sells. Table 7 presents this model's aggregate implications: although the volatility of output increases relative to that in models with no interactions in the costs of price adjustment (0.26% vs. 0.15%), business cycle fluctuations are substantially reduced relative to those in a Calvo setup.

I also solve the problem of a single-product firm in which technology shocks are drawn from the distribution assumed in the Benchmark model. Leptokurtic shocks are capable, on their own, to reproduce the kurtosis of price changes in the data. They fail to generate however a large number of small price changes: only 11% of price changes are now less than 4.5% in absolute value. As Table 7 indicates, leptokurtic shocks increase the volatility of output from 0.15% in the Standard model to 0.46%, a significant improvement, albeit smaller than in the Benchmark model. Both interactions in the costs of price adjustment, as well as leptokurtic shocks are thus a

[30] I define, here and in the next calibration, a small price change as a price change whose absolute value is less than 4.5%, the average magnitude of price changes in the Benchmark model.

necessary ingredient of a model capable of reproducing the microeconomic evidence and generating sizable business cycle fluctuations.

D. Discussion

To understand why the two departures I have made from Golosov-Lucas (2003) increase the model's ability to generate real effects from money shocks, consider the impulse responses of the aggregate price level, P, and consumption, C to a 1% shock to the growth rate of the money supply in Figure 5. Notice that the aggregate price level, P, is much more responsive in the Golosov-Lucas setup than it is in a world with economies of scale in price adjustement and leptokurtic marginal cost disturbances. This in turn, given the $M = PC$ constraint I impose, implies that consumption is more responsive in the setup I consider.

The response of the aggregate price level, $\pi = \log \frac{P}{P_{-1}}$, is, to a first-order approximation, the product of two terms: the fraction of adjusters, times the mean price change conditional on a price change:

$$\pi \approx \frac{1}{2} \sum_{i=1,2} \int \pi^i(z) dz = \underset{i,z}{\text{mean}} \left(\pi^i(z) | \pi^i(z) \neq 0\right) \times Fr\left(\pi^i(z) \neq 0\right)$$

where $\pi^i_z = \log(p_{iz}/p_{iz,-1})$ is the price change experienced by good i sold by firm z. Klenow and Kryvtsov (2005) perform this decomposition for the BLS price data and find that most movement in US inflation is associated with fluctuations in the mean price change, conditional on adjustment, as opposed to variation in the fraction of adjusting firms, which is fairly stable over time. Both versions of the menu-cost model I consider are consistent with this feature of the data. As Figure 6 illustrates, the fraction of firms that adjust increases from 22% to 25% on impact in the Benchmark setup and to 24% in the Golosov-Lucas parametrization of the model. In contrast, the mean price change of an adjusting firm is much more responsive to nominal disturbances. Adjusting firms increase their prices on average by 2.5% in the setup I consider, and by 3.5% in the standard menu-cost model considered by Golosov and Lucas (2003). The dampened response of the aggregate price level in the Benchmark model is thus due the fact the adjusting firms, although as numerous as in the Golosov-Lucas parametrization, respond, on average, less aggresively to the aggregate disturbance.

To see why this is the case, let $f(x)$ be the distribution of desired price changes of firms in the economy: $x = \log\left(\frac{p^*}{p_{-1}}\right)$, absent a money shock in the current period, and $h(x)$ be the fraction of firms of type x that find it optimal to adjust, i.e., the hazard. x captures both the contemporaneous productivity disturbance, but also the cumulative history of aggregate and idiosyncratic shocks since the previous price adjustment. Assume, for simplicity, that a money shock has a one-for-one effect on the firm's desired price (more on this below). A firm's desired price change, given the money shock, is therefore $\tilde{x} = x + \Delta m$, where Δm is the monetary disturbance. The change in the price level in this economy is

$$\pi = \int_x dx f(x) h(x) x$$

in the absence of the monetary disturbance, and

$$\pi' = \int_x dx f(x) h(x + \Delta m) (x + \Delta m)$$

if the economy is hit by an aggregate (money) shock of size Δm. The effect a monetary disturbance has on the inflation rate in this economy is thus, rearranging[31]:

$$\Delta p = \int_x dx f(x) \left(h(x + \Delta m) - h(x) \right) x + \Delta m \int_x dx f(x) h(x + \Delta m)$$

The second term in this expression is simply the intensive channel through which the money shock affects the price level: the fraction of adjusting firms times the size of the monetary disturbance. This term is present in state-dependent and time-dependent models alike and captures the fact that all adjusting firms will respond to the increase in the size of the money stock.

The first term is unique to models with menu costs and captures the selection effect. With menu costs, firms optimally choose the timing of their price changes and are more likely to adjust if the money stock reinforces the desire to change prices triggered by idiosyncratic disturbances, captured by x. Firms for which x is positive and desire (prior to the money shock) to increase prices, are more likely to pay the menu cost and adjust if the current money shock is positive and reinforces the impetus to adjust. In contrast, firms for which $x < 0$ are less willing to adjust in times of an expansionary monetary disturbance, as the monetary expansion erodes the firm's real price and offsets the initial incetive to lower the nominal price. Hence, in a menu cost model, the distribution of desired price changes (prior to the aggregate shock), conditional on adjustment, shifts to the right in times of a monetary expansion, thereby increasing the responsiveness of the aggregate price level to monetary shocks.

Figure 7 illustrates the ergodic density of desired price changes, pooled across all goods and firms in the economy,[32] in the Benchmark and Golosov and Lucas setups,

[31] I am indebted to Ariel Burstein, Christian Hellwig and Ivan Werning, as well as to a paper by Caballero and Engel (2006) for pointing out to me the transparency of discussing the intuition behind these results in terms of hazard functions and the density of desired price changes in the increasing hazard region. Derivations similar to the ones above have appeared in the two discussions mentioned above, as well as in Caballero and Engel (2006).

[32] Computed as a kernel density esimate of the distribution of desired price changes (demeaned by

as well as the adjustment hazards for $\Delta m = 0$ and $\Delta m = 2.5\%$. Note first, that the distribution of desired price changes is much more leptokurtic in our Benchmark setup (kurtosis = 13.5) than in the Golosov-Lucas setup (kurtosis = 2.9). Second, notice the difference in hazards in the two setups. In both cases, heterogeneity in productivity across firms and (in the Benchmark setup) in the desired price change of the other good produced by a given firm, smooth out what would otherwise be {0,1} step functions. In the Benchmark setup the hazard is non-negative for all values of x because of the cost complementarities in price adjustment, as opposed to the Golosov-Lucas model in which no firms change prices by a small amount. Finally, notice that a positive money shocks shifts the adjusment hazard to the left. Firms for which x is negative are less likely to pay their menu costs, while those for which x is positive are more willing to reprice.

The equation above illustrates that the strength of the selection effect is larger, the larger is the mass of firms $f(x)$, weighted by their desired price changes, x, in those regions of the parameter space in which the money shock triggers changes in the adjustment hazard: $[h(x + \Delta m) - h(x)]$. Because $f(x)$ declines faster for values of x away from zero for distributions with greater kurtosis, $f(x)$ is lower in the adjustment region (the region where $h(x + \Delta m) - h(x)$ is non-negative) in the Benchmark setup than in the Golosov-Lucas setup. The selection effect is thus weaker, thus explaining why firms, in the impulse responses computed above, adjust, on average, less aggressively to a money shock than they do in the Golosov-Lucas parametrization[33]. Moreover, the cost complementarities assumed in the model smooth out the hazard (it rises slower away from zero), and decrease the strength of the self-selection effect even more.

substracting time-specific means) in a simulation of the model.

[33]Gertler and Leahy (2005) illustrate a similar point using a menu-cost model in which technology shocks affect, in a given period, only a fraction of firms in the economy: the slope of the Phillips curve that their model generates depends on the fraction of firms that are subject to technology disturbances in a given period, or in our language, the kurtosis of technology shocks in the economy.

Figure 8 plots the ergodic distribution of all desired price changes, $f(x)$ (prior to the realization of the money shock) in the two setups considered here, as well as the distribution of price changes, conditional on adjustment, before: $f(x)h(x)$, (first row) and after a money shock $\Delta m = 2.5\% : f(x)h(x + \Delta m)$(second row). Notice that in the absence of the money shock the distribution of desired price changes, conditional on adjustment, is approximately symmetric around zero in both setups. Consider next what happens when the economy is hit by an aggregate money shock. In the Golosov and Lucas setup, where the mass of firms near their adjustment thresholds is large, the selection effect is strong and most firms that do change their prices are firms that need large price increases. Even though few firms adjust, the ones that do are exactly those that need large price increases and as result the aggregate price level is close to flexible. In contrast, although the distribution of desired price changes does shift to the right in the Benchmark setup I consider, the selection effect is much weaker and firm-level price stickiness translates into a more sluggish aggregate price level.

Note finally that the strength of the selection effect is linked to the distribution of non-zero price changes that the model generates. A bi-modal distribution of non-zero price changes, $f(x)h(x)$, with a large mass away from zero and no large price changes arises exactly because $f(x)$ is large in the region of increasing hazard. The Benchmark model considered above reduces the mass of adjusters in the increasing hazard region by allowing $f(x)$ to fall more rapidly away from zero than a Gaussian. Allowing for cost complementarities across a larger number of firms would, by smoothing out the hazard and rendering it flatter, necessitate a less leptokurtic distribution to match the micro-data and to reduce the strength of the selection effect. In contrast, increasing the substitutability of goods produced by a given firm etc. would make the hazard rise more rapidly: an even more leptokurtic distribution of idiosyncratic shock would be required to match the distribution of price changes in the micro-data.

Counterfactual experiments

I next quantify the role of self-selection and synchronization in explaining the aggregate properties of the menu-cost economy. I do so by using two counterfactual experiments, reported in the last rows of Table 7. In the first experiment, I use policy rules optimal in the Benchmark model but assume a constant adjustment hazard. This counterfactual, by holding constant the fraction and identity of the adjusting firms, allows me to gauge the combined role of firm synchronization and self-selection in reducing output variability in menu cost models.

Note in Table 7 that the standard deviation of output in this counterfactual is 2.1 times larger than in the original Benchmark model with multi-product firms. In contrast, shutting down synchronization and self-selection in the standard single-product model generates output fluctuates that are 9 times larger than originally. Clearly, synchronization/self-selection has a much stronger effect in the standard menu-costs models than in a model with multi-product firms and disturbances sufficiently leptokurtic to match the distribution of price changes observed in the data.

I next solve a second counterfactual, in order to pinpoint the exact source of the monetary neutrality in the single-product, Gaussian shocks menu-cost model. I maintain the assumption of a constant hazard and the original policy rules, but allow the fraction of adjusters to vary as in the original simulations of the menu-cost models. Output fluctuations in this second counterfactual are almost as large as in the first one, suggesting that fluctuations in the fraction of adjusting firms in the model play only a limited role, and most of the responsiveness of the aggregate price-level is due to selection.

Note finally that self-selection, although muted in the Benchmark setup I consider, still plays an important role and reduces the volatility of output in half. Why then, does the state-dependent model produce business cycle fluctuations that are of similar magnitude as those in the Calvo model where this effect is entirely absent?

If shocks to the growth rate of the money supply are, as in the data, persistent, a current monetary expansion increases the conditional expectation of the growth rate of the money stock in future periods. Forward-looking firms take into account these forecasts and over-adjust by responding stronger to the nominal shock than they would in a flexible-price world. The willingness of firms to "front-load" expected future increases in the money supply depends however on the size of the frictions that give rise to nominal price stickiness in the first place.

Figure 9 plots the price functions, conditional on adjustment, in the Calvo and menu-cost models, expressed as log-deviations of the optimal price from the one that firms would set in a flexible price world in which prices would respond one-for-one to the monetary disturbance. Clearly, Calvo firms front-load current prices much more aggressively in response to future expected increases in the money supply than state-dependent firms do: a 3% increase in the money growth rate triggers a price increase by Calvo firms that is almost 5% larger than what is optimal in a flexible price world. In contrast, the state-dependent firm's price increases by only 1%.

To understand why state-dependent firms refuse to front-load prices, even though they expect future increases in the growth rate of the money supply, consider Figure 10 in which I plot the two types of firms' values of adjustment and inaction, as a function of one of the two goods' past prices. Note that a state-dependent firm's value of inaction is much less sensitive to deviations of the past price from the optimum than is the value of inaction of a Calvo firm. If a state-dependent firm finds itself with a suboptimal price in a given period, it can always exercise its option to adjust. Its losses are therefore bounded by the size of the menu cost. In contrast, a time-dependent firm pays a hefty price every time its nominal price is suboptimal: because it has to wait for an exogenously fixed number of periods before it gets to reset its price, it will incur much larger losses from a suboptimal price relative to what a state-dependent firm would. A time-dependent firm's incentive to offset future deviations of its price from the optimum is therefore larger than that of a state-dependent firm and it adjusts more aggressively in response to persistent shocks to the growth rate of the money supply[34].

5. Conclusion

This paper has shown that standard single-product state-dependent pricing models are inconsistent with two facts regarding the behavior of individual good's prices: the large number of small price changes and excess kurtosis of price changes in the data. The large number of small price changes can be reconciled with state-dependent models if multi-product firms face interactions in the costs of adjusting prices: I find indeed substantial evidence that prices of products in narrow product categories within grocery stores adjust in tandem.

I then study the general equilibrium properties of a multi-product menu-cost economy calibrated to accord with this micro-economic evidence, and find that the model can, in fact, generate business cycle fluctuations from nominal disturbances that are almost as large as in Calvo-style time-dependent models. A key feature of the calibration, the leptokurtic distribution of idiosyncratic disturbances, implies, together with the assumption of economies of scale in the price adjustment technology, that the selection effect that plays an important role in standard menu cost economies is much weaker in this setup. This, as well as the fact that state-dependent pricing firms are less willing to front-load prices in response to expected future changes in the stock of the money supply, ensures that firm-level nominal price stickiness does translate into sluggish response of the aggregate price level to nominal shocks.

[34] Dotsey, King and Wolman (1999) point out this difference in the optimal price functions of time- and state-dependent firms, but find it quantitatively small in their baseline parametrization with no serial correlation in the growth rate of the money supply and time-varying costs of price adjustment.

References

[1] Barro, Robert, 1972, "A Theory of Monopolistic Price Adjustment," Review of Economic Studies, 39(1): 17-26

[2] Barsky, Robert, Mark Bergen, Shantanu Dutta, and Daniel Levy, "What Can the Price Gap Between Branded and Private Label Products Tell Us About Markups?," mimeo

[3] Benabou, Rolland, "Inflation and Efficiency in Search Markets,"Review of Economic Studies, 59(2): 299-329

[4] Bils, Mark and Peter Klenow, 2004, "Some evidence on the importance of sticky prices," NBER working paper 9069

[5] Bonomo, Marco and Carlos Viana de Carvalho, 2005, "Endogenous Time-Dependent Rules and Inflation Inertia," forthcoming, Journal of Money, Credit and Banking

[6] Burstein, Ariel, 2002, "Inflation and Output Dynamics with State-Dependent Pricing Decisions," mimeo

[7] Butler, J.S., and Rober Moffitt, 1982, "A Computationally Efficient Quadrature Procedure for the One Factor Multinomial Probit Model," Econometrica, 50: 761-764

[8] Caballero, Ricardo and Eduardo Engel, "Explaining Investment Dynamics in US Manufacturing: A Generalized (S,s) approach," Econometrica, 67(4): 783-826

[9] Calvo, Guillermo, 1983, "Staggered Prices in a Utility-Maximizing Framework," Journal of Monetary Economics, 12(3): 383-398

[10] Caplin, Andrew and Daniel Spulber, 1987, "Menu Costs and the Neutrality of Money," Quarterly Journal of Economics, 102(4): 703-725

[11] Caplin, Andrew and John Leahy, 1991, "State-Dependent Pricing and the Dynamics of Money and Output," Quarterly Journal of Economics,106(3): 683-708

[12] Cecchetti, Stephen, 1986, "The Frequency of Price Adjustment: A Study of the Newsstand Prices of Magazines," Journal of Econometrics, 31(3): 255-274

[13] Chevalier, Judith, Anil Kashyap and Peter Rossi, 2003, "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data," American Economic Review, 93(1): 15-37

[14] Danziger, Leif, 1999, "A Dynamic Economy with Costly Price Adjustments," American Economic Review, 89(4): 878-901

[15] Dhyne, Emmanuel, et. al., 2005, "Price Setting in the Euro Area. Some Stylized Facts from Individual Consumer Price Data," ECB working paper

[16] Dotsey, Michael, Robert King and Alexander Wolman, 1999, “State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output,” Quarterly Journal of Economics, 114(2): 655-690

[17] Dutta, Shantanu, Mark Bergen and Daniel Levy, 2002, “Price flexibility in channels of distribution: Evidence from Scanner Data,” Journal of Economic Dynamics and Control

[18] Fisher, Timothy, and Jerzy Konieczny, 2000, “Synchronization of Price Changes by Multi-Product Firms: Evidence from Canadian Newspaper Prices,” Economic Letters, 68: 271-277

[19] Gertler, Mark and John Leahy, 2005, “A Phillips Curve with an Ss Foundation,” mimeo

[20] Golosov, Mikhail, and Robert E. Lucas, Jr, 2003, “Menu Costs and Phillips Curves,” NBER working paper 10187

[21] Hoch, Stephen, Xavier Dreze and Mary E Purk, 1994, “EDLP, Hi-Lo, and Margin Arithmetic,” Journal of Marketing, 58: 16-27

[22] Hodrick, Robert and Edward Prescott, 1997, “Post-War US Business Cycles: An Empirical Investigation,” Journal of Money, Credit and Banking, 29(1): 1-16

[23] Hosken, Daniel and David Reiffen, 2004, “Patterns of Retail Price Variation,” Rand Journal of Economics, 35(1): 128-146

[24] Kackmeister, Alan, 2005, “Yesterday’s Bad Times are Today’s Good Old Times: Retail Price Changes in the 1890s Were Smaller, Less Frequent, and More Permanent,” Finance and Economics Discussion Series, Federal Reserve Board

[25] Kashyap, Anil, 1995, “Sticky Prices: New Evidence From Retail Catalogs,” Quarterly Journal of Economics, February, 245-274

[26] Khan, Aubhik and Julia Thomas, 2004, “Inventories and the Business Cycle: An Equilibrium Analysis of (S,s) Policies,” FRB Philadelphia Working Paper 04-11

[27] Klenow, Peter and Oleksiy Kryvtsov, 2004, “State-Dependent or Time-Dependent Pricing: Does it Matter for Recent US Inflation?” mimeo

[28] Krusell, Per and Anthony Smith, 1997, “Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns”, Macroeconomic Dynamics, 1(2): 387-422

[29] Lach, Saul and Daniel Tsiddon, 2005, “Small Price Changes and Menu Costs,” mimeo

[30] Lach, Saul and Daniel Tsiddon, 1996, “Staggering and Synchronization in Price-Setting: Evidence from Multiproduct Firms,” American Economic Review, 86(5): 1175-1196

[31] Lazear, Edward, 1986, "Retail Pricing and Clearance Sales," American Economic Review, 76: 14-32

[32] Levy, Daniel, Mark Bergen, Shantanu Dutta, and Robert Venable, 1997, "The Magnitude of Menu Costs: Direct Evidence from Large US Supermarket Chains," Quarterly Journal of Economics, 112(3): 781-826

[33] Miranda, Mario and Paul Fackler, 2002, "Applied Computational Economics and Finance," MIT press

[34] Nevo, Aviv, 1997, "Measuring Market Power in the Ready-to-Eat Cereal Industry," Econometrica, 69(2): 265-306

[35] Peltzman, Sam, 2000, "Prices Rise Faster Than They Fall," Journal of Political Economy, 108(3): 466-502

[36] Sheshinski, Eytan and Yoram Weiss, 1977, "Inflation and Costs of Price Adjustment," Review of Economic Studies, 44(2): 287-303

[37] Sheshinski, Eytan and Yoram Weiss, 1983, "Optimum Pricing Policy Under Stochastic Inflation: the Multi-Product Monopoly Case," Review of Economic Studies, 59: 331-359

[38] Sheshinski, Eytan and Yoram Weiss, 1992, "Staggered and Synchronized Price Policies UnderInflation," Review of Economic Studies, 50(3): 513-529

[39] Varian, Hal, 1980, "A Model of Sales," American Economic Review, 70(4): 651-659

[40] Warner, Elizabeth and Rober Barsky, 1995, "The Timing and Magnitude of Retail Store Markdowns: Evidence from Weekends and Holidays," Quarterly Journal of Economics, 110(2): 321-352

[41] Zbaracki, Mark, Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen, 2004 "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets," Review of Economics and Statistics, 86(2): 514-533

Table 1: Distribution of price changes conditional on adjustment

	AC Nielsen	Dominick's
Δp		
mean. %	0.0	1.5
standard deviation, %	13.2	10.4
kurtosis	3.5	5.4
fraction positive changes	0.50	0.65
|Δp|		
mean, %	10.4	7.7
fraction "small" changes	0.28	0.35
# obs.	7560	23029

Notes:

1. p is the natural logarithm of a store's price
2. Statistics are unweighted, based on all non-zero regular price changes, excluding top and bottom 1%.
3. I define a small price change as any change whose absolute value is lower than 1/2 the mean of |Δp|.

Table 2: Ordered Probit Marginal Effects

A. Dominick's

	I.		II.	
	price decrease	price increase	price decrease	price increase
Δ wholesale price	-0.08	0.13	-0.07	0.11
	(0.003)	(0.01)	(0.002)	(0.01)
Δ retail wage	-0.21	0.35	-0.25	0.44
	(0.09)	(0.15)	(0.08)	(0.14)
Fraction of price changes:				
within store			0.02	-0.02
			(0.01)	(0.02)
within product category			0.05	0.29
			(0.004)	(0.01)
withing manufacturer category			0.10	0.45
			(0.003)	(0.01)
% adjustments explained	0.0		17.1	
Log-likelihood	-62301		-52341	
# obs.	151614		151614	

Notes: marginal effects on probability of a price change reported. Standard errors reported in parantheses.
month and product-category dummies included in threshold equations

Table 2: Ordered Probit Marginal Effects

B. AC Nielsen

	I.		II.	
	price decrease	price increase	price decrease	price increase
Δ avg. competitor price	-0.43	0.37	-0.32	0.34
	(0.02)	(0.14)	(0.02)	(0.02)
Δ retail wage	-1.99	1.73	-1.63	1.73
	(0.49)	(0.43)	(0.43)	(0.46)
Fraction of price changes:				
across stores			0.02	0.02
			(0.01)	(0.01)
within store			0.07	-0.07
			(0.02)	(0.03)
within product category			0.34	0.62
			(0.01)	(0.02)
% adjustments explained	0.0		22.5	
Log-likelihood	-18672		-15825	
# obs.	36400		36400	

Notes: marginal effects on the probability of a price change reported. Standard errors reported in parantheses.
month, product-category and store dummies included in threshold equations

Table 3: The frequency and persistence of price changes

	AC Nielsen	Dominick's
Duration of Price Spells, months		
mean	4.0	5.2
median	3	3
iqr	3	6
Persistence of price changes		
Deviance Ratio: 12 months	0.96	0.80
Deviance Ratio: 24 months	N/A	1.02
Prob{sgn(Δp') = sgn(Δp)}	0.32	0.41

Notes:

1. Prob{sgn(Δp') = sgn(Δp)} is the probability that the next price change will have the same sign as the current one, computed as an equally weighted average of the probabilities of two consecutive positive/negative price changes.
2. The Deviance ratios are the mean absolute difference in a good's (detrended) price in periods that are 12 (24) months apart, relative to the mean absolute value of non-zero price changes.

Table 4: Fraction of variance explained by ex-ante heterogeneity

	AC Nielsen	Dominick's
Absolute value of price changes		
store	0.04	-
product	0.14	0.16
month	0.05	0.03
Duration of price spells		
store	0.03	-
product	0.12	0.08
month	0.08	0.10

Notes:

1. Fraction of variance attributed to store/product/month fixed effects reported.
2. Product category × manufacturer, as opposed to upc-specific effects are included in Dominick's data.

Table 5: Calibration targets

Moments:	Targets	Benchmark	Golosov-Lucas '03	Multi-product Gaussian shocks	Single-product Leptokurtic Shocks
Duration of price spells	4.5	**4.6**	**4.6**	**4.5**	**4.6**
mean (|Δp|), %	9.0	**8.8**	**9.0**	7.4	9.9
std (Δp), %	12.0	**11.8**	9.3	8.4	12.5
kurtosis(Δp)	4.5	**4.3**	1.3	1.7	4.0
Pr(|Δp|<mean(|Δp|)/2)	0.30	**0.33**	0.00	0.21	0.11
Deviance ratio: 24 months	1.02	**1.01**	0.84	0.89	1.12
"Over-identifying" restrictions					
Deviance Ratio: 12 months	0.88	0.88	0.82	0.85	1.02
Probability that next price change is in the same direction	0.37	0.52	0.30	0.35	0.53

Note: entries in bold are the moments used to pin down parameter values for the various calibrations

Table 6: Parameter Values

		Benchmark	Standard SDP	Multi-product, Gaussian shocks	Single-product, Leptokurtic Shocks
Parameters not explicitly solved for:					
Common:					
β	discount factor	0.997			
δ	persistence of money shocks	0.79			
σ^2_η	variance of money shocks	$2.05*10^{-5}$			
ψ	marginal disutility from work	2.4			
θ	elasticity of substitution	3			
Calibration specific					
ρ	persistence of technology shocks	-	0.50	0.50	0.50
σ^2_ε		-	-	$2.2*10^{-3}$	-
α_1	Beta(α_1,α_2)	-	-	-	0.05
α_2	Beta(α_1,α_2)	-	-	-	1.3
ε_{max}	upper bound of distb of shocks	-	-	-	0.40
Parameters explicitly solved for					
σ^2_ε	variance of technology shocks	-	$2.2*10^{-3}$	-	-
ξ	menu cost, % of SS labor bill	1.20	0.96	1.00	0.67
α_1	Beta(α_1,α_2)	0.05	-		
α_2	Beta(α_1,α_2)	1.30	-		
ρ	persistence of technology shocks	0.50	-	-	-
ε_{max}	upper bound of shock distribution	0.40	-		

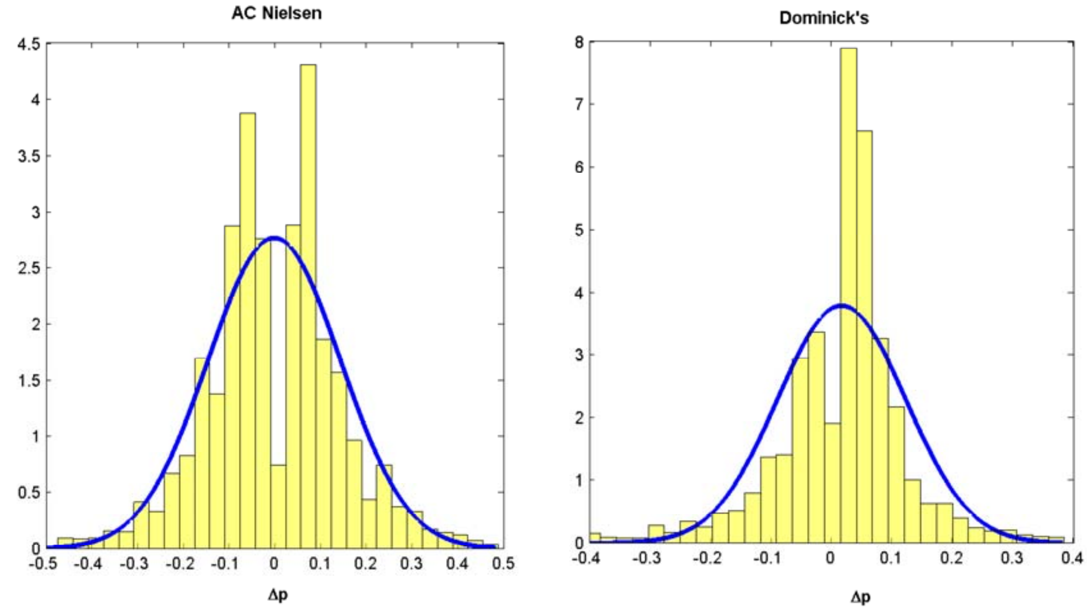
Table 7: Aggregate implications

	Benchmark		Golosov-Lucas '03		Multi-product Gaussian shocks	Single-product Leptokurtic Shocks
	menu-cost	Calvo	menu-cost	Calvo	menu cost	menu cost
Business Cycle Statistics						
$\sigma(y)$	0.61	0.75	0.15	0.73	0.26	0.46
$\rho(y)$	0.94	0.93	0.75	0.92	0.82	0.92
Counterfactual Experiments						
$\sigma(y)$ with Calvo timing (relative to original model)	2.1	-	9.2	-	4.8	3.0
$\sigma(y)$ with no self-selection (relative to original model)	1.8	-	8.6	-	4.4	2.5

Notes:

1. output data was detrended using an HP(14400) filter
2. Δ_t is the average price change of adjusting firms
3. F(ε|adjustment) is the distribution of technology shocks conditional on adjustment

Figure 1: Distribution of price changes conditional on adjustment



Note: superimposed is the pdf of a Gaussian distribution with the same mean and variance

Figure 2a: Fraction of small price changes by product category: Dominick's

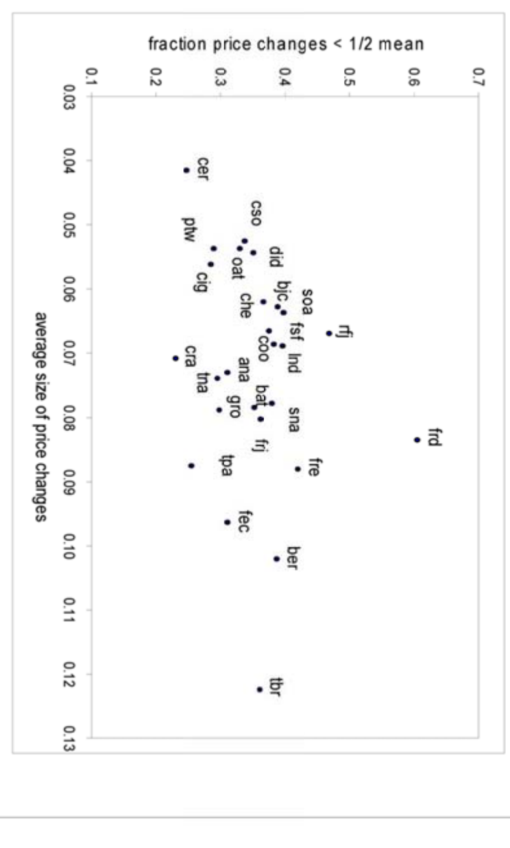


Figure 2b: Kurtosis of non-zero price changes by product category: Dominick's

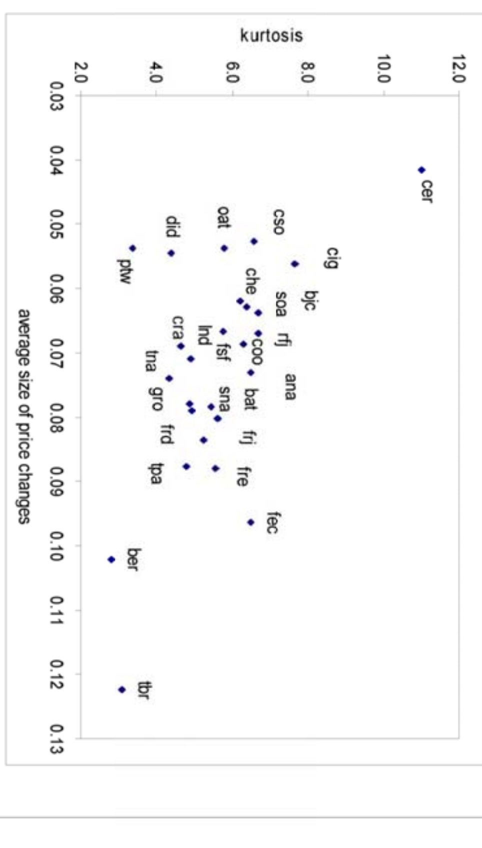


Figure 3: Inaction (Ss) regions for multi-product firms

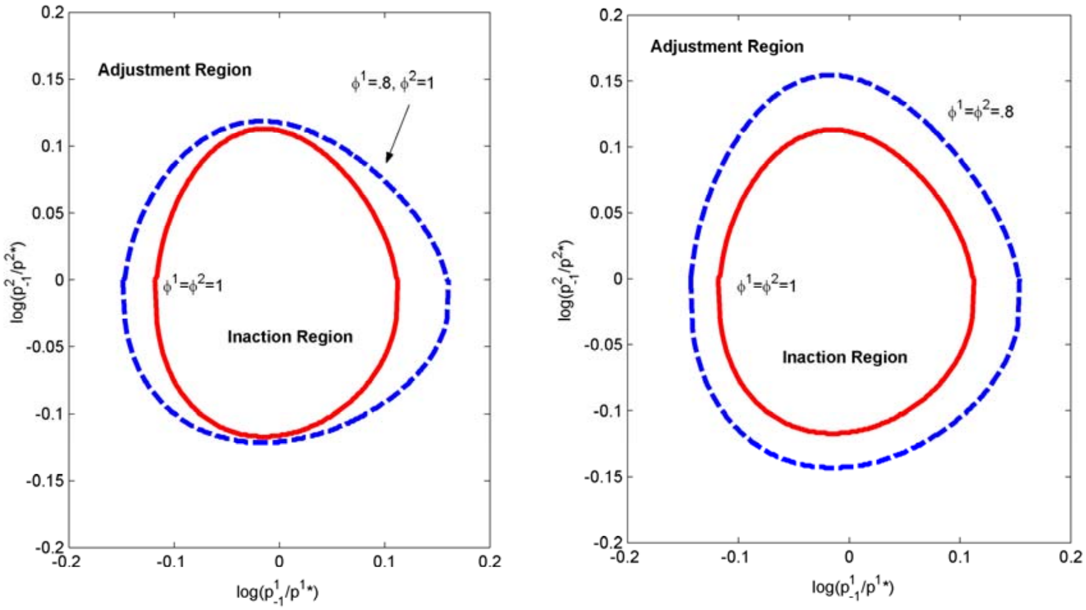


Figure 4: Distribution of non-zero price changes: Model vs. Data

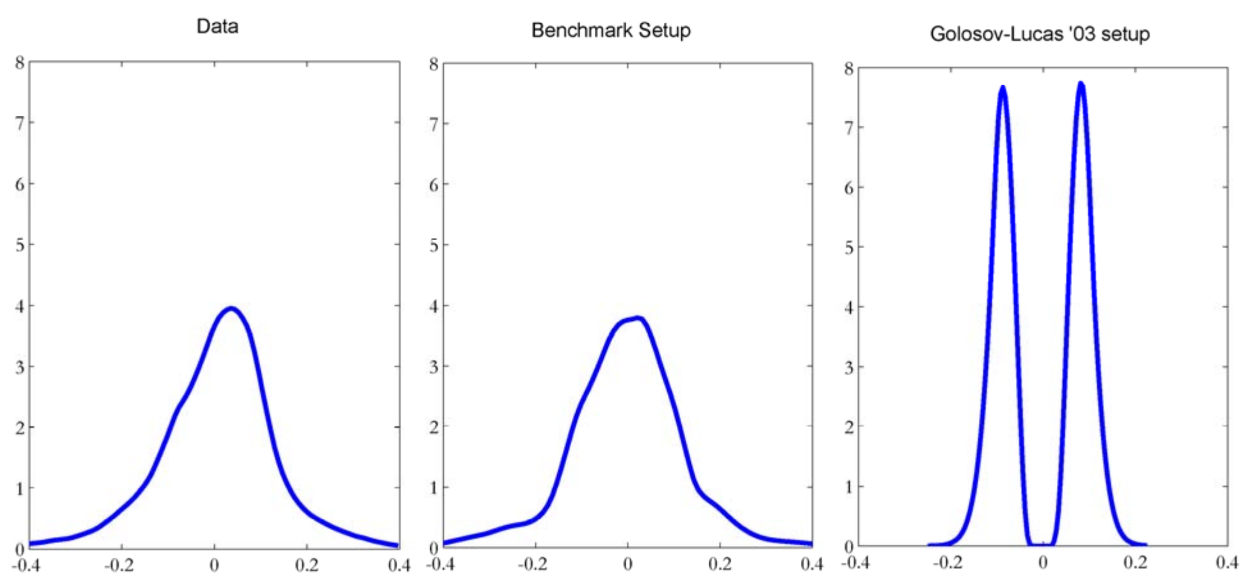


Figure 5: Impulse Response to 1% increase in money growth rate

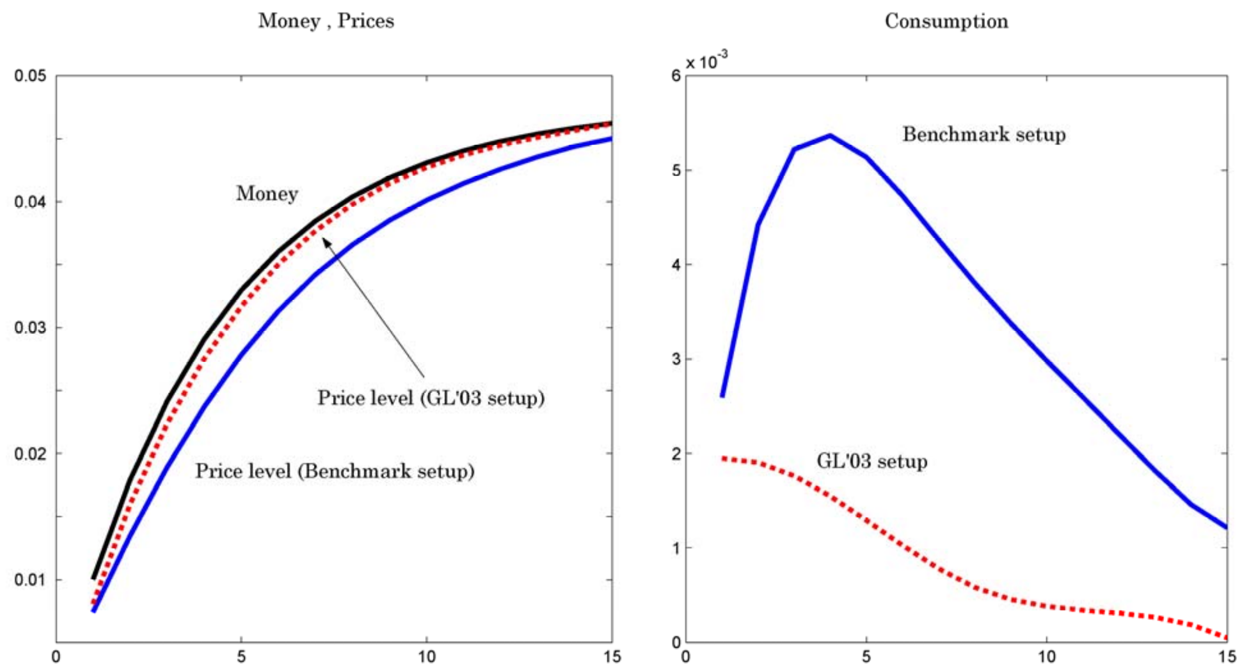


Figure 6: Impulse response to 1% increase in money growth rate

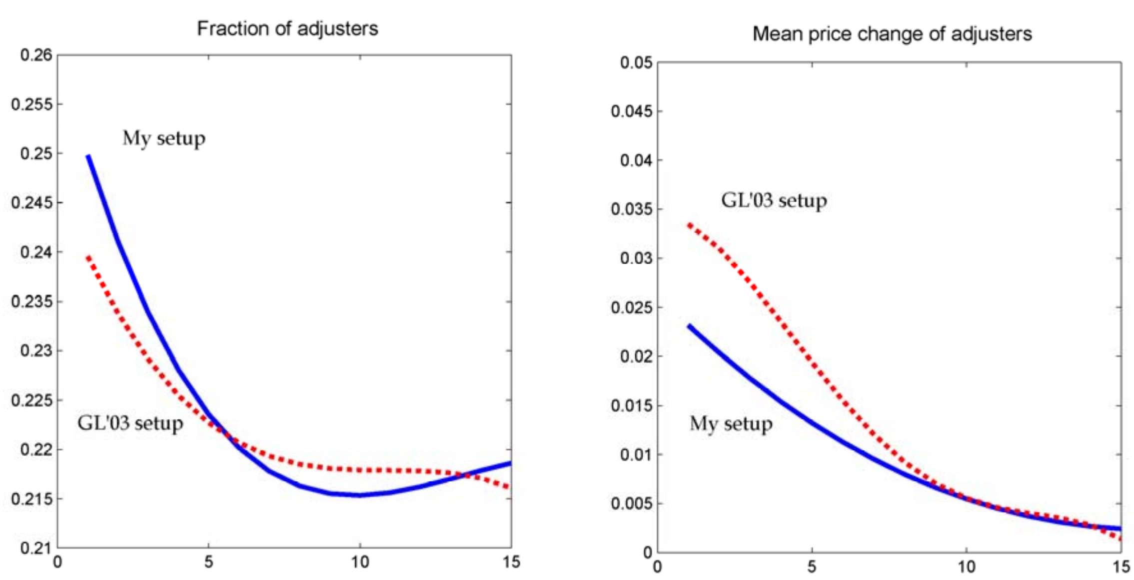


Figure 7: Adjustment hazard and ergodic density of desired price changes

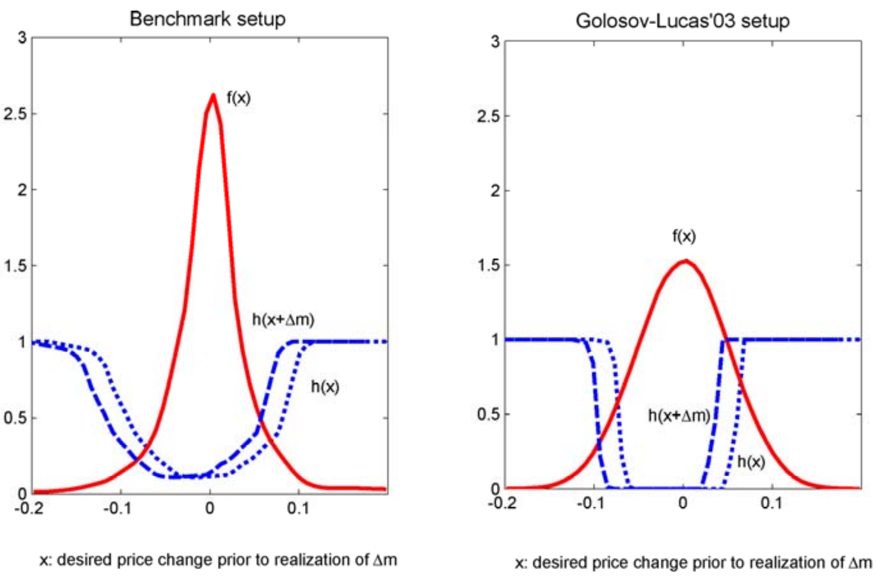
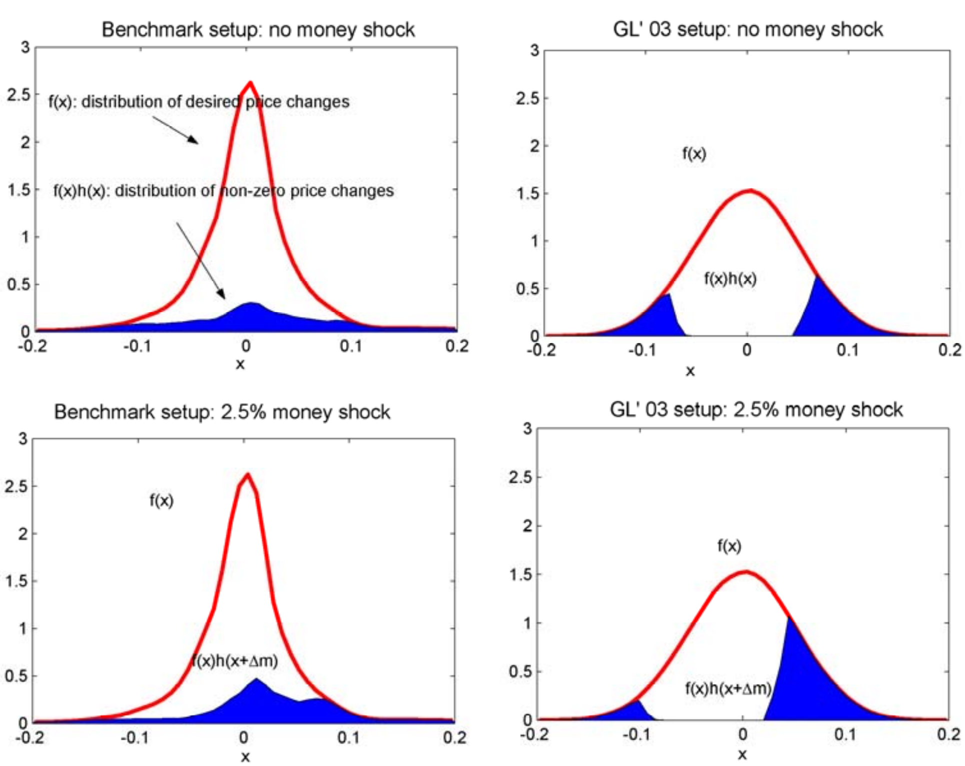


Figure 8: Effect of money shock on the distribution of non-zero price changes



Data Appendix

A. Construction of Price Series

AC Nielsen

I use the household-level data to construct weekly price series according to the following algorithm. For each store/good in the sample, I calculate the number (if any) of units sold at a particular price during the course of the week. If the store sells the product at a single price during the week, I assign this value to the weekly price series. If more than one price is available, the weekly price is the price at which the store sold the largest number of units. In case of a tie in the number of units sold at a particular price, the weekly price is the highest price at which the store sells in a given week[35].

Given that I use scanner price data, price observations for a particular store/good is only available when a customer purchases the product in a particular week. The original prices series are therefore frequently interrupted by gaps. I ignore the gaps if these last for four weeks or less, and use the latest available price before the gap to fill in the price series. Because I study the frequency and size of price changes, I require an uninterrupted prices series. If gaps larger than four weeks are present in the data, I keep only the longest spell of uninterrupted price observations, and discard the rest. I discard stores/goods with less than 100 weeks of continuous price observations.

Dominick's

The Kilts Center for Marketing makes available weekly price quotes for 86 of Dominick's stores. As noted above, this dataset is a by-product of a series of randomized pricing experiments conducted by Dominick's from 1992 to 1993. I only work with the prices of those stores/product categories that were part of control groups in order to avoid treating price changes arising due to experiments as regular

[35]Intuitively, if the number of units sold at two prices is equal, the highest price is likely to have been in effect for a longer time-period as consumers are more likely to buy at the lowest price.

price changes[36].

Dominick's stores are divided into three groups: high, low and medium-price stores, depending on the extent of local competition. Prices within, but also across groups, are strongly correlated, as Dominick's sets prices on a chain-wide basis[37]. Given that gaps in prices are a common occurrence for this dataset, I collapse store-wide prices into a single, chain-wide price, in order to reduce the number of missing observations. I work with medium-price stores only, as these account for the largest share of Dominick's stores. From this set of observations, I let the chain-wide price be the price of the store that has the largest number of observations[38]. For each gap present in this series, I fill in the gaps with the price of another store in the chain, whose pricing most closely resembles the price of the original store, provided that data for this store is available during this period, and the two store's prices coincide in the periods immediately before and after the gap[39]. My metric of the similarity of two stores' price policies is the number of periods in which the two stores set identical prices for a given product. On average 1% of price series are imputed using another store's price. The prices of the stores used to fill in missing data coincide with the price of the original store in an average of 96% of time periods for which data on both stores is available. Once again I discard from the sample those goods for which less than 100 uninterrupted weekly observations are available.

B. Time-Aggregation and Treatment of Sales

Retail prices are characterized by a large number of temporary price mark-downs (sales). Kackmeister (2005) reports that 40% of price changes arise due to sales in a dataset of forty-eight products sold in retail stores during 1997-1999. Hosken and

[36]Hoch, Dreze and Purk (1994) discuss Dominick's experiment in detail.

[37]See Peltzman (2000) for a discussion of Dominick's pricing practices.

[38]Prior to this step, I eliminate gaps smaller than 4 weeks following the algorithm used for the AC Nielsen data.

[39]This last constraint is dictated by our unwillingness to confound changes in stores with changes in the price of a particular store.

Reiffen (2004) find that 60% of the price decreases in their sample of twenty products sold in 30 locations during 1988-1997 are followed by a price increase in the following month. Several hypothesis have been advanced to explain this pattern of retail price variation, stressing informational frictions on consumer's side of the market (Varian 1980), demand uncertainty (Lazear 1986), or thick-market explanations (Warren and Barsky, 1995), to name a few. Instead of incorporating one (or more) of these explanations into the model economy, I follow Golosov and Lucas (2003) and filter out temporary price sales from the price series. This decision leads me to overestimate the importance of nominal rigidities, as I artificially increase the duration of price spells I ask the model to match, but my goal is to compare the performance of two competing sticky price models, rather than compare the models' performance to the data.

Although I eventually time-aggregate the weekly data into monthly observations, I first filter out sales using the original weekly data[40]. I eliminate sales according to the following algorithm[41]. For any price decrease, I check whether this price change is reversed in one of the four weeks following the original price cut. This definition eliminates both V-shaped price changes (price decreases immediately followed by price increases), but also gradual price decreases, provided these are eventually followed by a price increase after at most four weeks following the first price cut. If a sale is deemed to have taken place, I replace the "sales" price with the price in effect in the period immediately before the sale. Figure A1 (left panel) illustrates how the algorithm works. The thin line in the figure is the original price series, while the thick line is the "regular" price. Note for example that the first "sale" was implemented gradually, with the original price decrease followed by no price change in the first

[40]The alternative choice (of time-aggregating the data first and then eliminating sales) can produce spurious price changes if stores periodically put their prices on sale, at regular intervals.

[41]Dominick's dataset includes a "sales" variable that we could in principle use to eliminate temporary markdowns from our price series. This variable is however coded inconsistently and leaves out many temporary price cuts. I therefore choose to eliminate sales manually.

week of the “sale”, an additional price decrease in the second week, and finally a price increase four weeks after the original sale. The drawback of this definition is that it does not eliminate all temporary markdowns in case a price cut is gradually reversed. For example, the final price cut illustrated in the figure was followed by two consecutive price increases, and I have artificially introduced a new “sale” using the algorithm discussed above. To address this problem I repeat the algorithm above an additional three times, in order to eliminate sales that have been gradually implemented. The right panel of Figure 1 illustrates the resulting series of “regular prices” following the last iteration. Note that a single “regular” price change remains after all gradual price reversals are taken into account. Although this remaining price decrease is also reversed, it does not constitute a sale according to our definition because it lasts more than four weeks.

Note a final issue that will play an important role in our discussion of the size and frequency of price changes. By eliminating temporary price cuts, I have introduced an artificial small price change in the regular price series. Any time temporary price reductions are not completely reversed, or followed by price changes larger than the original price cut, setting the regular price equal to the price prior to the sale will artificially introduce a number of small price changes that are otherwise absent in the actual price data. Given that a key statistic in the micro-price data I use to calibrate the menu costs model is the fraction of small price changes in the data, I ignore artificially generated price changes arising due to the filtering of sales and only work with those changes in the regular price that have actually been observed in the original data.

Finally, given that most quantitative studies of sticky price models calibrates them to the monthly or quarterly frequency, I assume a period of one month in the model presented in text. I therefore time-aggregate weekly observations into monthly data, by constructing the monthly series using price data collected in the first week of the month.

Technical Appendix

The typical approach used in solving state-dependent pricing or inventory models is the simulation technique suggested by Krusell and Smith (1997) and applied by Willis (2002) and Khan and Thomas (2004) to models with non-convexities. I depart slightly from the standard method and use a solution technique free of simulations, one that draws heavily on collocation, a residual-based functional approximation method discussed at length in Miranda and Fackler (2002). A simulation-free solution technique used to solve models with heterogeneous agents was originally suggested by DenHaan (1997) in the context of an uninsurable idiosyncratic risks model.

Recall that the firm's problem is to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t \frac{U_c(c_t,n_t)}{U_c(c_0,n_0)} \Pi_t(\tilde{p}_t,)$$

where (for simplicity, I discuss the problem of a single-product firm)

$$\Pi_t(\tilde{p}_t) = \left(\frac{\tilde{p}_t}{\tilde{P}_t}\right)^{-\theta} \left(\frac{\tilde{p}_t}{\tilde{P}_t} - \frac{\tilde{w}_t}{\phi_t \tilde{P}_t}\right) c_t - \xi \frac{\tilde{w}_t}{\tilde{P}_t} \mathcal{I}\left(\tilde{p}_t \neq \frac{\tilde{p}_{t-1}}{g_t}\right).$$

In equilibrium, $\tilde{w}_t$ is constant at the steady-state level because of the preference structure assumed, and $\tilde{P}_t c_t = 1$.The unknown aggregate functions are $c(g, \boldsymbol{\mu})$, aggregate consumption as a function of the growth rate of the money supply, g and $\boldsymbol{\mu}$, the joint distribution of last period's firm prices and current technology, as well as Γ, the law of motion of $\boldsymbol{\mu}$.

Letting $\Xi = \{g, \boldsymbol{\mu}\}$ denote the aggregate state of the world, one can rewrite the firm's problem recursively as:

$$V^{adj}(\tilde{p}_{-1}, \phi; \Xi) = \max_{\tilde{p}} U_c \left(\Pi(\tilde{p}) - \xi \frac{\tilde{w}}{\tilde{P}}\right) + \beta E V(\tilde{p}, \phi'; \Xi') \qquad \text{(V1)}$$

$$V^{nadj}(\tilde{p}_{-1}, \phi; \Xi) = U_c \Pi(\tilde{p}_{-1}) + \beta E V\left(\frac{\tilde{p}_{-1}}{g}, \phi'; \Xi'\right), \tag{V2}$$

where $V = \max\left\{V^{adj}, V^{nadj}\right\}$ is the firm's value, V^{adj} and V^{nadj} is the value of adjustment and not adjustment, respectively and $\Pi(\tilde{p}) = \left(\frac{\tilde{p}}{\tilde{P}}\right)^{-\theta}\left(\frac{\tilde{p}}{\tilde{P}} - \frac{\tilde{w}}{\phi \tilde{P}}\right) c$. $\boldsymbol{\mu}$ evolves according to $\boldsymbol{\mu}' = \Gamma(g, \boldsymbol{\mu})$. The unknowns in this problem are $V^{adj}()$, $V^{nadj}()$, as well as $c()$ and $\Gamma()$. Following Krusell and Smith (1997), I approximate $\boldsymbol{\mu}$ with one moment. In particular, I have found that $\hat{\mu}_t = \int \tilde{p}_{t-1}(z)\phi_t(z)dz$ yields a large degree of accuracy. In the multi-product case $\boldsymbol{\mu}$ is the joint distribution of the two past nominal prices of the firm. I approximate this distribution once again with its first moment: $\hat{\mu}_t = \frac{1}{2}\int \sum_{i=1}^{2} \tilde{p}_{t-1}^{i}(z)\phi_t^i(z)dz$

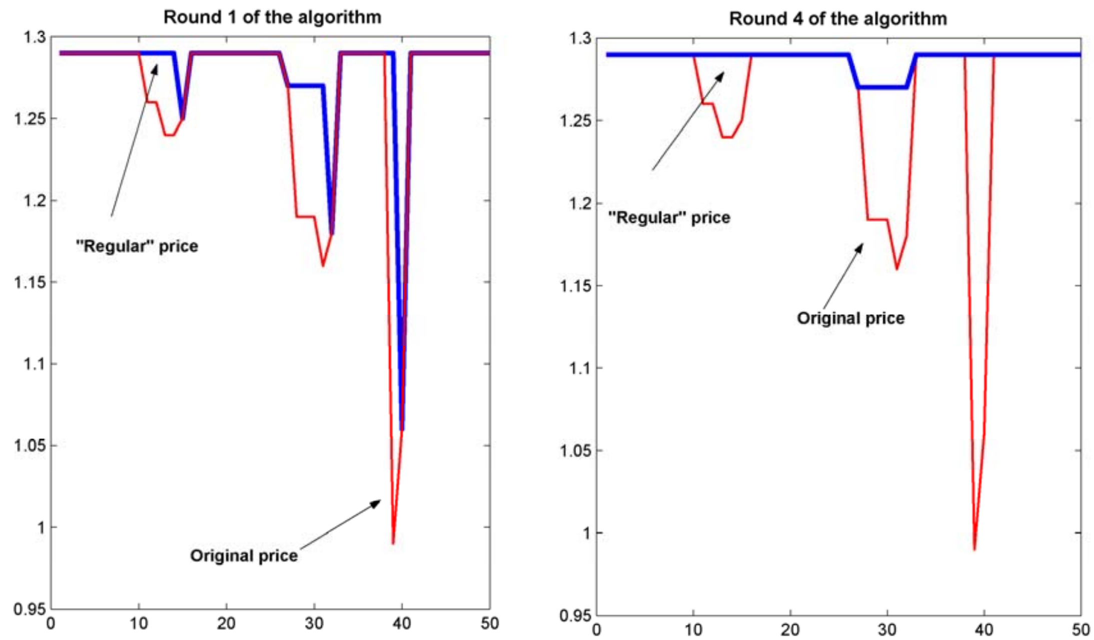
Given initial guesses for $c()$ and $\Gamma()$, I solve the functional equations in (V1-V2) using collocation. Specifically, I approximate each of the two value functions using a linear combination of N Chebyshev polynomials. To solve for the $2N$ unknown coefficients, I require that (V1) and (V2) hold at $2N$ nodes in the state space. This condition yields $2N$ equations I use to solve for the unknown coefficients. I solve the firm's maximization problem in (V1) using a Newton-type routine and evaluate the expectations on the RHS of the Bellman equation by discretizing the distribution of shocks and integrating using Gaussian quadrature.

To solve for the aggregate functions $c()$ and $\Gamma()$, I replace them with a linear combination of Chebyshev polynomials and solve for an equilibrium at each node used to discretize the state-space. For each aggregate node $(g_i, \hat{\mu}_i)$, I solve the firm's problem and recompute aggregate variables c_i and $\hat{\mu}_i'$. To calculate these objects, I need to integrate individual firms' decision rules. Given that I only use one moment of the joint distribution of idiosyncratic states, I assume away all variability in $p_{-1}\phi$ (and also that a firm's past prices are independent of each other). I discretize the cross-sectional distribution of ϕ and calculate, for each mass point in this distribution, the associated p_{-1} consistent with the assumption that $p_{-1}\phi$ is degenerate at $\hat{\mu}_i$ and the law of motion for ϕ_t. For each node $\{p_j,\ \phi_j\}$, I solve firm decision rules and integrate using Gaussian quadrature. Given aggregate quantities c_i and μ_i, I retrieve a new set of coefficients that characterize aggregate functions. For example, letting c be an $M \times 1$ vector of aggregate consumption that satisfy the equilibrium conditions at each node used to discretize the aggregate state-space, Φ be a $M \times K$ matrix of K Chebyshev polynomials evaluated at the M nodes, I find the K unknown coefficients γ_c by solving $\Phi\gamma_c = C$. This set of coefficients for all aggregate variables is used to re-solve the firm's problem, obtain a set of new aggregate variables at each node and calculate a new set of $\gamma = [\gamma_c, \gamma_\Gamma]$, etc.

To evaluate the accuracy of this solution method, I plot, in Figure A2, a time-series of aggregate consumption predicted by the approximant $\hat{c}(\mu, \Gamma)$ for a simulation of stochastic forcing processes, as well as the actual aggregate consumption calculated by integrating firm decision rules. I do so for the Benchmark model, the state-dependent setup. Note that the two series are close to each other: the variability of the actual consumption series explained by the approximant is 94%, suggesting that the additional, higher-order moments that we assume away explain little of the fluctuations of aggregate variables in simulations of the model economy (or comove with those state variables I do keep track of). Aggregate functions are even more accurate than the ones illustrated in the figure in the Calvo and single-product state-dependent models. Moreover, given that the model lacks strong strategic complementarities, firm decision rules are little affected by the mistakes I make in predicting aggregate variables. Indeed, a simulation of the model that assumes that aggregate variables are time-invariant (an assumption made by Golosov and Lucas (2003) generates results virtually undistinguishable from those obtained using the algorithm discussed above.

In Figure A2 I ask how accurate are the solutions to the firm's problem. I plot the left and right-hand side of the Bellman equation in V2, holding constant all other state-variables but $\tilde{p}_{-1}$, at a large number of nodes (larger than that used to pin down the coefficient on the basis functions). Note that the two value functions (predicted and actual), are close to each other. The difference in the two (the residuals) are, as the right panel of Figure A3 indicates, small in absolute value (less than 5×10^{-3}) and oscillate around zero.

Figure A1: Eliminating Sales



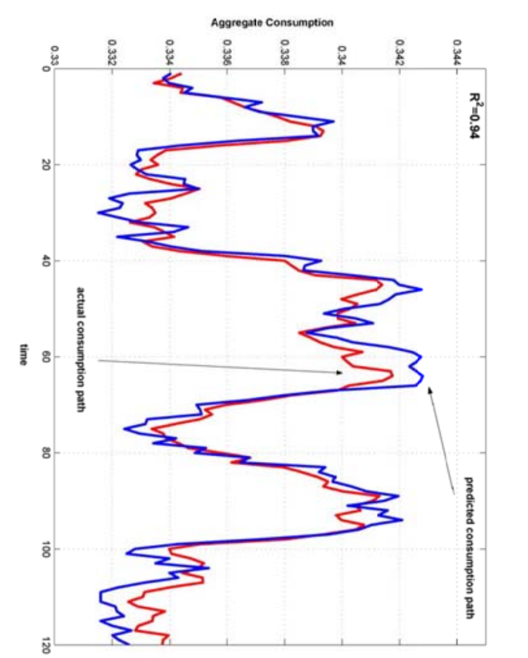


Figure A2: Accuracy of approximants

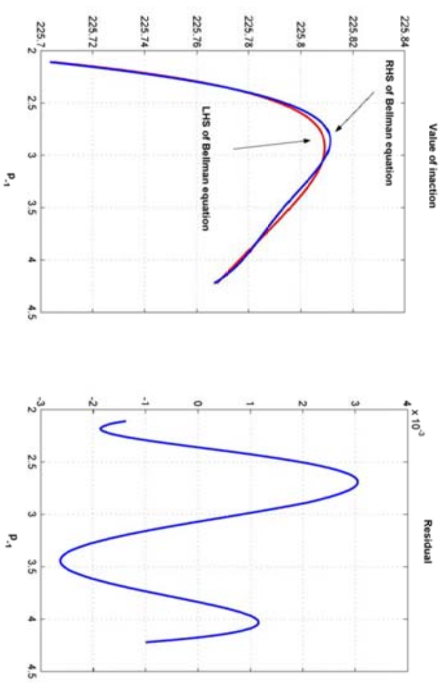


Figure A3: Accuracy of the solution of the firm's problem