

Hendel, Igal; Nevo, Aviv

Working Paper

The post-promotion dip puzzle: What do the data have to say?

CSIO working paper, No. 0057

Provided in Cooperation with:

Department of Economics - Center for the Study of Industrial Organization (CSIO), Northwestern University

Suggested Citation: Hendel, Igal; Nevo, Aviv (2004) : The post-promotion dip puzzle: What do the data have to say?, CSIO working paper, No. 0057, Center for the Study of Industrial Organization at Northwestern Univ., Evanston, Ill.

This Version is available at:

<http://hdl.handle.net/10419/38698>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

THE CENTER FOR THE STUDY
OF INDUSTRIAL ORGANIZATION
AT NORTHWESTERN UNIVERSITY

Working Paper #0057

**The Post-Promotion Dip Puzzle:
What Do the Data Have to Say?***

By

Igal Hendel

University of Wisconsin, Madison and NBER

and

Aviv Nevo

University of California, Berkeley and NBER

* We wish to thank David Bell for the data. We gratefully acknowledge support from the NSF and the Food System Research Group at UW.

ABSTRACT

One of the puzzles of store-level scanner data is the lack of a dip in quantity sold in the weeks following a promotion. Such a dip is predicted by a consumer inventory model. During a promotion consumers buy more, not only for current consumption, but stockpile for future consumption. The predictions of such a model have been confirmed by household-level data yet seem harder to find in aggregate brand- or category-level data. We re-examine this puzzle and find two things. First, the effects at the household-level are present, but are much smaller than previously found. Our estimates are different because we control for household heterogeneity in a more general way than most previous work. This suggests that since the effects are small they might be harder to spot in aggregate data. Second, we show that the dip is present in the aggregate data, once we control for additional promotional activity, like feature and display. The latter has an opposing dynamic effect that masks the existence of the post promotion dip.

1. Introduction

Most categories covered in store-level scanner data commonly exhibit temporary price reductions (sales), which naturally result in a large increase in the quantity sold. Since it is reasonable to assume that for many of these products variation in consumption is not large enough to explain this increase, the additional quantity purchased is likely to be stockpiled for future consumption. This has lead researchers to propose a model of household inventory behavior and test it using household-level data (for example, Shoemaker, 1979; Blattberg, Eppen and Lieberman, 1981; Neslin, Henderson and Quelch, 1985; Currim and Schneider, 1991; Gupta, 1988, 1991; Chiang, 1991; Bell, Chiang and Padmanabhan, 1999). Generally, the finding has been that when buying during a sale consumers buy more and accelerate their purchases (i.e., the duration to next purchase is longer and the duration since the previous purchase is shorter).

The household inventory model predicts that in the weeks after a sale, since the households have increased inventory, they will purchase less, holding everything else constant. This suggests that following the (observed) spike in quantity sold during a sale there should be a dip in the aggregate quantity sold. These dips have proven hard to find (Blattberg and Neslin, 1989, 1990; Grover and Srinivasan, 1992; Moriarty, 1985; Neslin and Shoemaker, 1983).

Neslin and Schneider Stone (1996) discuss eight possible arguments to sort out the apparent difference between the household-level data and the aggregate data. Our analysis below sheds light on two of these explanations, reconciling the apparent contradiction between the micro and aggregate findings. First, using panel data methods to control for cross household heterogeneity in a general way we show that household responses, while economically and statistically significant, are smaller than previously reported in the literature. Therefore, it is not surprising that these effects have been hard to detect in the aggregate

data. Second, we show that once we control for the dynamic effect of feature and display the post promotion dip is present in the store-level data.

Our first finding relates to Neslin and Schneider Stone (1996), who also provide (using a different approach) evidence that the size of the effect measured at the household level might be small. They estimate the effect of inventory, constructed under the assumption of constant consumption, on the probability of purchase. We, on the other hand, directly measure household responses to promotions by comparing the increase in quantity purchased and the change in inter-purchase duration time.

The measures we examine have been studied by previous work. The reason our findings differ from this literature is that we decompose the overall, or total, difference between sale purchases and non-sale purchases into two different components. The first component is the difference in an “average” household’s behavior in purchases made during sale versus non-sale periods. We will refer to this as the within (household) difference. The second component, the between (household) effect, is driven by the difference in the composition, or identity, of the households who buy during sale and non-sale periods. Both, within and between differences have interesting economic and marketing implications. However, the within estimator controls for heterogeneity in a more general way and under this general form of heterogeneity only it, the within estimator, consistently estimates the effects of stockpiling.

From a statistical point of view, the within and between differences will be the same if the variation across households is not systematic. In other words, if there is no correlation between a household’s propensity to purchase on sale and a household-specific effect on the variables in question (namely, quantity purchased, and inter-purchase duration). One can test if the difference between the two estimators is systematic. If the difference is systematic, as in our data, then for the purpose of testing a stockpiling theory we have to focus on the within effects. The between variation, although interesting for reasons like price

discrimination, has to be purged from the data as it biases the measurement of the reaction to promotions.

The economics behind the bias is intuitive. Those consumers that purchase more, and consume more, have a bigger incentive to wait for a sale and stockpile, in turn, having stockpiled will purchase less frequently. Such behavior creates a positive correlation between the propensity to purchase on sale and both quantity purchased and inter-purchase duration. Hence, when one compares sale and non-sale purchases, part of the differences is due to a composition effect. Namely, during sale periods quantity sold is higher because more intense consumers buy. The composition effect testifies to the heterogeneity across households who typically purchase on sale as oppose to those who do not.

Our results below suggest that the overall, or total, increase in quantity purchased during a sale and the change in inter-purchase time is similar in magnitude to the effects previously reported in the literature. However, we find that the within differences are smaller and our tests suggest that only they are a consistent estimate of the effect we want to measure. Our conclusion is that since the household level response is smaller than previously estimated the predicted dip in aggregate data might be smaller than expected.

We next turn to the aggregate (store-level) data and show that once we control for the effect of duration from last feature and display, the post promotion dip is present in the store-level data. This result is partly related to the fifth argument provided by Neslin and Schneider Stone (1996): retailers may extend display and/or feature activity beyond the period of the sale. This result is also consistent with other models of a long lasting dynamic effect of feature and display activity (Lattin and Bucklin, 1989). Our findings suggest that the long lasting effects of feature and display, if not accounted for, may be clouding the post promotion dip. This finding is closely related to the results provided in Van Heerde, Leeflang and Wittink (2000). The main difference is that we have to impose little structure on the data and therefore confirm our

results are driven by patterns in the data and not the modeling.

The rest of this paper is organized as follows. We outline a consumer inventory model and intuitively derive its predictions. Next we present the results from the household-level data, followed by the results from the aggregate data.

2. The Model

2.1 Implications of Stockpiling

A model of consumer inventory behavior is both intuitive and not new to this paper. We therefore only outline the elements of the model and the implications. For a formal model see, for example, Arrow, Harris and Marschak (1951), or Blattberg, Eppen and Lieberman (1981).

A consumer obtains utility from consuming the product, which is sold at a “regular” price with occasional temporary price reductions. The product is storable at a cost that is an increasing and convex function of inventory held. When the price is low the consumer balances the benefits from buying at the low price with the cost of holding the inventory. The optimal strategy is to buy when beginning-of-period inventory is below a threshold, which is a decreasing function of prices. In other words when prices are high it is optimal to buy only if the inventory is low, and when the prices are low it is optimal to buy even with a relatively large inventory. Similarly, the end-of-period, or target, inventory is decreasing in prices. If prices are low you will be willing to pay a higher inventory cost so that you can save in the future.

There are several implications of this optimal behavior which have been tested, and which we will examine below. First, the quantity purchased is a decreasing function of price. The standard neo-classical static economic model will also predict this effect: if price goes down consumers consume more. Here the effect exists even if one believes that consumption does not respond at all to prices.

Second, the timing of the purchase is changed. In this model, a purchase is made when the (beginning of the period) inventory falls below a threshold, which is a decreasing function of price. In other words, there is a range of inventory for which the consumer will buy when the product is on sale and not when it is at the regular price. This suggests that all else equal, when buying on sale, the duration to previous purchase should be shorter and the duration to next purchase should be longer (compared to a purchase during a non-sale period).

Third, aggregating consumer-level behavior we predict that the aggregate quantity sold should increase during a sale, both because consumers are buying more and because they are buying earlier. Furthermore, the effect should depend on the duration since the previous sale. The longer it has been since the previous sale the lower on average are the inventories held by consumers. Therefore the larger the quantity sold today. Finally, the quantity sold during non-sale periods increases in the duration from the last sale. The logic is the same as the one during sales.

2.2 Econometric Model and Estimation of Household-level Predictions

In order to test the above predictions we use the following econometric models. For the household panel we estimate

$$y_{it} = \alpha + \beta d_{it} + \gamma_i + \epsilon_{it} \quad (1)$$

where y_{it} are different measures of quantity purchased by household i at purchase instance t , d_{it} is an indicator if the product was on sale, γ_i is a household-specific effect and ϵ_{it} is a disturbance term (which is assumed to be uncorrelated with the right-hand side variables). This is probably the simplest model one could use to estimate the above relations. It can be generalized in many ways. However, this simple framework is rich enough to show how misleading estimates can be if one does not control for

heterogeneity in a general way.

Without the household-specific effects the estimation of equation (1) is straight-forward. There are several ways to estimate this equation once we introduce the household-specific effects. First, we could assume that the γ_i terms are uncorrelated with the indicator d_{it} , in which case OLS estimation of equation (1) yields consistent, but inefficient, estimates of β . These are the so called *Total* estimates (Hsiao, 1986). One can show that the *Total* estimate of β is equal to the difference between the mean of y_{it} during sale and non-sale purchases.

Second, one could construct the *Between* estimator of β , which is obtained from the following regression

$$\bar{y}_i = \alpha + \beta \bar{d}_i + \mu_i \quad (2)$$

where a “bar” denotes the average of the variable over t , and the error term $\mu_i = \gamma_i + \bar{\epsilon}_i$. If the panel is unbalanced (i.e., the number of observations is different for each household) then this equation can be estimated by GLS, where the weights are proportional to the inverse of the number of observations for each household. If we continue to assume that the γ_i terms are uncorrelated with the indicator d_{it} , then the estimates will be consistent, but not fully efficient.

Third, we can estimate equation (1) using GLS. We will refer to these estimates as the *Random Effects* estimates. Under the assumption that the γ_i terms are uncorrelated with the indicator d_{it} , this will yield consistent and efficient estimates (Hsiao, 1986).

Finally, we could relax the assumption that the household-specific effects, γ_i , are uncorrelated with the right-hand side variable. We do this by estimating a different intercept for each household. We can either estimate equation (1) with these additional parameters or estimate, by OLS, the transformation

$$y_{it} - \bar{y}_i = \beta(x_{it} - \bar{x}_i) + v_{it} \quad v_{it} = \epsilon_{it} - \bar{\epsilon}_i \quad (3)$$

Note, that in this transformed equation the household-specific effects no longer appear. Therefore, estimated with OLS it will yield consistent estimates under a wide range of assumptions on the correlation between the household-specific term and the right-hand side variables. We will refer to these estimates as the *Within* estimates.

We note that if the household specific terms are uncorrelated with d_{it} then both the *Random Effects* and the *Within* estimates will be consistent, but only the *Random Effects* estimates will be efficient. On the other hand, under the alternative assumption in which they are correlated, only the *Within* estimates will be consistent. This can be used to form a Hausman test (Hausman, 1978). In section 4 we use this test to evaluate which of the estimators is appropriate for this data. For more details on this test or the above estimators see Hsiao (1986).

2.3 Econometric Model and Estimation of Store-level Predictions

In order to test the aggregate predictions we estimate the model

$$\log(q_{jst}) = \delta_1 \log(p_{jst}) + \delta_2 dur_{jst} + dumvars + \epsilon_{jst} \quad (4)$$

where q_{jst} is the quantity of UPC j sold at store s at week t , p_{jst} is the price, $\log(\cdot)$ denotes the natural logarithm, dur_{jst} denotes the duration from the previous sale (which we define exactly in the next section) and ϵ_{jst} is a disturbance term. The regression also includes feature and display dummy variables as well as store and UPC-specific intercept terms, denoted as *dumvars*. In the specifications below we also include non-linear terms in dur_{jst} . Finally, in some of the specifications we also add the duration from last feature and duration from last display.

We estimate equation (4) using OLS. In principle, since this is an aggregate demand equation one

might worry that the error-term is correlated with some of the right-hand side variables. We control for difference across UPC's and stores by including the relevant fixed effects. We also control for additional promotional activity. In one of the products we examine, soft-drinks, one might worry about seasonality (Chavelier, Rossi and Kayshap, 2003). We, therefore, control for predicted high periods of demand for soft-drinks like 4th of July. What is left in the error-term is hopefully unanticipated demand shocks that are not correlated with the current price, which is set before the demand shock is realized.

The key parameter of interest is the effect of duration. As we noted in the previous section a prediction of the stockpiling model is that quantity demanded should increase with the duration since previous sale. Therefore, we expect that δ_2 will be positive. We expect this to be true both, in the whole sample and when we split it into sale and non-sale periods.

3. The Data

We estimate the above models using data on laundry detergents and soft-drinks categories taken from the Stanford Basket data set (Bell and Lattin, 1998). The data are drawn from two separate metro markets in a large US city and cover a two-year period from June 1991 to June 1993. For the purposes of this paper we do not separate the two markets.² Overall we have 12,673 purchases of laundry detergent and 64,970 purchases of soft-drinks (27,906 purchases of cans of cola or flavored soda, which will be our primary focus) by 1,012 households. The aggregate data comes from eight different stores in these two markets.

For the household panel we know the UPC of the product purchased, the number of units

²See Hendel and Nevo (2002a) for a cross-market analysis.

purchased, if a store was visited and detailed demographics of the household. From the aggregate data we know that total quantity sold for each UPC, the weekly price, feature and display. We also have a description of each UPC, including size and brand information.

In order to perform the above tests we need to define a “sale”. Eyeballing the data it is clear that for most products there is a “regular” price and occasional price reductions. In order to define a sale systematically we need to define a “regular” price and we need to define the threshold for a “sale”. In the results below we define the regular sale as the modal price for each UPC in each store over the 104 weeks of data. A sale is defined as any price at least 5 percent below the regular price. We explored various alternative definitions. First, we explored defining the regular price as the maximum price in the last x weeks, where we varied x between 3 and 6. Second, we define the threshold for a sale at 0, 10, 25 and 50 percent. Some statistics from these different definitions can be found in Hendel and Nevo (2002a). The important thing to note is that the results below are robust to these different definitions.

In order to estimate equation (4) we need to define the duration variable. The main problem is accounting for similar products. For example, if we are looking at a six pack of Diet Coke do we examine the duration from the last time it was on sale? Or do we consider the last time any Diet Coke product was on sale (e.g. 12 pack)? Or do we look if any diet cola (e.g., Diet Pepsi) was on sale? For the results below we measured the duration as the time since any UPC of the same brand was on sale. In other words, for the above example we focused on all Diet Coke UPC’s.

The means and standard deviations of the main variables are provided in Table 1.

4. Results from Household-Level Data

There are fundamentally two predictions taken from the theory that we test using the household-

level data. First, we test whether there is an increase in the quantity purchased during a sale. Next, we test if there is a change in the inter-purchase time. As we pointed out both these predictions have been tested before. Our goal is to show that the way these differences are computed is critical for the economic interpretation of the findings. More precisely, most of the findings in the literature confound two distinct effect, one of which should be purged in order to quantify the effect of promotions. We first present our results, we then discuss how they compare to previous findings, explain the differences and their implications for the post-promotion dip puzzle.

Table 2 presents the results from the household level data. For each of the two categories we examine five variables, namely, what we called y_{it} in equation (1). The first three, quantity, units and size, measure the quantity sold, while the last two measure inter-purchase time. For each of these we present the average during non-sale purchases. A sale is defined as any price at least 5 percent below the modal price of the UPC-store over the entire period.³ The rest of the columns present the difference between the average during non-sale purchases and during sale purchases, which is the parameter β in equations (1) through (3).⁴

This difference is computed four different ways. First, we compute the *Total*, which is the difference between the average of all sale and non-sale purchases. The *Total* difference averages purchases over time and across households ignoring the panel structure of the data. Hence, it reflects two different components: (i) each household's purchases on sale are likely to differ from non-sale ones (a within

³As we discuss in Section 3 we performed several robustness tests to this definition. The results in the table are not sensitive to the different definitions we tried.

⁴Notice the unit of observation is a household purchase. Visits to the supermarket that did not generate a purchase are irrelevant for our analysis. They generate no purchase, the same as a period without a visit to the supermarket.

household difference), and (ii) households differ in their propensity to purchase on sale (a between households' difference). Both, within and between differences have interesting economic implications. However, only the former directly relates to stockpiling. Hence, the latter should be purged in order to isolate the effect of promotions on household behavior, and the implications on aggregate behavior.

We separate these effects next. The second way we look at the difference between sale and non-sale purchases is by computing the *Within* difference. It is equivalent to computing the difference between sale and non-sale purchases for *each* household and taking a (weighted) average across the households. The weights are proportional to the number of observations: households with more observations get a higher weight. The advantage of this statistic is that it controls for heterogeneity across households in a general way. Next, we compute the *Between* difference, which is the slope coefficient in the regression of the household average of the variable being measured (quantity, inter-purchase time, etc.) when regressed on the fraction of purchases the household makes during sales. This statistic measures the variation across households, as opposed to the *within* statistic that measures the effect of promotions on a typical household. Finally, we compute the *Random Effects* estimate. This measure, like the *Total*, uses both the within and between household variation. However, it weights the different sources of the variation optimally.

The first three rows, in each category, of Table 2 present the above statistics for three different quantity measures. Generally, we see that for both categories the quantity (i.e., units*size) increases during sale purchases, as predicted by the theory. The magnitude of the effects is larger (in percentage terms) for soft-drinks. The *Within* and *Between* effects are positive and bigger for soft-drinks as well. Recall that the *Within* effect measures how much more the “average” household buys on sale relative to their non-sale purchase. The *Between* effect, on the other hand, captures the difference in the quantity purchased

between a household that typically buys on sale and a household that typical does not buy on sale. Notice that for both detergents and soft-drinks the *Within* effect is smaller than the *Between* effect.

Is the difference between the *Within* and *Between* effects statistically significant? If the *Within* is statistically different from the *Between* effect (and hence, from the *Total*) then one cannot rely on the *Total* difference to infer stockpiling. In this particular case, since the *Between* effect is larger than the *Within*, one would overestimate the effect of stockpiling. It is not that buyers expand their purchases during sales, but rather that more intense buyer are those that buy more frequently on sale. Consumer heterogeneity would lead us to overstate the individual responses to sales. In order to test this we compute the so called *Random Effects* estimator and use it to compute a Hausman test.⁵ Even though the results for the *Random Effects* estimator are similar to the *Within* estimates, the statistical test rejects the hypothesis that they are the same. Therefore, only the *Within* estimators are consistent.

Studying the effects on *Size* and *Units* we learn two things. First, the increase in total quantity purchased comes from different sources in each product category. In detergents consumers buy more units and larger sizes, while soft-drinks are purchased in fewer but bigger containers during sales. In Hendel and Nevo (2002a) we examine these differences across the categories more carefully.

Second, we once again notice the difference in the *Within* and *Between* estimates. We again computed the *Random Effects* estimator and reject that hypothesis that the household effects are uncorrelated with the propensity to purchase on sale. This, in a way, should not be surprising since we know that households vary in their tendency to purchase on sale and the factors determining this tendency could also be correlated with the quantity purchased. The implication is that the *Total* difference is not a

⁵We also computed a Breusch and Pagan (1980) Lagrangian multiplier test. The results were the same.

good measure of consumers' responses to promotions as it mixes responses with composition effects.

While the findings on quantity purchased are consistent with the predictions of the stockpiling theory, they are also consistent with alternative theories. Therefore, we follow past work and also examine inter-purchase duration differences. The last two rows, for each category, in Table 2 present the results. The two variables measure the number of days between the current purchase and either the previous purchase or the next purchase. The duration measured, is the duration between the purchases of any product in these categories.

Like the results for the quantity variables we find a large variation between the *Total*, *Within* and *Between* estimators. Moreover, a Hausman test again confirms that only the *Within* estimates are consistent. In the case of *Duration to previous purchase* the decomposition of the *Total* effect is even more important. Not only do the effects differ in magnitude but they even have opposing signs. The *Within* estimate confirms the predictions of the theory, duration from previous purchase should be shorter during a sale purchase. Interestingly, if we were to rely on the *Total* difference we would incorrectly conclude that the data contradicts the theory. The *Total* difference in duration backwards is large and positive. However, after decomposing the *Total* difference we see that there is no contradiction with the theory. The positive *Total* difference is dictated by a large *Between* difference, which dominates the negative *Within* difference.

Notice that the opposing signs of the *Between* and *Within* estimates are consistent with each other. A negative within effect confirms the theoretical prediction, that sales accelerate purchases. On the other hand, a positive between difference testifies to household heterogeneity of the following form. The inter-purchase duration is longer for households who typically purchase on sale than the duration of households who typically do not buy on sale. In other words, on average households who tend to buy more on sale also tend to buy less frequently, which is consistent with an inventory model. Sale-prone households buy

less often, but purchase a much larger quantity. The between quantity difference for soft-drinks is 144% ($7.20/5.00$), while the between difference in duration is about 100%, suggesting that on average typical sales-buyers consume 40% more than households who do not typically buy on sale. Since they are more intense buyers they gain more from stockpiling.

This cross-household pattern, although consistent with a stockpiling theory, is not the effect we are trying to measure. It does not tell us whether the average household is buying earlier because of the promotion. It tells us that buyers are heterogeneous, which is interesting because it testifies about the firms' incentives to have sales. The numbers suggest differences across households; intense users are more willing to wait for a promotion. Hence, cyclical prices (sales) help the seller discriminate between the less intense buyer, who more frequently will end up paying the modal price, and the more intense buyer who is willing to wait for promotions. Notice that one sort of heterogeneity, consistent with our findings, is the presence of deal prone and non-deal prone buyers. Where the former buy from sale-to-sale, while the latter are not sensitive to promotions. In its extreme form, this heterogeneity would lead to large total effects due to between differences, but no within effects of promotions.

The *Within* estimate is instead the true test of the theory. All the *Within* effects are statistically significant but are relatively small, compared to the *Total* difference. They predict roughly a change of 5-10 percent in the inter-purchase duration. As a side point we note that the *Within* estimates generally predict larger effects on quantity than on duration, which can be interpreted as evidence of consumption effects. If the consumption rate did not change one would expect the effects to be the same: if you buy 25 percent more it should take you on average 25 percent longer till your next purchase.

With the exception of Shoemaker (1978)⁶, the rest of the literature that tested these predictions provided either a *Total* estimate or some form of a random-effects estimator (not necessarily the efficient version). Blattberg, Eppen and Lieberman (1981) compute what we call the *Total* estimate for four categories. The magnitude they report (in their Tables 2 and 3) is similar to our *Total* estimates. For example, they find that in the four categories the increase in duration to next purchase is 23-36 percent. Our *Total* estimates are 20-60 percent, much larger than our *Within* estimates. Neslin, Henderson and Quelch (1985) using very similar data find acceleration of purchases, for coffee and bathroom tissue, as a reaction to advertised price cuts. They test the impact of promotions through regressions of: quantity purchased and time elapsed since the previous purchase on a proxy for inventory (derived from past behavior) and dummies for different promotional activities. They find substantial effects of promotions on duration to next purchase as well as on quantity purchased. Consistent with the previous literature they report total effects, without decomposing the household response from the composition effect. Subsequent work (Currim and Schneider, 1991; Gupta, 1988, 1991) find effects of the same magnitude as Blattberg, Eppen and Lieberman (1981). Neslin and Schneider Stone (1996) summarize that the estimates of accelerated purchase are between 14 and 50 percent, consistent with our *Total* estimates but again somewhat larger than our *Within* estimates.

Once we isolate individual responses to promotions we find that while the effects are still present they seem to be smaller than originally thought. This can explain why they might be hard to find in aggregate data. This explanation has been suggested by Neslin and Schneider Stone (1996). In their empirical analysis they estimate the effect of inventory, constructed under the assumption of constant consumption,

⁶While Shoemaker (1978) does not call his estimate a “within” estimate, his procedure is identical to what we call *Within*. Indeed, like him, we also find relatively small effects.

on the probability of purchase. Our findings add in several ways. First, our analysis requires fewer assumptions. For example, we do not need to make assumptions to generate the unobserved inventory series. Indeed, a small estimated effect of inventory could be driven by measurement error in the construction of the inventory series, a problem we do not face. Second, our analysis decomposes the overall, or total, difference between sale purchases and non-sale purchases into two different components, the within and between effects. While we claim that only the within effects truly measures the importance of stockpiling, both the within and between effects have interesting economic and marketing implications.

5. Results from Store-Level Data

In the previous section we showed that the household-level estimates are consistent with the inventory model but much smaller than previously found. Therefore, it is likely that the effects will be harder to find in aggregate data. In particular it is likely that the expected post-promotion dip is masked by other effects. In this section we show that indeed for our data the post-promotion dip is initially missing. However, once we allow for a simple control for other dynamic effects, a post-promotion dip is detected

Table 3 presents the results of regressing the log of quantity sold, measured in 16 ounce units, as a function of price, measured in dollars per 16 ounce, current promotional activity and duration since previous promotional activity. Each observation is a brand-size combination in a particular store. For each category we look at four regressions. The first two use all the sample. The last two include only non-sale observations.⁷ As before, a sale is defined as any price at least 5 percent below the modal

⁷We do not present the results for the sale sample. They generally give us the expected result even before we control for the additional duration, and become larger once we do control for the additional variables. Pesendorfer (2002) also reports results for the sale sample and finds the expected sign. His focus is not an inventory model and therefore

price for that product in that store. Within each sample we present two regressions that differ in the controls. Duration from previous sale/feature/display is measured as number of weeks, divided by 100, from previous sale/feature/display for that brand in that store for any size.

In all the regressions the price, feature and display coefficients are of the expected signs. The key coefficient, from the viewpoint of a stockpiling model, is the coefficient on the effect of duration. As explained in Section 2 we expect this coefficient to be positive: the longer since the last sale the higher the demand, everything else constant. A post-promotion dip will be reflected by a positive coefficient on duration from last sale.⁸

The first regression in each sample finds that for detergents (and other categories not reported in the table) the effect of duration is negative, i.e., rather than a post-promotion dip we find a boost in sales. This effect is true for the all sample and for the non-sale sample. It is the opposite of expected and is consistent with findings in the literature. Indeed this is the basis for the so called post-promotion-dip-puzzle. The effect for soft-drinks is positive instead.

In the next column in each sample we add controls for the duration from previous feature and display. We speculate that both display and feature have dynamic effects that are of opposite sign to the effect we are after. For example, feature or display generate awareness of the product, and could impact demand (positively) even after they are over. That is, it is reasonable to assume the effects of these activities shifts out demand in later periods as well. This explanation is consistent with many neo-classical theories of “advertising” (see for example Bagwell (2002)) or behavioral economic theories. Our purpose here is not to separate these theories, or to test them, just to point out that they have a different effect than that of

he does not examine the non-sale sample.

⁸The regression also includes a square term to control for non-linear effects. Within the range of the observed data the marginal effect is determined by the linear term and therefore for the rest of the discussion we focus on this term.

a sale. Once we control for their effect we can recover the post-promotion dip. Since we have previously established that the effects we are after are small it is likely that they will be masked by the counter effects of feature and/or display. They will be masked because during a sale the probability of feature and/or display increases, but also because there could have been a feature and/or display a few weeks prior to the sale.

By adding duration from feature and display we control for these effects. Indeed once these controls are included the coefficients on duration from sale all become positive, as expected. For the soft-drinks category they increase substantially in magnitude. The magnitude of the effects also makes sense. They are largest for soft drinks, where the demand right after a sale is predicted to be almost 7 percent lower than 4 weeks later. For detergents the demand right after a sale is roughly 3 percent lower than 4 weeks later.

The explanation we provide above has been suggested by Neslin and Schneider Stone (1996). Van Heerde, Leeflang and Wittink (2000) propose a “sophisticated distributed lag analyses of weekly sales data in the hope of measuring the post-promotion dip statistically.” Indeed their model shows that adding various lags and leads can help to find the dip. Our analysis is closely related. However, we impose little structure on the data and show that a simple control is enough to uncover the effects.

6. Conclusions

In this paper we investigate one of the mysteries of scanner data: the lack of a post-promotion dip in store-level data. Previous work, using household-level data, found an increase in quantity and acceleration in the timing of purchase. On the other hand, dips in aggregate weekly scanner data have been hard to find. We re-examine two explanations that have been previously proposed. First, by

applying panel data methods we show that the magnitude of the effects previously reported from household-level data were too high. Second, we show that a simple control for the duration since last feature/display is sufficient to recover the effects at the aggregate level.

Our main findings, for the data we examine, are as follows:

- The effects at the household level are smaller than previously reported. For example, we find that timing is accelerated between 5 and 10 percent, compared to 14 to 50 percent found by previous work.
- This provides support to the conjecture made by Neslin and Schneider Stone (1996) that the effects might be smaller than expected and therefore hard to detect.
- Once we control for duration from previous feature/display, a post-promotion dip can be detected in the weekly store-level data. This can be found using a simple linear regression and does not require more elaborate models of lags and leads.

Our analysis is not a substitute for a structural model of the effects of promotions. In order to answer many of the interesting questions regarding dynamic consumer stockpiling behavior one needs a structural dynamic model like the one proposed and estimated in Erdem, Imai and Keane (2003) or Nevo and Hendel (2002b). Since our findings are consistent with stockpiling they suggest that researches estimating demand should control for current inventories or duration from last purchase.

References

- Arrow, K, T. Harris, J. Marschak (1951). "Optimal Inventory Policy," *Econometrica*, Vol. 19, No 3, 250-272.
- Bagwell, K., (2002) The Economic Analysis of Advertising, forthcoming *Handbook of Industrial Organization*.
- Bell, D., J. Chiang and V. Padmanabhan (1999), "The Decomposition of Promotional Response: An Empirical Generalization," *Marketing Science*, 18, 504-26.
- Bell, D. and J. Lattin (1998), "Shopping Behavior and Consumer Preference for Retail Price Format: Why 'Large Basket' Shoppers prefer EDLP," *Marketing Science* 17, 66-88.
- Blattberg R., G. Eppen and J. Lieberman (1981), "A Theoretical and Empirical Evaluation of Price Deals in Consumer Nondurables," *Journal of Marketing*, 45, 116-129.
- Blattberg, R. and S. Neslin (1989), "Sales Promotion: The long and Short of It," *Marketing Letters* 1, 81-97.
- Blattberg, R. and S. Neslin (1990), *Sales Promotions*, Prentice Hall.
- Breusch, T. And A. Pagan (1980), "The Lagrange Multiplier and Its Implications to Model Specification in Econometrics," *Review of Economic Studies*, 47: 239-253.
- Chevalier, J., A. Kashyap and P. Rossi (2000), "Why Don't Prices Rise During Peak Demand Periods? Evidence from Scanner Data," *American Economics Review*, 93, 1, 15-37.
- Chiang J. (1991), "A Simultaneous Approach to the Whether, What and How Much to Buy Questions," *Marketing Science*, 10, 297-315.

- Currim I. and L. Schneider (1991), "A Taxonomy of Consumer Purchase Strategies in a Promotion Intense Environment," *Marketing Science* 10, 91-110.
- Erdem, T., S. Imai and M. Keane (2003), "Consumer Price and Promotion Expectations: Capturing Consumer Brand and Quantity Choice Dynamics under Price Uncertainty," *Quantitative Marketing and Economics*, 1, 5-64.
- Grover R. and V. Srinivasan (1992), "Evaluating the Multiple Effects of Retail Promotions on Brand Loyal and Brand Switching Segments," *Journal of Marketing Research*, 29, 76-89.
- Gupta S. (1988), "Impact of Sales Promotions on When, What, and How much to Buy," *Journal of Marketing Research*, 24, 342-55.
- van Heerde H., Leeflang P. and D. Wittink (2000), "The Estimation of Pre- and Postpromotion Dips with Store-Level Scanner Data," *Journal of Marketing Science*, 37, 383-395.
- Hausman, J. (1978), "Specification Tests in Econometrics," *Econometrica* 46(6), 1251-71.
- Hendel, I. and A. Nevo (2002a), "Sales and Consumer Inventory," NBER Working Paper no. 9048.
- Hendel, I. and A. Nevo (2002b), "Measuring the Implications of Sales and Consumer Stockpiling Behavior," University of California, Berkeley, mimeo.
- Hsiao, C. (1986) *Analysis of Panel Data*, Cambridge University Press.
- Lattin J. And R. Bucklin (1989), "Reference Effects of Price and Promotion on Brand Choice Behavior," *Journal of Marketing Research* 26, 299-310.
- Moriarty, M. (1985), "Retail Promotion Effects on Intra- and Inter-brand Sales Performance," *Journal of Retailing* 6, 27-48.

Neslin S., C. Henderson and J. Quelch (1985), "Consumer Promotions and the Acceleration of Product Purchases," *Marketing Science*, 4(2), 147-165.

Neslin, S. and L. Schneider Stone (1996), "Consumer Inventory Sensitivity and the Postpromotion Dip," *Marketing Letters*, 7, 77-94.

Neslin, S. and R. Shoemaker (1983), "A Model for Evaluating the Profitability of Coupon Promotions," *Marketing Science* 2, 361-388.

Pesendorfer, M. (2002), "Retail Sales. A Study of Pricing Behavior in Supermarkets," *Journal of Business* , 75, 33-66.

Shoemaker R. (1979), "An Analysis of Consumer Reactions to Product Promotions," in *Educators' Conference Proceedings*, Chicago: American Marketing Association, 244-248.

Table 1
Descriptive Statistics

Variable:	Laundry detergents		Soft-drinks	
	mean	std	mean	std
Quantity (16 oz.)	98.2	361.3	247.2	1056.0
Price (\$/16 oz.)	1.28	0.45	0.79	0.23
Sale	0.16	-	0.40	-
Feature	0.04	-	0.18	-
Display	0.07	-	0.11	-
Sale and Feature	0.04	-	0.16	-
Sale and Display	0.04	-	0.09	-
Sale and Feature and Display	0.02	-	0.06	-

A sale is defined as any price at least 5 percent below the modal price, of a UPC in a store over the observed period.

Table 2
Differences in Household Purchasing Patterns Between Sale and Non-Sale Purchases

Variable:	Avg during non-sale		Difference during sale							
	Total		Within	Between	Rand Effects					
Laundry Detergents										
Quantity (16 oz)	4.79	(0.04)	1.55	(0.07)	1.14	(0.07)	2.65	(0.28)	1.20	(0.07)
Units	1.07	(0.01)	0.09	(0.01)	0.08	(0.01)	0.13	(0.03)	0.08	(0.01)
Size (16 oz)	4.50	(0.03)	0.91	(0.05)	0.63	(0.05)	1.65	(0.21)	0.67	(0.04)
Days from prev	44.38	(0.68)	6.70	(1.12)	-2.01	(1.03)	29.28	(4.86)	-1.52	(1.02)
Days to next	43.75	(0.67)	8.56	(1.14)	1.95	(1.04)	26.01	(5.02)	2.33	(1.03)
Soft-Drinks										
Quantity (16 oz)	5.00	(0.26)	5.04	(0.31)	3.01	(0.34)	7.20	(0.59)	3.84	(0.29)
Units	4.18	(0.17)	-2.34	(0.20)	-1.75	(0.26)	-2.96	(0.33)	-2.34	(0.20)
Size (16 oz)	2.82	(0.13)	4.31	(0.15)	2.73	(0.16)	5.97	(0.27)	3.43	(0.14)
Days from prev	24.71	(2.30)	8.85	(2.75)	-2.47	(2.07)	21.19	(6.14)	-0.74	(1.99)
Days to next	21.49	(2.31)	12.89	(2.77)	2.50	(1.99)	24.26	(6.28)	4.04	(1.92)

A sale is defined as any price at least 5 percent below the modal price, of a UPC in a store over the observed period. The column labeled *Within households* controls for a household fixed effect, while the column labeled *Between households* is the regression of household means. Standard errors are provided in parentheses.

Table 3
Demand as a Function of Duration from Previous Promotional Activity

Variable	Laundry Detergents				Soft-drinks			
	all sample		non-sale periods		all sample		non-sale periods	
log(price per 16 oz)	-2.51 (0.56)	-2.46 (0.02)	-2.45 (0.02)	-2.40 (0.02)	-1.83 (0.03)	-1.83 (0.02)	-1.53 (0.05)	-1.50 (0.05)
duration from previous sale	-0.25 (0.10)	0.50 (0.12)	-0.21 (0.10)	0.67 (0.13)	1.23 (0.26)	2.27 (0.25)	0.50 (0.24)	1.72 (0.28)
(duration from previous sale) ²	-0.22 (0.22)	-1.05 (0.25)	-0.30 (0.22)	-1.44 (0.26)	2.32 (0.35)	-3.58 (0.36)	-1.76 (0.34)	-3.18 (0.40)
feature	0.44 (0.03)	0.46 (0.02)	0.56 (0.03)	0.56 (0.07)	0.12 (0.02)	0.13 (0.02)	0.55 (0.04)	0.56 (0.04)
display	1.22 (0.06)	1.19 (0.02)	1.21 (0.02)	1.19 (0.02)	1.54 (0.02)	1.52 (0.02)	1.36 (0.04)	1.32 (0.04)
duration from previous feature	-	-0.71 (0.09)	-	-0.84 (0.11)	-	-0.14 (0.13)	-	-0.50 (0.17)
(duration from previous feature) ²	-	1.16 (0.12)	-	1.41 (0.14)	-	0.36 (0.17)	-	-0.69 (0.23)
duration from previous display	-	-0.47 (0.08)	-	-0.37 (0.08)	-	-1.73 (0.12)	-	-1.39 (0.15)
(duration from previous display) ²	-	0.13 (0.11)	-	0.04 (0.12)	-	1.37 (0.18)	-	0.90 (0.22)
N =	41,995	41,995	35,314	35,314	37,024	37,024	22,135	22,135

The dependent variable in all regressions is the natural logarithm of quantity purchased (measured in 16 ounce units). Each observation is a brand-size combination in a particular store. Duration from previous sale/feature/display is measured as number of weeks, divided by 100, from previous sale/feature/display for that brand in that store for any size. All regressions include brand and store dummy variables. The regressions in the soft-drinks

category are for the sub-sample of cans and include a dummy variables for high demand holiday weeks (July 4, labor day, Thanksgiving and Christmas).