Comment

Anil K Kashyap, University of Chicago and NBER

Introduction

This paper brings a new perspective to the debate on the foundations of price stickiness. The authors consolidate and extend an approach pioneered by Caballero and Engel (2007) and introduce new data to evaluate the responsiveness of prices to nominal shocks. While there are many interesting findings, their interpretation of the results rests heavily on an identifying assumption that I find doubtful.

In the remainder of these remarks I begin by laying out two competing visions for the direction that this literature has taken (the dominant one and my own). I then explain my doubts about the critical assumption underlying the authors’ analysis that calls their interpretation of the evidence into question. I close by introducing a few stylized facts that I hope will inform future debates over potential modeling strategies for micro pricing data.

The Microeconomic Foundations of Price Rigidity

Macroeconomists studying price setting seem to have changed their motivations for looking at micro data and trying to model firm level pricing decisions. The Lucas critique led many economists to search for a stable microeconomic model of pricing that could guide their empirical work. The idea, expressed nicely by Romer (1996, 241) is that

[S]ome critics of traditional Keynesian models argue that the models’ assumptions about price stickiness are inconsistent with any reasonable model of microeconomic behavior; they therefore conclude that microeconomic theory provides a strong case against the models’ relevance. More generally, if the conditions needed for nominal stickiness appear implausible or inconsistent with
microeconomic evidence, this would suggest that gradual nominal adjustment is unlikely to be important. If the needed conditions appear realistic, on the other hand, this would support the importance of nominal stickiness.

This perspective guided early work such as Cecchetti (1985), where individual price data were collected and analyzed with the hope of having a stable micro model that could be aggregated to generate macro predictions. It was recognized that a single set of parameters would not describe the behavior of all firms’ pricing of all products, but the hope was that the amount of heterogeneity needed would be modest so that a tractable model might be possible. Unfortunately, as the empirical work on pricing has accumulated this goal has thus far eluded us.

Nonetheless, macroeconomists have remained obsessed with having microeconomic foundations for the macroeconomic models that we use. But the motivation for seeking these foundations seems to have shifted away from Romer’s requirement of plausibility and realism toward the mere possibility that a model can be concocted that has any optimizing foundations at all—even if it poorly fits the data!

For example, in a quest to fit the inertial responses of aggregate prices to changes in interest rates, many of the new Keynesian dynamic stochastic general equilibrium models suppose that firms index their prices during times the prices are not fully reset (e.g., Christiano, Evans, and Eichenbaum 2005). While I appreciate the desire to improve the fit of the macro model, the one thing that we have definitely found from micro studies is that prices are not indexed. So I think we have gone too far in this direction.

One reason why the literature has run into trouble is that the research was spawned by macroeconomists making breakthroughs in the 1980s working under the assumption that the frictions that impede the continuous adjustment of prices arise from the price-setting technology that is available to firms. These seminal contributions relate to models with fixed or quadratic costs of changing prices (such as Caplin and Spulber [1987] or Rotemberg [1982]), as well as the Calvo (1983) pricing model, and in a Taylor (1980) style time-dependent pricing model. But at roughly the same time, microeconomic theorists, who were perhaps more aware of the actual facts about micro price behavior, were developing very different models. These models (such as Sobel [1984] or Varian [1980]) assumed that the central challenge for firms was to devise ways to use prices to discriminate between heterogeneous consumers. Thus, they built models where the action in the models was explicitly driven by variation in customer characteristics.

The subsequent macro literature has largely ignored this distinction
and has been content to work with firm-level pricing frictions. Yet, in the intervening time a great deal of evidence favoring the price discrimination approach has accumulated. I will present some of this evidence in the next section, but it has profound implications for this paper.

In particular, the paper’s central idea is to use the “revelation principle” to learn about optimal prices and price-setting frictions. The revelation principle holds that whenever we see a price change, the new price that is set reveals the statically optimal price given the firm’s various (complicated) constraints. Hence, by focusing on the prices just after they are adjusted we can settle many of the key debates about important pricing questions.

Unfortunately, as they recognize, this is simply not true if price discrimination drives pricing decisions. With a price discrimination motive a firm might opt to change prices in the absence of any shocks, which means that the observed sequence of price changes will be a mixture of responses to shocks and attempts to exploit heterogeneity in demand. This undermines the empirical strategy in the paper.

The Prevalence of Price Discrimination

In many respects price discrimination is so pervasive that most consumers take it for granted. Consider the sequence of transactions that I engaged in to get to this conference. I had to buy an air ticket, which as we all know involves navigating constantly changing prices. I then had to get to the airport. I used a car service, which offers the university a frequent buyer discount, so that payment also involved some price discrimination. Upon landing in Boston I took a cab to the hotel that was set based on an administered price, although in some cities there is a time of day surcharge. Upon arriving at the hotel I got dinner, which came with a menu that offered certain daily discounts. I checked in to my room and was able to obtain the discounted rate that the NBER had negotiated. While eating at the conference, I am guessing that the NBER gets a discounted rate on the catering.

I do not think this kind of discounting is pervasive for the shelter component of the CPI, where we do not normally see this kind of price discrimination. But at least in Chicago during the Great Recession one even heard stories of landlords offering renters a free month for signing a one-year lease.

The point of these anecdotes is not to suggest that prices are fully flexible, but rather to say the prices appear to be set based on customer characteristics and this raises the possibility that pricing strategies over
the seasonal or business cycle might vary. If that were the case, it would matter for monetary transmission.

Fortunately the explosion of micro data sets that have become available in the last decade mean that we can move well beyond this kind of casual empiricism to evaluate this possibility. One of the nice things about the paper is that it exploits the IRI/Symphony data set for much of the empirical analysis. This is a data set that economists are just beginning to study, but it has many appealing features. First, it contains information on retailers in forty-seven market areas, thus it can be used to study competition within markets and across cities with very different market structure. Second, it contains not only the prices paid, but also the quantities purchased at each price, so that it is possible to study substitution responses of consumers across items. Third, it can be merged with information on purchases by individual households, where some data about the characteristics of the household are known. The main drawback of the data set is that it only contains data for seven years starting the early 2000s, so that it does not tell us anything about behavior when inflation is high. As shown by Gagnon (2009), when inflation rises substantially some price-setting rules seem to change.

To show the importance of price discrimination I want to display several facts about the behavior of prices for branded products that illustrate the kind of price discrimination that is pervasive in the IRI/Symphony data set. Figure C1 shows the price (per ounce) of the top selling pair of carbonated beverages at a randomly selected grocery store in Pittsfield, Massachusetts. The solid line represents the price for the best-selling package (a case of 12 ounce cans of Pepsi), and the dashed line shows the price for the second best-selling package in this store (a twelve pack of classic Coca Cola). To make the graph easily readable the data that are displayed covers only March 2004 to December 2006.

The graph shows two very obvious facts. First, Coca Cola is put on discount sale much more often than Pepsi Cola. Second, the temporary discounts of the two products rarely coincide. A formal statistical test rejects the hypothesis that the discounts occur independently at a 2 percent level of confidence. It seems unlikely that anything about the price-setting technology would explain this staggering of the discounts, whereas standard stories about consumer heterogeneity that presume some people who are discount hunters would naturally explain it (Chevalier and Kashyap 2012).

This basic pattern holds for many types of grocery products across most stores. To give a better sense of this, table C1 shows data on top
Fig. C1. Coke and Pepsi prices in Pittsfield, Massachusetts
### Table C1
Actual and Expected Propensities of Temporary Discounts, Various Goods

<table>
<thead>
<tr>
<th>Top 4 24 oz. Bottles of Ketchup</th>
<th>Store 250517 Charlotte</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td>Description</td>
</tr>
<tr>
<td>1</td>
<td>Hunts Ketchup</td>
</tr>
<tr>
<td>2</td>
<td>Heinz Ketchup</td>
</tr>
<tr>
<td>3</td>
<td>Private Label Ketchup</td>
</tr>
<tr>
<td>4</td>
<td>Del Monte Ketchup</td>
</tr>
</tbody>
</table>

Test: Expected Number of Weeks with no sales = Actual, p value = 0.000
Expected Number of Weeks with one sale = Actual, p value = 0.000

<table>
<thead>
<tr>
<th>Top 4 16/18 oz. Jars of Peanut Butter</th>
<th>Store 652159 Pittsfield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td>Description</td>
</tr>
<tr>
<td>1</td>
<td>Skippy Peanut Butter</td>
</tr>
<tr>
<td>2</td>
<td>Peter Pan Peanut Butter</td>
</tr>
<tr>
<td>3</td>
<td>Private Label Peanut Butter</td>
</tr>
<tr>
<td>4</td>
<td>Jif Peanut Butter</td>
</tr>
</tbody>
</table>

Test: Expected Number of Weeks with no sales = Actual, p value = 0.000
Expected Number of Weeks with one sale = Actual, p value = 0.000
### Top 4 Standard Sized Boxes of Facial Tissue
Store 534239 Hartford

<table>
<thead>
<tr>
<th>Rank</th>
<th>Description</th>
<th>No. of UPCs on Sale</th>
<th>Expected % of Weeks</th>
<th>Actual % of Weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Scotties Facial Tissue 2 Ply</td>
<td>0</td>
<td>70.17%</td>
<td>77.71%</td>
</tr>
<tr>
<td>2</td>
<td>Marcal Facial Tissue 2 Ply</td>
<td>1</td>
<td>26.14%</td>
<td>10.84%</td>
</tr>
<tr>
<td>3</td>
<td>Kleenex Facial Tissue 2 Ply (White)</td>
<td>2</td>
<td>3.50%</td>
<td>11.45%</td>
</tr>
<tr>
<td>4</td>
<td>Kleenex Facial Tissue 2 Ply (Assorted)</td>
<td>3</td>
<td>0.20%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Test: Expected Number of Weeks with no sales = Actual, $p$ value = 0.000

Expected Number of Weeks with one sale = Actual, $p$ value = 0.000

### Top 4 16 oz. Margarine Products
Store 653776 Eau Claire

<table>
<thead>
<tr>
<th>Rank</th>
<th>Description</th>
<th>No. of UPCs on Sale</th>
<th>Expected % of Weeks</th>
<th>Actual % of Weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I Can’t Believe It’s Not Butter Vegetable Oil Spread</td>
<td>0</td>
<td>35.66%</td>
<td>37.09%</td>
</tr>
<tr>
<td>2</td>
<td>Imperial Vegetable Oil Spread</td>
<td>1</td>
<td>43.20%</td>
<td>42.38%</td>
</tr>
<tr>
<td>3</td>
<td>Parkay Vegetable Oil Spread</td>
<td>2</td>
<td>18.02%</td>
<td>15.23%</td>
</tr>
<tr>
<td>4</td>
<td>Blue Bonnet Vegetable Oil Spread</td>
<td>3</td>
<td>2.96%</td>
<td>5.30%</td>
</tr>
</tbody>
</table>

Test: Expected Number of Weeks with no sales = Actual, $p$ value = 0.325

Expected Number of Weeks with one sale = Actual, $p$ value = 0.000
selling products at four different stores for four different products. The first panel of the table shows the staggering of sales for the top selling 24 ounce bottles of ketchup at a store in Charlotte, North Carolina. This category is interesting because it is dominated by three national brands (Hunts, Heinz, and Del Monte) and (at least in this store) their discounting strategies are relatively similar; each of the big three offers temporary sales between 15 percent and 26 percent of the time. Yet, the percentage of times when none of the four top sellers is offering a discount is less than would be expected (if the timing of the deals were independent), while the percent of the time when exactly one of the types of the four top sellers is on discount is higher than would be expected. The deviations of the actual percentage of deals relative to the expected are statistically significant.

The second panel of the table shows results for 16 to 18 ounce jars of peanut butter in the same Pittsfield, Massachusetts, store for which the Coca Cola and Pepsi Cola data were taken. Peanut butter also has three large national brands (Skippy, Peter Pan, and Jif), but these three brands account for only about one-third of the total category sales and the discounting propensities are very different across the brands: Peter Pan goes on sale about six times as often as Jif in this store. Nonetheless, the number of times when one of the four top sellers is on sale is still significantly larger than expected, while the number of times when none are being discounted is lower than expected.

The third panel shows another common pattern using data on facial tissues from a store in Hartford, Connecticut. In this category the branded goods are even less important than in the first two. Among the top four selling individual items are two types of Kleenex, which differ in their color. These two items also sell for the same price (including when they go on sale—the assorted colors were missing for a few weeks). These kind of very similar goods are common in many categories: often different “flavors” are priced identically. When estimating demand it is usually impossible to differentiate between such goods because there is not enough price variation to distinguish them. So for many purposes we might want to aggregate items like this, but to keep things simple I am including each of them.

The staggering we saw for peanut butter and ketchup is not present for facial tissue in this store: sales occur much less often than would be expected. The reason for the lack of staggering is that propensity to discount in this category changes dramatically over the calendar. Almost all of the discounts come in March, April, and September, and this
seasonal pattern overwhelms any other considerations. Strong seasonal patterns of consumer demand and discounting is another common pattern for some goods sold in grocery stores (Chevalier, Kashyap, and Rossi 2003).

The last panel in table 1 shows pricing for margarine. The market shares of the leading brands is about as fragmented as with facial tissues, and the propensity for different items to be put on sale is almost as diffused as in peanut butter. For this category there is not much evidence of staggering; this is not uncommon in categories with weak brand presence.

I read table 1 as suggesting that retail grocery prices are often set strategically. There are many other papers that tie this sort of behavior to specific models based on price discrimination. It seems very implausible to me to build a microeconomic model of price setting that rules out this type of behavior for many goods. Unfortunately if this is the case, then the revelation principle will not hold.

Implicit in the foregoing discussion is a distinction between pricing strategies for branded and unbranded or “private label” goods. Historically it has been hard to establish general facts about private label goods since they are typically different across chains. But new data is opening up this area. For instance, in the A. C. Nielsen HomeScan data set, goods are classified according to whether they are branded or private label. This is a national panel data set that tracks purchases of the participating households across a wide range of categories. Between 2004 and 2009 more than 90,000 households were part of the data set and 16.5 percent of the median household’s expenditure was on private label goods. Thus, branded goods are not all that matter for grocery prices.

Private label goods differ from branded goods in several important ways. First, they are almost always substantially cheaper than the branded goods. It is not uncommon for their prices to be lower than a branded good, even when the branded good is on sale. Second, these prices are set without any input from the wholesaler that supplies them. The timing of a Coca Cola or Pepsi discount depends in part of the preferences of Coca Cola and Pepsi and not just the retailer. The retailer can set the price of its house brand of cola on its own. Third, it seems to be the case that there are some consumers who simply will not consider these items. This is easiest to see for drugs, where branded items such as Tylenol and private label acetaminophen are literally identical yet the private label version will sell at a substantial discount.
For all these reasons, the pricing patterns for private label goods differ and categories where private label goods are dominant are often also different. For instance, figure C2 shows the price of the top selling 18 ounce jar of peanut butter at a random store in the western part of Texas near New Mexico, with a market share of about 30 percent; the second most popular peanut butter in this story is also a private label item. The price for this item fluctuates much more than is typical for branded goods. Interestingly, the top selling Skippy and Peter Pan jars each only go on sale 6.6 percent of the time and 4.7 percent of the time, respectively. From table 1, we saw that in the Massachusetts store where the private label peanut butter was much less popular, these brands were put on discount much more often.

I have not seen previous work systematically showing whether branded goods are discounted less often when private label goods have high market share. But this would be exactly what would be expected if one believes that temporary discounts are primarily designed to appeal to customers who are more strategic in their shopping. If such customers are already prone to buy the private label alternative, there is much less of a reason to offer discounts on the branded products.

Kruger (2012) provides evidence suggesting that something like this might be going on. He compares the fraction of branded goods spending on promoted items to total category spending across seventy-three categories of goods; his data are from IRI’s Infoscan Reviews for 2007 (which covers the entire nation). He finds that the simple correlation be-
between branded goods discounting and private label share is weak, but there is a strong correlation between private label share and the relative use of private label discounts compared to branded discounts. This is potentially important to recognize since many studies ignore private label goods and Kruger’s evidence suggests doing so will lead to an interesting type of omitted variable bias: in categories where private label items are popular, discounting activity will be more concentrated on these unbranded items relative to their branded counterparts.

The data in figure C2 also point to another tricky issue that arises in analyzing microdata: How should we determine the “regular” price for an item? The authors use a variant of the Eichenbaum, Jaimovich, and Rebelo (2011, henceforth EJR) proposal for calculating reference price in their analysis. I have calculated discounts by comparing actual prices relative to the Kehoe and Midrigan (2010, henceforth KM) regular price. Figure C3 shows the two alternatives, along with the actual price. The authors’ implementation of EJR simply takes the modal price over a thirteen-week centered window. The KM algorithm is instead an iterative procedure that searches forward and backwards to look for recurring prices. The general point from figure C3 is that it is often difficult to infer regular prices for many private label items.

While neither the modified EJR nor the KM algorithm works particularly well with the tricky private label data, having experimented with these filters on many series, I have found that the KM procedure delivers more reasonable results in most cases. For comparison, figures C4 and C5 show the two regular prices for an 18 ounce jar of Peter Pan
peanut butter at the Pittsfield, Massachusetts, store. In situations like this where discounts are fairly common, using the modal price can deliver a normal price that is much more volatile than the KM price. I believe most people would agree that the KM price looks more like what we are trying to recover via these filtering techniques. I think it would be a step forward if this literature shifts to using the KM algorithm for identifying normal/regular prices.
Implications of Strategic Price Setting for Macroeconomics

The conclusion that a substantial part of price variation is driven by price discrimination motives that derive from heterogeneous consumers implies a number of challenges and questions that macroeconomists working with pricing data have tended to ignore. For instance, inferring the size of demand shocks will often be hard when people substitute across items. If there are some customers who are indifferent between Coke and Pepsi and buy whichever is on sale, then it will make sense for the two goods to be put on sale at different times. However, failure to account for both prices can lead to faulty inference: if Coke prices do not change and Pepsi goes on sale, the quantity of Coke sold may plummet. Absent data on both goods one could easily conclude that idiosyncratic demand shocks are very important for Coke when in fact preferences are stable but both prices matter. For many macro models, calibrating the size of idiosyncratic demand shocks is very important.

Similarly, if price fluctuations depend more on consumer characteristics than firms’ price-setting technologies, then many underexplored questions become important. Perhaps most important is how do customers update their willingness to pay for different products in the face of macroeconomic shocks? Rotemberg (2009, 2010) takes some first steps in this direction, but there are many questions that have received little attention. For example, can central bank communications lead consumers to update their willingness to pay to ratify an inflation target? Do undershooting and overshooting of an inflation target have symmetric effects on consumer behavior? Is deflation special in any way?

One interesting way to pursue some of these questions might be to exploit some of the evidence developed in this paper on aggregate shocks. In particular, the two VAT changes in Mexico are particularly interesting since one was relatively large and presumably impossible to miss, while the other was more modest and perhaps less salient. I look forward to following the authors’ future work on these issues.

Endnotes

I thank the National Science Foundation for a grant administered by the NBER and the Chicago Booth Initiative on Global Markets and Center for Research on Securities Prices for research support in preparing this document. I benefitted greatly from conversations with Judy Chevalier and Günter Hitsch about these issues and from research assistance by Cecilia Gamba and Conor Devitt. Information on my outside compensated activities...
can be found at http://faculty.chicagobooth.edu/anil.kashyap/. As a condition of use, SymphonyIRI reviews all papers using their data to check that the data are not described in a misleading fashion. However, all analyses in this paper based on SymphonyIRI Group Inc. data are my own work, not SymphonyIRI Group Inc.’s, and all other errors herein are also my own. For acknowledgments, sources of research support, and disclosure of the author’s material financial relationships, if any, please see http://www.nber.org/chapters/c12755.ack.

1. The HomeScan data include the following categories of goods: baby needs, cosmetics, cough and cold remedies, deodorant, diet aids, ethnic haba, feminine hygiene, first aid, women’s fragrances, grooming aids, hair care, medications/remedies/health aids, men’s toiletries, oral hygiene, sanitary protection, shaving needs, skin care preparations, vitamins, baby food, baking mixes, baking supplies, bread and baked goods, breakfast food, candy, carbonated beverages, cereal, coffee, condiments, gravies, and sauces, cookies, crackers, desserts, gelatins, syrup, flour, canned fruit, gum, jams, jellies, spreads, juices, canned or bottled drinks, nuts, packaged milk and modifiers, pasta, pet food, pickles, olives and relishes, prepared food dry mixes, prepared food read ted y to serve, salad dressings, mayonnaise, toppings, canned seafood, shortening, oil, snacks, noncarbonated soft drinks, soup, spices, seasoning, extracts, sugar, sweeteners, table syrups, molasses, tea, canned vegetables, dried vegetables and grains, frozen baked goods, frozen breakfast foods, frozen desserts/fruit/lemons, ice, ice cream, novelties, juices, frozen drinks, frozen pizza/snacks/hors d’oeuvres, frozen prepared foods, frozen unprepared meat/poultry/seafood, frozen vegetables, butter and margarine, cheese, cottage cheese, sour cream, toppings, dough products, eggs, milk, pudding, desserts-dairy, snacks, spreads, dips-dairy, yeast, yogurt, deli dressings/salads/prepared foods, fresh meat, deli packaged meats, fresh produce, charcoal, logs, accessories, detergents, disposable diapers, fresheners and deodorizers, household cleaners, household supplies, laundry supplies, paper products, personal soap and bath additives, pet care, tobacco and accessories, paper products, personal soap and bath additives, beer, liquor, wine, automotive, batteries and flashlights, cookware, electronics, records, tapes, floral, gardening, glassware, tableware, hardware, tools, housewares, appliances, insecticides/pesticides/rodenticides, kitchen gadgets, lightbulbs, electric goods, photographic supplies, seasonal goods, soft goods, stationery, school supplies, and toys and sporting goods.

2. I thank Günter Hitsch for this calculation.

3. The Symphony/IRI data only give the metro area where a store is located, and for this relatively rural area it would not be accurate to associate the store with a particular town.

References


Christiano, Lawrence, Charles Evans, and Martin Eichenbaum. 2005. “Nominal


