Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

The Cyclicality of Sales, Regular and Effective Prices: Comment

Etienne Gagnon, David Lopez-Salido, and Jason A. Sockin

2015-052

Please cite this paper as:
Gagnon, Etienne, David Lopez-Salido, and Jason A. Sockin (2015). "The Cyclicality of Sales, Regular and Effective Prices: Comment," Finance and Economics Discussion Series 2015-052. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.052.

The Cyclicality of Sales, Regular and Effective Prices: Comment*

Etienne Gagnon
Federal Reserve Board

David López-Salido
Federal Reserve Board

Jason Sockin
Federal Reserve Board

July 9, 2015

Abstract

Coibion, Gorodnichenko, and Hong (2015) argue that the CPI underestimates the deceleration in consumer prices during economic downturns because the index fails to account for the reallocation of consumer spending from high- to low-price stores. We show that these authors' measures of inflation with and without store switching suffer from several methodological deficiencies, including an excessive truncation of price adjustments and the lack of a treatment for missing observations. When we address these deficiencies, the authors' key regression results no longer suggest that greater store switching during downturns is a statistically or economically significant phenomenon.

JEL Codes: D12, E31, E32.

Keywords: Outlet substitution bias, effective prices, inflation measurement.

*The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or any other person associated with the Federal Reserve System. We are grateful to Information Resources Inc. (IRI) for providing the scanner data. All estimates and analyses in this paper are by the authors and not by IRI. Comments and suggestions can be directed to `etienne.gagnon@frb.gov` and `david.lopez-salido@frb.gov`.

1 Introduction

Coibion, Gorodnichenko, and Hong (2015, henceforth "CGH") argue that the U.S. CPI understates the response of consumer prices to economic downturns because the index—which collects information on the prices posted by firms but not on the quantities sold—fails to account for the reallocation of consumer spending from high-price to low-price stores as labor market conditions deteriorate. Using retail data, CGH show that the difference in sensitivity to labor market slack between inflation ignoring quantity movements ("posted price inflation") and inflation taking those movements into account ("effective price inflation") is "*quantitatively large: a 2 percentage point rise in the unemployment rate lowers inflation in effective prices by 0.2–0.3 percentage points relative to inflation in posted prices for a given UPC.*" Applying their range of estimates to the 5-percentage-point jump in the U.S. unemployment rate during the Great Recession suggests that official inflation might have overestimated actual inflation by 0.4–0.7 percentage point. These figures are notable because they indicate, for the first time, that the severity of what is known as "outlet substitution bias" in the price measurement literature fluctuates in an economically significant manner with the business cycle.[1] They also shed light on a potential contributor to the remarkable stability of U.S. nonenergy inflation during and after the Great Recession despite large increases in resource slack.[2]

Our paper shows that CGH's conclusions are unwarranted because their measures of posted and effective price inflation suffer from several methodological deficiencies that, on net, bias the inference toward finding a cyclical role for store switching. One key deficiency is CGH's overly severe truncation of price movements prior to calculating inflation, a procedure presumably employed to control for outliers, but that, ultimately, dampens cyclical movements in posted price inflation more than in effective price inflation. In a nutshell, CGH obtain their measure of posted price inflation by aggregating price movements at the lowest possible level of disaggregation—the item level, where an item corresponds to a universal product code ("UPC") sold in a specific store. Prior to aggregating, they truncate all monthly item price movements that exceed, in absolute terms, 100 percent on an annualized basis, that is, only 8.3 percent on a monthly basis. Most promotional discounts and many

[1] An outlet substitution bias has been known to affect the CPI for several decades but its effects have typically been discussed in the context of structural transformations in the retail industry's competitive landscape unfolding over decades. See, for example, Denison (1962), Oi (1990), and Reinsdorf (1993). The outlet substitution bias is akin to the "sourcing substitution bias" discussed in the trade and producer prices literature. For recent reviews of this related bias, see Houseman *et al.* (2011) and Nakamura *et al.* (2014).

[2] See Robert E. Hall's 2011 presidential address to the American Economic Association, in which he documents this stability and invites macroeconomists to reconsider the long-held view that producers find it desirable to expand output by cutting prices in times of extreme slack.

regular price movements meet that threshold, so truncation greatly reduces the amplitude of movements in their posted price inflation series. CGH derive their measure of effective price inflation in a similar manner, with the crucial difference being that they apply their truncation procedure at a higher level of aggregation for which the threshold is less binding. Because CGH dampen movements in posted price inflation—including cyclical responses—more than movements in effective price inflation, they overestimate the importance of store switching.

When we substitute CGH's aggressive truncation procedure with a less invasive method to control for outsize price adjustments, we find that the difference in sensitivity to slack between posted and effective price inflation loses much of its statistical and economic significance. Notably, CGH's weighted panel regressions, which are arguably the best benchmarks for the size of the cyclical outlet substitution bias in official statistics, have point estimates of the sensitivity to slack with the wrong ordering. The only cases for which we still find a significantly larger sensitivity of effective price inflation to labor market slack are those placing relatively large weights on small markets and sparsely traded items.

This latter pattern is revealing of the effects of a second source of bias in CGH's analysis: the treatment of missing observations. Missing item prices account for over a third of all observations in CGH's sample of weekly scanner data. When computing posted price inflation, CGH systematically fill in most missing monthly item price adjustments with zeros (possibly unintentionally). As was the case with CGH's truncation method, their imputation of missing item price adjustments artificially lowers the cyclical response of posted price inflation relative to that of effective price inflation, with the consequence that the importance of store switching is further overstated during downturns.

To address this second problem, we follow an imputation procedure for missing item prices similar to that employed by the BLS for the CPI. The procedure ensures that item-level inflation sums up to the actual changes in item prices over long horizons even when missing observations are pervasive. In the process, we also fix a number of other unsatisfactory aspects of CGH's methodology, including an improper stitching of the 2001–2007 and 2008–2011 subsamples, the lack of a treatment for clearance sales, and an error in the calculation of effective price inflation for months with five weeks. Together, our data filters result in posted and effective price series that are better behaved than those originally derived by CGH. In particular, our price indexes are much less subject to spurious level jumps or implausibly large divergences between the level of posted and effective prices over the sample period. Importantly for our question of interest, the economic and statistical significance of the difference in sensitivity between posted and effective price inflation to local labor market conditions vanishes from all panel regressions considered by CGH. We thus conclude that CGH's regressions do not offer conclusive evidence, as they claim, "*that effective price inflation is [...] more cyclically sensitive than inflation in posted prices.*"

The paper is organized as follows. Section 2 presents the data and methodology used by CGH. Section 3 discusses the main deficiencies of CGH's methodology and our strategies to remedy them. Section 4 presents our corrected panel regression results. Section 5 concludes with some general remarks on the cyclical importance of the outlet substitution bias and on the challenges researchers face when using high-frequency scanner data.

2 Dataset, definitions, and methodology

2.1 Description of the IRI dataset and terminology

CGH's scanner dataset is made available to researchers by Information Resources Inc. ("IRI"). The dataset contains weekly price and quantity information from 2001 through 2011 on items belonging to 29 personal care, housekeeping, and food product categories, as well as cigarettes and photographic supplies. The data come from a sample of over $2,000$ supermarkets and drugstores operating in 50 U.S. markets. The content of the dataset is detailed in Bronnenberg, Kruger, and Mela (2008).

For clarity of exposition, we define an "item" as a UPC sold in a specific store. An "observation" corresponds to the information collected by IRI on an item in a particular week; this information includes the number of units sold and total revenues, along with characteristics such as the presence of a promotional display. We call the history of an item's price in the sample its "price trajectory." We refer to the combination of a product category and a market as a "stratum." Because our focus is on store switching, we drop all observations pertaining to private labels (that is, to UPCs that are specific to a retail chain), as CGH also do.

2.2 Posted and effective price inflation

All posted and effective price inflation series used in CGH's key regressions are computed at the stratum level and are aggregated to a monthly frequency. The concept of posted price inflation seeks to capture changes in the prices posted by firms ignoring contemporaneous movements in quantities, whereas the concept of effective price inflation seeks to take these quantity movements into account. In principle, a true cost of living index should factor in quantity changes in order to account for substitution over time and across outlets (see, for example, Diewert (1999)). In practice, the use of quantities in the construction of official price indexes is an exception rather than the norm. One reason is that the process of gathering quantity information is often challenging or even impossible, in part because some retailers are reluctant to disclose sales volumes or lack the technology to track those volumes in real time.[3] Even if statistical agencies could observe quantities along with item prices, the computation of an effective price index tracking substitution across outlets would necessitate, for each market, the gathering of *many* observations of items with identical characteristics, which under current data collection methods would be onerous. The IRI dataset is largely immune to the above shortcomings because it tracks the prices and quantities of identical items sold at multiple retailers in each market, with the data collection done electronically rather than in person at the stores. However, its product and outlet coverage is far more limited than that of the CPI.[4]

Following CGH, we define the "posted price," $P_{mscj,t}$, as the price paid for an item in week t , where m, s, c, and j are indices for markets, stores, product categories, and UPCs, respectively. The posted price is calculated as the item's total revenue that week, $TR_{mscj,t}$, divided by the item's total number of units sold, $TQ_{mscj,t}$. We define the "effective price" of UPC c in market m as

$$P^{eff}_{mcj,t} = \frac{\sum_{s \in S(m)} TR_{mscj,t}}{\sum_{s \in S(m)} TQ_{mscj,t}},$$

where $S\left(m\right)$ is the set of participating stores in market m. This effective price can be alternatively expressed as a weighted sum of prices paid at the stores in market m, where store weights are proportional to the number of units sold by each store for that UPC. Effective price inflation at the stratum level is obtained by aggregating the log changes in $P^{eff}_{mcj,t}$ across UPCs sold in the stratum,

$$\pi^{eff}_{mc,t} = \sum_{j \in J(m,c)} w_{mcj,t} \; \log\left(\frac{P^{eff}_{mcj,t}}{P^{eff}_{mcj,t-1}}\right),$$

where $J\left(m,c\right)$ is the set of UPCs sold at participating stores in the stratum. The aggregation uses some relative weights $w_{mcj,t}$ to be discussed shortly. Inflation in posted prices is defined similarly, with the distinction that the aggregation takes place at the item level rather than

[3] To be sure, the BLS gathers information on household consumption patterns every other year, updates the sample of visited stores frequently based on an on-going point-of-sale survey, and works with sampled stores to assess the relative importance of items joining the CPI basket. However, the quantity information thus gathered is unsatisfactory for the purpose of computing a monthly effective price index because it is collected too infrequently and made available with too substantial a lag.

[4] IRI censors the identity of retailers but the dataset is known to exclude independent stores, online retailers, and Wal-Mart.

at the UPC level,

$$\pi_{mc,t}^{pos} = \sum_{(s,j)\in I(m,c)} w_{mscj,t} \log\left(\frac{P_{mscj,t}}{P_{mscj,t-1}}\right),$$

where $I(m,c)$ is the set of items sold in the stratum and $w_{mscj,t}$ is an item's relative weight. Both the UPC-level weights, $w_{mcj,t}$, and the item-level weights, $w_{mscj,t}$, are assumed to change infrequently, so that the reallocation of consumer spending from high-price stores to low-price stores is captured by $\pi_{mc,t}^{eff}$ (through changes in $P_{mcj,t}^{eff}$) but not by $\pi_{mc,t}^{pos}$.

2.3 UPC and item weights

Like CGH, we consider three sets of weights to aggregate changes in $P_{mcj,t}^{eff}$ to stratum-level effective price inflation. "Uniform" UPC weights assign equal weights to all UPCs in a market. "Market-specific" UPC weights are proportional to total spending across stores in each market for all UPCs in a calendar year. "Common" UPC weights are proportional to total spending across all stores in the IRI sample of UPCs during the year. Similarly, we consider three kinds of weights to aggregate changes in $P_{mscj,t}$ to stratum-level posted price inflation. "Uniform" item weights give equal weights to all items in a stratum. "Market-specific" item weights are the product of the market-specific UPC weights defined above and a store's yearly market share in the stratum. Likewise, "common" item weights are the product of the above common UPC weights and a store's share of total annual sales in the product category in the IRI sample.[5]

We make a couple observations regarding these weighting schemes. First, if we had a balanced panel of items, then the item-level market-specific and common weights used for posted price inflation would aggregate to the corresponding UPC-level market-specific and common weights used for effective price inflation. In practice, basket turnover and missing observations may create some differences in UPC-level weights between the two inflation measures, especially if no treatment of missing observations is employed. Second, even if we had a balanced panel of items, CGH's uniform item weights used for posted price inflation (which depend on the number of items in a market) would generally not aggregate up to the uniform UPC weights used for effective price inflation (which depend on the number of UPCs in a market). For this reason, and because uniform weights treat sparsely-traded items and UPCs on the same basis as big sellers, we see the regression results associated with uniform weights as less reliable than those derived for market-specific and common weights for gauging the inflation bias due to store switching.

[5] CGH's paper and technical appendix discuss the aggregation of UPC prices to the stratum level but not the aggregation of item prices to the UPC level or higher. Our description of the item weights is based on our reading of their computer code.

2.4 The cyclicality of posted and effective prices

To assess the cyclicality of price changes with respect to economic conditions, CGH first aggregate the micro data to a monthly frequency and calculate monthly $\pi^{pos}_{mc,t}$ and $\pi^{eff}_{mc,t}$ series (our next section details these steps). They then adopt the following baseline empirical specification for their panel regressions,

$$Y_{mc,t} = \beta \; U_{m,t} + \lambda_t + \theta_{mc} + error_{mc,t},$$

where $Y_{mc,t}$ is the 12-month rate of either posted or effective price inflation, $U_{m,t}$ is the seasonally-adjusted unemployment rate prevailing in month t and market m, and λ_t and θ_{mc} are month and stratum fixed effects, respectively. To the extent that local inflation slows when local labor market conditions deteriorate relative to those in the sample, the point estimates of β should be negative. And to the extent that consumers reallocate their expenditures toward cheaper stores at a faster pace during downturns, then, as CGH find, the estimate of β should be smaller (that is, more negative) for $\pi^{eff}_{mc,t}$ than for $\pi^{pos}_{mc,t}$.[6]

Following CGH, we consider eight variations of the above baseline specification. Arguably the two most important variations use the market-specific and common weights to aggregate price changes into monthly $\pi^{eff}_{mc,t}$ and $\pi^{pos}_{mc,t}$ series at the stratum level, and then use the stratums' respective expenditure shares as panel weights.[7] By accounting for the importance of items, UPCs, and stratums, these regressions come closest to measuring the bias in official statistics, which use importance weights at all these levels. Another two variations similarly use market-specific and common weights to derive the $\pi^{eff}_{mc,t}$ and $\pi^{pos}_{mc,t}$ series, but then treat each panel equally, so that small stratums (for example, razors in Eau Claire, WI) receives the same weights as large ones (for example, carbonated beverages in Chicago, IL) in the estimation of β. The last four variations considered by CGH use uniform weights to compute $\pi^{eff}_{mc,t}$ and $\pi^{pos}_{mc,t}$ at the stratum level, along with equal panel weights; by over-emphasizing sparsely-traded items and UPCs as well as small markets, they provide, in our view, less reliable estimates of the bias afflicting official inflation statistics than the aforementioned four variations.[8] Like CGH, we account for possible sectoral and cross-sectional correlation

[6] A smaller estimate of β for $\pi^{eff}_{mc,t}$ than for $\pi^{pos}_{mc,t}$ is also consistent with retailers having a greater recourse to sales, or households being more responsive to sales, during economic downturns. CGH rule out these alternative explanations by showing that the frequency and depth of sales, as well as the share of goods bought on sale, are insensitive to local labor market conditions in the IRI sample.

[7] We follow CGH in using panel weights proportional to each stratum's median monthly share of total expenditures in the sample.

[8] One of these four regressions uses no month and stratum fixed effects, another only stratum fixed effects, and yet another both month and stratum fixed effects. The remaining variation keeps the stratum fixed effects but replaces the month fixed effects with stratum-specific linear time trends.

in all eight specifications by computing Driscoll and Kraay (1998) standard errors.

3 Shortcomings of CGH's implementation and our solutions

3.1 Truncation of monthly item and UPC price movements

Prior to computing $\pi^{pos}_{mc,t}$, CGH truncate all item-level monthly price adjustments, $\Delta \log (P_{mscj,t})$, that exceed 100 percent on an annualized basis, or, equivalently, 8.3 percent on a monthly basis. Similarly, prior to computing $\pi^{eff}_{mc,t}$, CGH truncate all effective price movements, $\Delta \log \left(P^{eff}_{mcj,t}\right)$, that exceed that threshold. CGH's paper and supporting material do not mention this truncation; our conjecture is that it was introduced to ensure that the regression estimates are not driven by outliers.[9] Unfortunately, the thresholds misleadingly tilt the regression coefficients in the direction of finding greater cyclical responses for $\pi^{eff}_{mc,t}$ than $\pi^{pos}_{mc,t}$.

Table 1 shows that the chosen thresholds are easily met. For posted price inflation (column 1), between a fifth and a quarter of underlying monthly $\Delta \log (P_{mscj,t})$ used by CGH are truncated. Among nonzero adjustments (not shown in the table), the proportion reaches a staggering 70 percent of raw price changes in the sample. The reason why the threshold is so easily met by item price changes is simple: Most promotional discounts and many regular price movements exceed 8.3 percent in absolute terms. The proportion of truncated item price changes is especially large in product categories for which active promotional discounting is a feature of how goods are marketed to consumers. For example, over a third of nonmissing monthly item price adjustments are truncated for carbonated beverages, frozen dinners, and frozen pizzas. For effective price inflation (column 4), the proportion of underlying $\Delta \log \left(P^{eff}_{mcj,t}\right)$ that are truncated, which is around 40 percent, is even larger than that of $\Delta \log (P_{mscj,t})$.

To explore the effects of truncation on CGH's findings while still controlling for possible outliers, we replace CGH's aggressive truncation procedure with the exclusion of items that exhibit one or more outsize adjustments.[10] Our trimming ensures greater uniformity in the item-level price movements that underpin the calculation of $\pi^{pos}_{mc,t}$ and $\pi^{eff}_{mc,t}$. Dropping the affected items' entire price histories leads to the trimming out of only about half a percent

[9] Our replication files, which are available upon request, detail where this and other possible errors occur in CGH's computer code.

[10] In particular, we drop from our sample an item's entire price trajectory whenever $P_{mscj,t}/P_{mscj,t'} < 0.15$ or $P_{mscj,t}/P_{mscj,t'} > (1/0.15)$ for some week t, where $t' \leq t-1$ is the item's previous nonmissing price observation.

of observations accounting for a similarly modest fraction of store revenues.

Table 1 shows that CGH's truncation procedure dramatically lowers the volatility of their $\pi^{pos}_{mc,t}$ and $\pi^{eff}_{mc,t}$ series (expressed on a 12-month rate basis), with the effect being more pronounced for posted price inflation. Absent truncation, the expenditure-weighted standard deviations of posted and effective price inflation are similar, at 4.5 percentage points and 4.6 percentage points, respectively.[11] This tiny difference suggests that the effects of swings in quantities on effective prices, which may be important at the UPC-market level, tend to cancel each other as the data are aggregated to the product-category or higher levels. Truncation more than halves the standard deviation of $\pi^{pos}_{mc,t}$, to 1.9 percentage points, whereas it only cuts that of $\pi^{eff}_{mc,t}$ by about a third, to 3.1 percent.

The relatively large reduction in the volatility of posted price inflation is crucial for understanding why truncation might bias the analysis toward finding a counterfactually important role for store switching. Intuitively, in a univariate context, a linear regression of π_t on U_t produces a coefficient $\beta = cov(\pi_t, U_t)/var(U_t)$. Roughly speaking, CGH's aggressive truncation scales down π_t to $\lambda\pi_t$, with $0 < \lambda < 1$, so that the standard deviation of $\lambda\pi_t$, the covariance term $cov(\lambda\pi_t, U_t)$, and the estimate of β are also scaled down by the same factor λ relative to their true value.[12] Because it is easier for item price changes to meet CGH's truncation threshold than for effective price adjustments at the UPC-market-level, CGH end up scaling down $\pi^{pos}_{mc,t}$ by a factor λ^{pos} that is smaller than the factor λ^{eff} affecting $\pi^{eff}_{mc,t}$. As a result, CGH's estimates of β^{pos} are more severely biased toward zero than those for β^{eff}, thereby leading to an overstatement of the incidence of store switching.

3.2 Treatment of missing observations

CGH do not discuss the presence of missing observations but table 2 shows that they are pervasive in their sample. On average, 39.0 percent of raw observations in the IRI dataset are missing each week. A majority of weekly observations are missing for 9 out of the 31 product categories, with a proportion as high as 68.0 percent for razors.[13] Unfortunately, IRI provides no information that would allow us to ascertain the nature of each missing

[11] These statistics use market-specific item and UPC weights to derive stratum-level inflation series. The standard deviations of the resulting inflation series are then aggregated to the product-category and sample levels using as weights the stratums' spending shares over the full sample period.

[12] A similar line of reasoning applies to CGH's multivariate panel regressions, with the complications that the estimates of β further depend on the covariance between the local unemployment rate and the month and stratum fixed effects, as well as, in the case of weighted regressions, how the incidence of truncation varies with the panel weights.

[13] These statistics are lower bounds because IRI only reports an item's information if a transaction occurs during the week and the store transmits its data; therefore, the statistics count observations that are missing between the first and last recorded transactions in the sample but not those outside of that period.

observation. One aspect we can infer, though, is that only a small proportion of missing observations are due to a failure of stores to report data to IRI. As column 4 indicates, only about 2 percent of the sample's observations come from stores for which all (or essentially all) observations are missing during the week. The vast majority of missing observations are thus attributable to other factors, which include stockouts, out-of-season items, and the absence of weekly transactions.

CGH derive item-level monthly posted price inflation by summing weekly price changes during the month. If all weekly $\Delta \log(P_{mscj,t})$ are missing in a month, then CGH's computer code overwrites the missing monthly entry with a zero.[14] This handling of missing observations is unsatisfactory on several levels. First, a monthly price change may be missing even if the price is observed on occasion during the month, so that there is a loss of information. Second, there is no guarantee that the sum of weekly changes will coincide with the actual change in an item's price over the month. For example, suppose that an item is discounted in the final week of the previous month, stocked out in the first week of the current month, then sold at its original regular price over the remainder of the month. The price changes for the first and second weeks of the month are missing because either the current or previous weekly pice is not observed, so that the price increase at the end of the promotional sale is never registered; CGH would thus record that the price is unchanged during the entire month, as if it had stayed at its promotional level. To the extent that there are systematic patterns in the way weekly observations are missing—say because items are likely to be missing after sales because of stockouts and satiated consumer demand—then CGH's handling could create systematic biases in the measurement of posted and effective price inflation. Moreover, because item prices are more likely to be missing than market-wide UPC effective prices, any bias is likely to affect $\pi^{pos}_{mc,t}$ and $\pi^{eff}_{mc,t}$ to differing degrees. Third, and perhaps most consequential for CGH's analysis, the systematic imputation of missing monthly item price changes with zeros artificially dampens movements in the posted price inflation series. As a consequence, these imputations are likely to tilt the inference toward finding smaller cyclical responses for posted price inflation than for effective price inflation, leading to the finding of a counterfactually large role for store switching.

To control for selection biases due to missing weekly prices and to undo CGH's zero-imputations of missing monthly price changes, we handle missing prices using a procedure similar to that used by the BLS for the CPI. For as long as an item price is missing, our imputation method imputes the last observed price using the inflation rate in the stratum,

[14] It is unclear to us if CGH intended to have this overwriting of missing monthly price changes with zeros. The overwriting of missing observations with zeros is a feature of the particular Stata function called by CGH to aggregate weekly data to a monthly frequency.

but then corrects any imputation errors by effectively linking the next observed weekly item prices with the previous nonmissing weekly prices. As an illustration, suppose that an item is heavily discounted in week $t-1$, stocked out in week t, then transacted at its earlier regular price in period $t+1$. We impute the first missing price change, $\Delta p_{mscj,t}$, by the average amount of inflation in the stratum, $\pi^{pos}_{mc,t}$. In period $t+1$, we compute the change in the item's price from its imputed level in period t, $\Delta p_{mscj,t+1} = \log(P_{mscj,t+1}) - \left(\log(P_{mscj,t-1}) + \pi^{pos}_{mc,t}\right)$. This approach ensures that the sum of $\Delta p_{mscj,t}$ and $\Delta p_{mscj,t+1}$ equals the actual change in the item's price from week $t-1$ to week $t+1$.[15] When item prices are missing for several consecutive periods, we keep imputing the level of the item's price with posted price inflation in the stratum until the item's price is available again. For consistency, we follow the same imputation strategy for missing monthly effective prices, which, as we noted earlier, are less likely to be missing than weekly item prices.

3.3 Incorrect time aggregation of $\pi^{pos}_{mc,t}$ and $\pi^{eff}_{mc,t}$

CGH's aggregation of weekly effective price inflation to a monthly frequency contains a mistake for months that have five weeks. To compute monthly effective price inflation, they first create a *weekly* variable that contains the average paid price *over the entire month* by dividing total revenues that month by total quantities sold that month. They next calculate the *four-week* change in this weekly variable and then take, for each month, the average of those weekly four-week changes to obtain a measure of monthly effective price inflation. A problem arises for months that have five weeks because the four-week change in the monthly average price is always zero for the fifth week of the month. As a result, CGH's measure of monthly effective price inflation understates actual inflation by a fifth in months with five weeks. This time aggregation mistake dampens the measured response of effective prices. As a result, it may actually lead CGH to understate the flexibility of effective prices, and thus to underestimate the bias in posted price inflation due store switching. Table 3 provides an illustration of this error.

This issue aside, the monthly aggregation of effective prices requires at least one pair of weekly observations—one observation in each month—separated by exactly four weeks; otherwise the four-week change will be missing for the month, resulting in a loss of information. Relatedly, the time aggregation of weekly posted price inflation is also unsatisfactory because of CGH's treatment of missing weekly and monthly price observations noted earlier. Our imputation procedure for missing observations allows us to circumvent this problem. We define an item's monthly posted price as the average weekly price—observed or imputed—during

[15] Gagnon and López-Salido (2014) follow the same approach, with the distinction that they only use prices for frequently-traded items to compute inflation.

the month, and then compute monthly $\pi^{pos}_{mc,t}$ by aggregating changes in those item-level monthly averages.

3.4 Spurious index jumps due to clearance sales

Table 4 shows that, on average, 2.3 percent of items in the IRI dataset disappear every month and do not return. A slightly larger proportion of items (about 2.4 percent) join the sample each month, consistent with modest growth in stores' product offerings over time.[16] Selection effects associated with the entry and exit of items are a well known source of bias in the official CPI over long horizons.[17] A number of authors have also explored how such selection effects distort the dynamic response of price indexes to shocks—in our case, the response to variation in local labor market conditions.[18] A fully satisfying treatment of biases due to basket turnover is beyond the scope of this paper because it would require the judgmental linking of millions of entering and exiting items. That said, there is one major source of downward bias in CGH's price indexes that we can readily address and that potentially creates a wedge in the cyclical response of posted and effective price inflation. "Clearance sales" is the phenomenon by which retailers sometimes offer extra discounts on items about to permanently disappear from the shelves. In this case, a failure to link the disappearing item's price with that of its replacement—even when no quality adjustment is required—can lead to a downward bias in the index over time because the index fails to capture the rise in the price as consumers reallocate their spending to new or existing items. The disappearance of an item after a clearance sale permanently lowers the posted price index but, if the UPC to which the item belongs remains in the sample, effective price inflation will remain anchored by the prices of continuing items.[19] Table 4 shows that clearance sales are a material concern in the IRI dataset. On average, the price of an exiting item is over 8 percent lower than the price that prevailed a quarter before the item's exit (that is, 14 to 26 weeks earlier), with a majority of product categories having a price drop of over 10 percent.[20] The probability

[16] In computing these statistics, we exclude all observations in the first and last 13 weeks of a store's presence in the sample to reduce the risk of confounding store and item turnover. Some of the growth in product offering likely reflects a broadening of some product category definitions around sample enlargements.

[17] Notably, a "new good bias" and a "quality change bias" are known to slant upward measured changes in the CPI, leading to an underestimation of the rise in living standards over long periods. See, among many contributors, the edited volume by Boskin et al. (1996), Bresnahan and Gordon (1996), Gordon (2006) and the conference summaries of the Ottawa Group on Price Indices.

[18] Broda and Weinstein (2010) provide evidence of cyclical creation of destruction of barcodes in scanner data. See also Berger *et al.* (2009), Nakamura and Steinsson (2012), and Gagnon, Mandel, and Vigfusson (2014) for some related evidence and theory on the response of trade prices to exchange rate movements.

[19] Note that incorrect imputations of missing market-wide effective price changes can similarly create spurious shifts in the level of the effective price indexes.

[20] We identify promotional sales using the IRI's sales flag, which is extracted from price records using a proprietary algorithm. This sales flag is not directly comparable to the BLS sales flag, which is based on

that an item is on sale in its final week in the sample is higher, at 30.5 percent (column 3), than for the typical item in the sample, at 23.4 percent (column 4), contributing to these relatively lower prices upon exiting.

We limit the incidence of clearance sales on our price indexes through a simple fix: We drop the last quarter (that is, 13 weeks) of every item's price trajectory in the sample. This trimming out horizon is chosen to be long enough that it encompasses the vast majority of clearance sales.

3.5 Incorrect stitching of subsamples

In November 2012, the IRI dataset's sample coverage was extended through the end of 2011 with the release of data for the years 2008 to 2011. The inclusion of these latter years in the analysis is important because these years witnessed the largest swings in unemployment rates. We note that CGH improperly stitched the 2001–2007 and 2008–2011 subsamples. In short, CGH create item identifiers for each subsample, mapping each item to a unique integer. They then aggregate the two subsamples using those item identifiers. Because the two subsamples do not contain the same number of items, the price history of, say, item n in the first subsample is typically stitched with the price history of a different item in the second subsample that had been attributed the n-th identifier. Our implementation corrects for this error.

4 Corrected regression results

The upper panel of table 5 reproduces CGH's original panel regression estimates from their table 1 for the eight specifications they consider. They emphasize the last four columns because these regressions control for the relative importance of items and UPCs in the computation of the $\pi^{pos}_{mc,t}$ and $\pi^{eff}_{mc,t}$ series. The statistical object of interest is the difference between the regression coefficients on the local unemployment rate for 12-month posted price inflation and the corresponding coefficient for 12-month effective price inflation. Across CGH's regressions using either market-specific or common weights (columns 5 through 8), this difference is statistically significant, ranging from 0.084 to 0.130, implying that a 5-percentage point rise in the local unemployment rate is associated with an extra fall in

in-store observations by price collectors. Nonetheless, we take some comfort in that the fraction of items exiting the IRI sample (2.3 percent) multiplied by the fraction for which the IRI price-reduction flag is one upon exit (30.5 percent) is, at 0.7 percent, similar to the fraction of items subject to a clearance sale in the "processed food" and "other goods" categories of the U.S. CPI, at 0.6 percent, as reported by Nakamura and Steinsson (2008) in table 7 of their supplementary materials.

effective price inflation of 0.4–0.7 percentage point. A similar set of regressions are presented in the first four columns for which CGH use uniformly weighted items and UPCs in the computation of 12-month posted and effective price inflation series. Again, the difference in sensitivity to slack between posted and effective price inflation is economically large and statistically significant.

The middle panel of table 5 shows our regression results once we replace CGH's severe truncation of price changes with our more standard treatment of outliers.[21] As expected, removing CGH's severe truncation of observations boosts the measured sensitivity of both posted and effective price inflation to local slack. The only two point estimates that do not increase in absolute size relative to their values in the upper panel pertain to effective price inflation when we use the stratum-level expenditure shares as panel weights (columns 7 and 8); these measures are arguably the least susceptible to CGH's truncation because large markets tend to have less volatile effective prices because of the averaging of item prices across large sets of stores. Importantly for assessing the relevance of store switching, the regression estimates that aggregate across stratums using their relative weights (columns 7 and 8)—which provide the most reliable estimate of the bias affecting official statistics—now have the wrong ordering, with posted price inflation being marginally more cyclically responsive than effective price inflation (although not statistically or economically so). Moreover, when the stratums are given the same weights in the panel regression (columns 5 and 6), the difference in coefficients is a third smaller than reported by CGH and we can no longer reject that the regression coefficients are the same at standard significance levels.

An arguably more direct statistical test for the presence of a cyclical store switching bias hypothesizes that the response to the local unemployment rate of effective price inflation is larger (that is, less negative) than that of posted price inflation. The data do not reject this hypothesis for the weighted panel regressions but do reject it at the 10 percent and 5 percent level of statistical significance for market-specific and common weights, respectively, when the panels are uniformly weighted.

Our results illustrate that, to some degree, CGH's findings of an economically large difference in the cyclical response to slack between the two inflation measures is driven by their severe truncation procedure. That said, given the methodological deficiencies noted in the previous section that, in addition to severe truncation, plague CGH's estimates, the coefficients reported in the middle panel of table 5 do not yet provide satisfactory estimates of the cyclical sensitivity of inflation to labor market slack. Indeed, a number of patterns apparent in the middle panel are worrisome. We note that the estimated cyclical biases

[21] These regressions results also correct for CGH's improper stitching of the 2001–2007 and 2008–2011 subsamples, which only has a limited effect on the regression coefficients at the margin.

are much larger for the panel regressions that apply uniform item and UPC weights in the calculation of $\pi^{pos}_{mc,t}$ and $\pi^{eff}_{mc,t}$ (columns 1 through 4) than for those that use expenditure shares (columns 5 through 8). By placing infrequently-traded items and UPCs on the same footing as big sellers, the first four regressions might be more susceptible to selection effects associated with sparse data, which could explain the contrasting results. Likewise, the fact that the difference in cyclical sensitivity for the panel regressions that use market-specific or common item and UPC weights retains some significance when the stratums are equally weighted (columns 5 and 6) but not when the stratums are weighted by expenditures shares (columns 7 and 8) also worries us that biases other than truncation could be affecting small stratums. We are further motivated to explore other methodological deficiencies because, when we cumulate CGH's posted and effective price inflation series, we find that the resulting price indexes are plagued with large jumps in their levels. Notably, even after controlling for truncation and proper sample stitching, we find that a majority of stratums saw effective prices increase by more than posted prices over the sample period despite the large rise in unemployment, a fact that is inconsistent with consumers taking advantage of lower prices.

Before discussing the regression estimates with all our corrective filters, we note that the application of our filters results in posted and inflation price indexes that are substantially better behaved than those derived using CGH's original methodology.[22] Although we cannot ascertain that all methodological issues have been addressed—for instance, we do not have a formal way of accounting for sample entry and exit besides our simple procedure for clearance sales—we nonetheless see our inflation series as considerably more likely to provide reliable estimates than those produced by CGH.

The bottom panel of table 5 shows our regression results after implementing all of our corrections, that is, including a standard treatment of outsize item price movements, a robust imputation procedure for missing observations, comparable definitions of monthly posted and effective price inflation, a proper stitching of the subsamples, and our simple control for clearance sales. There are two main take-away messages. The first is that the significance of the difference in cyclical responses between posted and effective price inflation has vanished from all specifications considered by CGH. The coefficients pertaining to the weighted panels with inflation series constructed using either market-specific or common weights (columns 7 and 8) changes only moderately with respect to the case with no truncation, but in a direction that makes it even less likely that effective price inflation is more cyclically sensitive than posted price inflation. The corresponding unweighted panel regressions (columns 5 and 6) now also have point estimates with posted price inflation that are marginally more sensitive

[22] Stratum-level posted and effective price indexes are available on our replication page, along with the corresponding indexes consistent with CGH's original methodology.

to slack than effective price inflation, although not in a statistically significant manner. Similarly, the coefficients using inflation measures without regards to item and UPC weights (columns 1 through 4) either have the wrong coefficient ordering for store switching to matter or have no significance. These latter results differ markedly from those shown in the upper and middle panels, and suggest that CGH's original results were driven to some degree by the lack of a proper treatment for sparsely traded items and UPCs.

The second message is that the application of our filters points to inflation—both posted and effective—being more responsive to local labor market conditions than estimated by CGH. Our preferred estimates suggest that a one-percentage point increase in the local unemployment rate lowers 12-month inflation by about 0.15 percentage point, a figure roughly two to three times larger than the range of estimates originally reported by CGH in columns 5 through 8. Of note, these estimates are nearly identical to those presented by Beraja, Hurst, and Ospina (2015) using the Nielsen scanner database, which has a broader coverage of the retail sector than the IRI database. These authors estimate that a one percentage point increase in the local unemployment rate lowers retail prices (measured on a posted basis) by 0.46 percentage point over a three-year period, a figure is almost exactly three times as large as our point estimate for yearly posted price inflation.

5 Conclusion

We have identified several methodological deficiencies in the way CGH calculate posted and effective price inflation that, when addressed, make their finding of effective prices' relatively large cyclical response to slack disappear. It could be tempting to conclude from our analysis that the cyclical bias due to store switching in official price indexes is an economically unimportant phenomenon. However, we are reluctant to draw such a conclusion for a number of reasons. Most importantly, the IRI sample is not representative of the universe of outlets, notably excluding online retailers and the largest U.S. brick-and-mortar retailer, Wal-Mart, which pursued competitive pricing strategies and saw their market shares rise over the sample period. The IRI sample's product coverage is also limited and might thus not be representative the extent of cyclical outlet substitution occurring in other economic sectors. Furthermore, a number of studies have uncovered related evidence that stores adjust prices and households respond to them in a manner that correlates with the business cycle (see Kryvtsov and Vincent (2015) and Nevo and Wong (2014)).

While some methodological deficiencies were specific to CGH's study, others serve as cautionary tales that large scanner datasets can be a mixed blessing compared with the smaller datasets of consumer prices collected by statistical agencies for the production of official price indexes. On the one hand, scanner datasets such as IRI's promise to enable research that would be impossible with the micro datasets behind official indices such as the CPI. For example, the high frequency of data collection allows the study of short-lived economic phenomenons (see, for example, Gagnon and López-Salido (2015)) while the extensive store coverage, coupled with multiple UPC-level observations across stores, permits the study of price dispersion (for example, Kaplan and Menzio (forthcoming)) and retail competition for particular products. Ultimately, the availability of quantity information along with prices could bring us closer to computing satisfactory cost-of-living indexes.

On the other hand, high-frequency scanner data present some important challenges to users relative to micro CPI databases. One challenge is the large size of the datasets, which makes access to a computer cluster a must for computationally intensive operations, such as the filtering out of promotional sales. Another challenge is the high proportion of missing observations, which is partly due to the granularity of the data. Whereas price collectors working for statistical agencies may observe why an item is missing, there is typically no such information stored in scanner dataset. Our implementation of an imputation procedure similar to that used by the BLS reduces some of the potential biases caused by missing observations. Finally, another challenge is sample turnover. The BLS devotes significant resources toward linking departing and entering items in order to reduce the extent of the new goods and quality change biases in the U.S. CPI. In doing so, it collects information on whether the item was discounted prior to leaving the sample, and can make judgmental adjustments. Such an approach is more difficult to implement in large scanner datasets for which millions of item prices would possibly need to be linked on a judgmental basis. That said, it seems possible to avoid some significant downward bias in price indexes by trimming out the last several observations of each item's price trajectory, as we have done with our sample.

References

[1] Beraja, Martin, Erik Hurst, and Juan Ospina (2015). "The Aggregate Implications of Regional Business Cycles," mimeo, University of Chicago.

[2] Berger, David, Jon Faust, John H. Rogers, and Kai Steverson (2009). "Border Prices and Retail Prices," International Finance Discussion Papers 972, Board of Governors of the Federal Reserve System.

[3] Bresnahan, Timothy F., and Robert J. Gordon (1997). "The Economics of New Goods," Studies in Income and Wealth, Volume 58, National Bureau of Economic Research.

[4] Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale Jorgenson (1996). "Toward a More Accurate Measure of the Cost of Living," Final Report to the Senate Finance Committee.

[5] Broda, Christian, and David E. Weinstein (2010). "Product Creation and Destruction: Evidence and Price Implications," *American Economic Review*, vol. 100(3), pages 691–723.

[6] Bronnenberg, Bart J., Michael W. Kruger, and Carl F. Mela (2008). "Database Paper: The IRI Marketing Data Set," *Marketing Science*, vol. 27(4), pages 745–748.

[7] Bureau of Labor Statistics, Handbook of Methods, Washington DC, 2009.

[8] Coibion, Olivier, Yuriy Gorodnichenko, and Gee Hee Hong (2015). "The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications," *American Economic Review*, vol. 105(3), pages 993–1029.

[9] Denison, Edward (1962). "The Sources of Economic Growth in the United States and the Alternatives before Us," New York: Committee for Economic Development.

[10] Diewert, W. Erwin (1999). "Axiomatic and Economic Approaches to International Comparisons," NBER Chapters, in: *International and Interarea Comparisons of Income, Output, and Prices*, pages 13–107, National Bureau of Economic Research, Inc.

[11] Driscoll, John C., and Aart C. Kraay (1998). "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *Review of Economics and Statistics*, vol. 80(4), pages 549–60.

[12] Fabiani, Silvia, Martine Druant, Ignacio Hernando, Claudia Kwapil, Bettina Landau, Claire Loupias, Fernando Martins, Thomas Mathä, Roberto Sabbatini, Harald Stahl, and Ad Stokman (2006). "What Firms' Surveys Tell Us about Price-Setting Behavior in the Euro Area." *International Journal of Central Banking*, vol. 2(3), pages 3–48.

[13] Gagnon, Etienne and David López-Salido (2014). "Small Price Responses to Large Demand Shocks," Finance and Economics Discussion Series 2014-18, Board of Governors of the Federal Reserve System.

[14] Gagnon, Etienne, Benjamin R. Mandel, and Robert J. Vigfusson (2014). "Missing Import Price Changes and Low Exchange Rate Pass-Through." *American Economic Journal: Macroeconomics*, vol. 6(2), pages156–206.

[15] Gordon, Robert J. (2006). "The Boskin Commission Report: A Retrospective One Decade Later," *International Productivity Monitor*, Centre for the Study of Living Standards, vol. 12, pages 7–22.

[16] Greenlees, John S., and Robert McClelland (2011), "New Evidence on Outlet Substitution Effects in Consumer Price Index Data," *Review of Economics and Statistics*, vol. 93(2), pages 632–646.

[17] Hall, Robert E. (2011). "The Long Slump," *American Economic Review*, vol. 101(2), pages 431–69.

[18] Houseman, Susan, Christopher Kurz, Paul Lengermann, and Benjamin Mandel (2011). "Offshoring Bias in U.S. Manufacturing," *Journal of Economic Perspectives*, vol. 25(2), pages 111–32.

[19] Kaplan, Greg and Guido Menzio (forthcoming). "The Morphology of Price Dispersion," with Greg Kaplan, *International Economic Review.*

[20] Klenow, Peter J. and Oleksiy Kryvtsov (2008). "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" *Quarterly Journal of Economics*, vol. 123 (3), pages 863–904.

[21] Klenow, Peter J. and Benjamin A. Malin (2011). "Microeconomic Evidence on Price-Setting." *Handbook of Monetary Economics*, Volume 3, Chapter 6, in B. M. Friedman and M. Woodford (editors), pages 231–284. North Holland: Elsevier.

[22] Kryvtsov, Oleksiy and Nicolas Vincent (2014). "On the Importance of Sales for Aggregate Price Flexibility," Working Papers 14-45, Bank of Canada.

[23] Nakamura, Alice O., W. Erwin Diewert, John S. Greenlees, Leonard I. Nakamura, and Marshall B. Reinsdorf (2014). "Sourcing Substitution and Related Price Index Biases," mimeo, University of British Columbia.

[24] Nakamura, Emi and Jón Steinsson (2008). "Five Facts about Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics*, vol. 123(4), pages 1415–1464.

[25] Nakamura, Emi and Jón Steinsson (2012). "Lost in Transit: Product Replacement Bias and Pricing to Market," *American Economic Review*, vol. 102(7), pages 3277–3316.

[26] Nevo, Aviv and Arlene Wong (2015). "The Elasticity of Substitution Between Time and Market Goods: Evidence from the Great Recession," NBER Working Papers 21318, National Bureau of Economic Research, Inc.

[27] Oi, Walter Y. (1992). "Productivity in the Distributive Trades: The Shopper and the Economies of Massed Reserves," in *Output Measurement in the Service Sectors*, Zvi Griliches, editor, NBER Books, National Bureau of Economic Research, Inc.

[28] Reinsdorf, Marshall (1993). "The Effect of Outlet Price Differentials on the U.S. Consumer Price Index," in *Price Measurements and Their Uses*, Murry Foss, Marylin Manser, and Allan Young, editors, NBER Books, National Bureau of Economic Research, Inc.

Table 1: Incidence of CGH's truncation and its dampening effect on inflation

Product category	Posted prices: share of $\mathbf{\Delta p}_{mscj,t}$ truncated (percent) (1)	Posted prices: $\boldsymbol{\sigma}\left(\pi^{pos}_{mc,t}\right)$ untrunc. (percent) (2)	Posted prices: $\boldsymbol{\sigma}\left(\pi^{pos}_{mc,t}\right)$ trunc. (percent) (3)	Effective prices: share of $\mathbf{\Delta p}^{eff}_{mcj,t}$ truncated (percent) (4)	Effective prices: $\boldsymbol{\sigma}\left(\pi^{eff}_{mc,t}\right)$ untrunc. (percent) (5)	Effective prices: $\boldsymbol{\sigma}\left(\pi^{eff}_{mc,t}\right)$ trunc. (percent) (6)
Beer	14.4	1.9	1.2	21.2	1.7	1.5
Blades	10.6	1.7	0.9	44.1	1.9	1.5
Carb. beverages	33.4	5.2	2.1	34.6	5.1	3.6
Cigarettes	4.1	6.2	3.5	36.1	7.4	5.7
Coffee	18.3	6.4	2.9	43.0	7.5	4.4
Cold cereal	26.7	3.9	1.4	46.3	3.8	2.0
Condiments	11.6	4.4	1.9	38.8	5.0	3.3
Deodorant	13.4	2.0	0.6	54.1	2.5	1.7
Diapers	14.7	3.3	1.8	39.7	4.3	3.9
Facial tissue	27.4	6.3	2.5	44.7	5.8	3.7
Frozen dinners	34.2	4.5	1.5	43.3	4.3	2.6
Frozen pizza	36.4	5.1	1.9	41.1	4.8	3.2
Hot dogs	37.5	6.0	1.9	39.9	5.8	3.3
Household cleaners	15.7	2.9	1.3	42.5	3.1	2.1
Laundry detergent	25.4	3.8	1.5	48.2	4.2	2.6
Margarine/butter	26.3	7.5	3.6	30.8	7.3	5.0
Mayonnaise	18.9	7.9	3.5	32.5	8.2	5.3
Milk	16.7	5.7	3.8	18.2	5.9	5.0
Paper towels	23.5	4.7	2.1	43.5	5.0	3.2
Peanut butter	19.1	5.9	2.9	29.7	6.3	4.3
Photography	10.7	7.0	1.7	66.0	6.3	3.8
Razors	10.9	3.3	1.3	65.8	4.1	2.8
Salty snacks	26.1	4.0	1.6	34.5	3.9	2.7
Shampoo	14.6	2.1	0.7	77.7	2.6	1.8
Soup	20.5	4.6	1.8	45.5	5.1	3.1
Spaghetti sauce	25.4	4.7	2.0	49.9	4.8	3.1
Sugar/substitutes	10.3	3.0	1.6	29.4	3.0	2.5
Toilet tissue	25.8	4.5	2.2	43.6	5.1	3.3
Tooth brushes	12.2	2.7	0.8	61.6	3.3	2.0
Tooth paste	19.1	2.7	1.0	51.0	3.1	1.9
Yogurt	30.0	4.8	2.1	31.4	4.1	2.9
Mean						
Unweighted	20.5	4.5	1.9	42.9	4.7	3.2
Expenditures-weighted	23.6	4.5	2.0	37.5	4.6	3.1
Observations-weighted	22.0	*n.a.*	*n.a.*	42.7	*n.a.*	*n.a.*

Source: Authors' calculations using IRI data.
Notes: The statistics exclude private labels. We pool raw observations across markets and months to derive the shares of truncated posted and effective price changes at the product-category level. Stratum-level inflation is measured on a 12-month basis. The averaging of statistics across product categories uses either uniform weights ("Unweighted"), sample expenditures weights ("Expenditures-weighted"), or raw observations weights ("Observations-weighted").

Table 2: Statistics on missing weekly item observations in the IRI sample

Product category	**Mean weekly sales** (thousands of $) (1)	**Mean weekly observations** (thousands) (2)	**Missing observations** (percent of sample obs.) Total (3)	Reporting issues (4)
Beer	10,833	302	31.4	2.0
Blades	952	120	57.4	2.0
Carb. beverages	14,209	433	27.0	2.1
Cigarettes	6,017	410	61.7	1.9
Coffee	3,270	238	43.1	2.1
Cold cereal	7,588	312	24.4	2.1
Condiments	792	88	42.3	2.5
Deodorant	1,029	433	61.3	1.9
Diapers	1,518	118	55.3	2.2
Facial tissue	946	44	25.7	1.8
Frozen dinners	6,641	505	25.0	2.0
Frozen pizza	3,364	166	28.3	2.2
Hot dogs	945	132	40.3	2.2
Household cleaners	1,978	62	21.3	2.3
Laundry detergent	3,640	171	35.1	1.9
Margarine/butter	1,532	73	13.0	2.2
Mayonnaise	1,276	54	24.4	2.7
Milk	4,631	118	23.5	2.1
Paper towels	2,106	43	22.5	1.9
Peanut butter	931	51	24.1	2.6
Photography	229	33	64.9	5.1
Razors	144	29	68.0	6.0
Salty snacks	9,039	502	27.9	1.9
Shampoo	1,127	432	64.4	1.9
Soup	4,056	432	33.3	2.2
Spaghetti sauce	1,919	175	32.4	2.5
Sugar/substitutes	410	37	36.0	3.2
Toilet tissue	3,483	60	20.2	1.8
Tooth brushes	602	214	63.7	2.1
Tooth paste	1,284	266	49.7	1.9
Yogurt	4,350	248	15.5	2.1
Total	100,842	6,299	*n.a.*	*n.a.*
Mean				
Unweighted	*n.a.*	*n.a.*	37.5	2.4
Expenditures-weighted	*n.a.*	*n.a.*	31.5	2.1
Observations-weighted	*n.a.*	*n.a.*	39.0	2.1

Source: Authors' calculations using IRI data.
Notes: The statistics exclude private labels. We pool raw observations across markets and months to obtain product-category figures. We categorize a missing observation under "Reporting issues" whenever over 95 percent of its store's weekly observations for the product category are missing. The averaging of statistics across product categories uses either uniform weights ("Unweighted"), sample expenditures weights ("Expenditures-weighted"), or raw observations weights ("Observations-weighted").

Table 3: Illustration of CGH's time aggregation mistake for effective prices

Week	**Month**	**Monthly effective price** (1)	**Four-week change in monthly effective price** (2)	**CGH's monthly effective price inflation** (3)
1	1	1	.	.
2	1	1	.	.
3	1	1	.	.
4	1	1	.	.
5	2	1.25	0.25	0.2
6	2	1.25	0.25	0.2
7	2	1.25	0.25	0.2
8	2	1.25	0.25	0.2
9	2	1.25	0	0.2

Notes: The table illustrates CGH's incorrect time aggregation of weekly effective prices for months that have five weeks. We suppose that a UPC's first and second months in the sample contain four weeks and five weeks, respectively. We assume that the UPC's effective (log) price across stores in the market is 1 in the first month and 1.25 in second month (column 1), consistent with a 0.25 log increase between the two periods. To measure monthly effective price inflation, CGH first compute the four-week change in the weekly series of these monthly effective prices (column 2). For weeks 5 through 8, the four-week change coincides with the actual monthly effective price change. For week 9, the four-week change is zero because weeks 5 and 9 belong to the same month. CGH then calculate the average four-week change during the month (column 3), so they report a change of only 0.2 for month 2.

Table 4: Item turnover and clearance sales in the IRI sample

Product category	**Monthly exit rate** (percent) (1)	**Monthly entry rate** (percent) (2)	**On sale** (percent) (3)	**On sale final week** (percent) (4)	**Price drop final week** (percent) (5)
Beer	1.4	1.7	20.3	21.2	2.6
Blades	2.4	2.4	16.2	25.4	12.3
Carb. beverages	2.0	2.1	32.2	30.7	2.2
Cigarettes	2.3	2.0	6.8	4.0	-1.6
Coffee	1.8	2.2	22.1	30.7	13.3
Cold cereal	2.1	2.2	23.0	42.1	17.0
Condiments	1.3	1.2	14.8	26.7	13.8
Deodorant	2.7	2.6	22.4	34.4	18.5
Diapers	3.9	4.2	24.6	35.0	9.6
Facial tissue	2.9	2.9	23.6	33.8	11.7
Frozen dinner	2.3	2.7	32.2	44.8	14.4
Frozen pizza	1.9	2.1	33.5	38.0	7.6
Hot dogs	1.5	1.5	27.4	28.5	3.9
Household cleaners	2.0	2.9	17.8	33.6	19.3
Laundry detergent	3.1	3.2	25.6	42.6	14.2
Margarine/butter	1.4	1.3	19.2	29.5	7.6
Mayonnaise	1.6	1.7	15.6	30.4	11.7
Milk	1.8	2.1	13.9	18.8	2.9
Paper towels	4.2	4.2	19.8	32.7	9.0
Peanut butter	1.2	1.4	16.8	31.5	12.0
Photography	3.0	2.0	21.0	18.7	13.4
Razors	3.5	3.7	24.6	28.4	15.5
Salty snacks	3.7	3.8	25.1	27.3	3.3
Shampoo	3.2	3.2	25.8	33.5	16.6
Soup	1.3	1.8	20.3	37.1	20.1
Spaghetti sauce	1.3	1.3	24.8	33.9	15.4
Sugar/substitutes	1.4	1.7	12.1	28.9	19.8
Toilet tissue	3.5	3.6	21.5	35.0	8.4
Tooth brushes	2.8	3.0	21.0	29.5	20.3
Tooth paste	2.5	2.5	23.6	36.0	18.5
Yogurt	2.3	2.6	24.5	35.8	7.6
Mean					
Unweighted	2.3	2.4	21.7	30.9	11.6
Expenditures-weighted	2.3	2.4	23.4	30.5	8.2

Source: Authors' calculations using IRI data.
Notes: The statistics exclude private labels. We pool raw observations across markets and months to obtain product-category figures. "On sale" is the fraction of nonmissing weekly observations for which the IRI sales flag is activated. "On sale in final week" is the corresponding fraction using only the last observation of price trajectories. "Price drop final week" is the percent (in log changes) by which an item's last observed price is below its mean price over the previous 14 to 26 weeks. With the exception of "On sale," all statistics exclude observations within 13 weeks of a store's entry in or exit from the sample. The averaging of statistics across product categories uses either uniform weights ("Unweighted") or sample expenditures weights ("Expenditures-weighted").

Table 5: Response of 12-month posted and effective price inflation to local unemployment rate

	Uniformly-weighted items and UPCs				Expenditure-weighted items and UPCs			
					Market-specific	Common	Market-specific	Common
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CGH's original estimates								
Inflation								
Posted prices	-0.084^{**}	-0.087	-0.061^{***}	-0.164^{**}	-0.077^{**}	-0.075^{**}	-0.052^{**}	-0.059^{**}
	(0.041)	(0.053)	(0.017)	(0.067)	(0.021)	(0.023)	(0.026)	(0.029)
Effective prices	-0.120^{*}	-0.126	-0.219^{***}	-0.288^{***}	-0.201^{***}	-0.205^{***}	-0.136^{***}	-0.146^{***}
	(0.067)	(0.087)	(0.024)	(0.105)	(0.031)	(0.033)	(0.027)	(0.028)
Test p-value								
$\hat{\beta}^{pos} = \hat{\beta}^{eff}$	0.246	0.332	< 0.001	0.011	< 0.001	< 0.001	0.019	0.026
GLSS' estimates without truncation								
Inflation								
Posted prices	-0.237^{***}	-0.241^{**}	-0.153^{***}	-0.253^{*}	-0.189^{***}	-0.182^{***}	-0.134^{***}	-0.137^{**}
	(0.090)	(0.115)	(0.054)	(0.139)	(0.050)	(0.058)	(0.052)	(0.063)
Effective prices	-0.403^{***}	-0.435^{***}	-0.444^{***}	-0.452^{***}	-0.268^{***}	-0.276^{***}	-0.088^{**}	-0.123^{***}
	(0.107)	(0.137)	(0.033)	(0.154)	(0.030)	(0.038)	(0.042)	(0.048)
Test p-value								
$\hat{\beta}^{pos} = \hat{\beta}^{eff}$	< 0.001	< 0.001	< 0.001	0.002	0.139	0.090	0.386	0.819
$\hat{\beta}^{pos} < \hat{\beta}^{eff}$	< 0.001	< 0.001	< 0.001	0.001	0.070	0.045	0.807	0.591
GLSS' estimates with all data filters								
Inflation								
Posted prices	-0.067	-0.058	-0.223^{***}	-0.428^{**}	-0.256^{***}	-0.253^{***}	-0.142^{***}	-0.157^{***}
	(0.120)	(0.157)	(0.036)	(0.189)	(0.044)	(0.051)	(0.040)	(0.047)
Effective prices	-0.083	-0.072	-0.236^{***}	-0.416^{***}	-0.222^{***}	-0.210^{***}	-0.090^{**}	-0.098^{**}
	(0.100)	(0.132)	(0.032)	(0.161)	(0.034)	(0.038)	(0.039)	(0.043)
Test p-value								
$\hat{\beta}^{pos} = \hat{\beta}^{eff}$	0.486	0.633	0.507	0.718	0.284	0.193	0.084	0.045
$\hat{\beta}^{pos} < \hat{\beta}^{eff}$	0.243	0.317	0.254	0.641	0.858	0.904	0.958	0.978
Specification								
Stratum fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed effects	No	No	Yes	Linear trend	Yes	Yes	Yes	Yes
Weighted regressions	No	No	No	No	No	No	Yes	Yes

Source: "CGH's original estimates" are reproduced from table 1 in Coibion, Gorodnichenko, and Hong (2015). All other numbers are the authors' calculations using IRI data.
Notes: The derivation of "GLSS' estimates without truncation" includes a proper stitiching of the subsamples and the exclusion of price trajectories with outsize adjustments. The derivation of "GLSS' estimates with all data filters" further includes comparable time aggregations of weekly price adjustments and treatments for missing prices and clearance sales. Driscoll and Kraay (1998) standard errors are reported in parentheses below point estimates. Statistical significance at the 10, 5, and 1 percent levels is indicated with one, two, and three stars, respectively.