Response to "The Cyclicality of Sales, Regular and Effective Prices: Comment" by Gagnon, López-Salido and Sockin.

Olivier Coibion, UT Austin and NBER

Yuriy Gorodnichenko, U.C. Berkeley and NBER

Gee Hee Hong*, International Monetary Fund

First draft: Oct 1st, 2015

This Draft: Aug 6th, 2018

Abstract: We address how using different censoring thresholds and imputation procedures affects the baseline results of Coibion, Gorodnichenko and Hong (2015). Higher censoring thresholds introduce measurement error and outliers that generate wide variability in results across weighting schemes, but methods that explicitly control for outliers confirm the results of Coibion et al. (2015) for all censoring thresholds. We also illustrate how the BLS's approach to imputing missing prices can introduce a cyclical bias into measures of posted price inflation when store-switching is present in the data.

Keywords: Sales, Price Changes, Store-Switching, Inflation Measurement.

JEL codes: E3, E4, E5.

To explain monetary non-neutrality, macroeconomists have long emphasized price stickiness as a likely mechanism, with many papers trying to assess how inflexible prices are in the data. In Coibion, Gorodnichenko and Hong (2015, CGH henceforth), we use a scanner-level dataset that includes information on both prices

* Coibion: Department of Economics, University of Texas-Austin, 2225 Speedway, BRB 3.156 C100, Austin, TX 87812 (e-mail: ocoibion@austin.utexas.edu); Gorodnichenko: Department of Economics, University of California-Berkeley, 530 Evans Hall #3880, Berkeley, CA 94720 (e-mail: ygorodni@econ.berkeley.edu); Hong: International Monetary Fund (e-mail: ghong@imf.org).

We are grateful to Eric Sims for helpful comments and suggestions. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

and quantities. This allows us to characterize not just how average retail prices ("posted prices") evolve over the course of the business cycle but also how the average prices *paid* by consumers ("effective prices") change. These effective prices can differ from posted prices even for a given Universal Product Code (UPC) if consumers change the retailers from which they do their purchases in response to changing economic conditions.

CGH provides several pieces of empirical evidence consistent with store-switching behavior by consumers. First, while average posted prices in a metropolitan area decline little in response to hikes in local unemployment, the effective prices paid by households fall more sharply than posted prices. Second, for a typical UPC within an area, a rise in unemployment leads to a larger share of goods being purchased from the bottom end of the cross-retailer price distribution. Third, expenditures at high-price retailers experience relatively larger declines during downturns than expenditures in low-price retailers. Fourth, using a rich panel data set reporting individual household consumption at these retailers, we find that households reallocate their expenditures toward low-price retailers when local economic conditions deteriorate. Thus, due to store-switching on the part of consumers, prices paid by households are more flexible than prices posted by retailers.

In their comment, Gagnon, Lopez-Salido and Sockin (2017, GLSS henceforth) challenge the first of these four pieces of evidence. They make three main arguments. The first two are on the choice of censoring thresholds to deal with outliers. GLSS argue that if one significantly raises the censoring thresholds in the pricing data or replaces the censoring with droppings all price sequences that contain an outlier, two out of the six main empirical specifications of CGH regarding the difference in how effective and posted prices respond to local economic conditions become insignificantly different from zero. Out of the remaining four specifications, two display even stronger results than in CGH while the remaining two display smaller differences but generally remain statistically significant. In response to this, we discuss at length the tradeoffs associated with censoring thresholds and provide a metric for interpreting how higher thresholds affect our data. In short, higher thresholds serve mainly to increase measurement error and the role of outliers. Consistent with this, methods that systematically address outliers confirm our baseline findings for all censoring thresholds.

The third comment of GLSS is that one should follow the Bureau of Labor Statistics' (BLS) price imputation procedure to deal with missing price observations. When an item's price cannot be observed in a given store, the BLS infers its change in price using price changes of the same good in other stores. However, this procedure is problematic in the presence of store-switching. We demonstrate that because low-price and high-price retailers do *not* follow the same pricing strategies over the business cycle, this imputation leads to systematic errors in the predicted price paths at higher-price retailers and, as a result, *attenuates* the difference between the cyclicality of effective and "measured" posted-price inflation. Thus, the fact that GLSS find a smaller difference in sensitivity of posted and effective prices to unemployment with this imputation procedure is *exactly what one would expect when store-switching is important* and therefore should not be interpreted as evidence against this type of consumer behavior.

I. Censoring Thresholds, Weights, and Outliers

A. Censoring Thresholds

The first issue raised by GLSS is the censoring (winsorization) threshold applied to price changes. GLSS point out that as many as 70 percent of non-zero price changes can be affected by our threshold. However, this claim is misleading for two reasons. First, the vast majority of these affected price changes are sales. By definition, the effects of sales on prices are immediately reversed and therefore have no effect on longer run price levels and annual inflation rates. The most relevant metric for the censoring threshold is how the choice of a specific threshold affects *regular* price changes. Second, they count only non-zero price changes but this ignores the fact that prices are unchanged much of the time. In Appendix Tables 1 and 2, we present the share of censored regular price changes out of all non-zero price changes as well as out of all non-missing price observations for censoring thresholds going from 1 (our baseline) to 12 (the GLSS value). Our censoring threshold binds for approximately 30 percent of *non-zero* regular price changes, which corresponds to approximately 1.5 percent of all non-missing observations.

The censoring threshold used by GLSS, in contrast, binds for less than one-tenth of one percent of all non-zero price changes, which corresponds to less than one hundredth of a percent of non-missing observations, allowing for a larger role of outliers.[2]

Given that censoring thresholds are largely chosen at the discretion of researchers, it is crucial to understand how changing thresholds affect the underlying data and subsequent results. GLSS propose one interpretation: a low censoring point reduces the volatility of posted price inflation which can attenuate its sensitivity to changes in unemployment. Specifically, letting π^{p^*} denote the unobserved true rate of posted-price inflation for a given market and category of goods, their interpretation is that one observes π^p where $\pi^p = \lambda\pi^{p^*}$ and $\lambda(T)$ is an attenuation factor increasing in the censoring threshold (T), i.e. $\lambda' > 0$. Thus, GLSS argue, if raising the censoring point increases the variance of π^p, this is good because it brings π^p closer to π^{p^*}, so one should use high censoring thresholds. They also assume that effective price inflation is invariant to the censoring threshold. But the standard argument for the need to use censoring methods is simply that there is noise (e.g., measurement error) in the underlying data, i.e. $\pi^p = \pi^{p^*} + \theta$ where θ is i.i.d. noise whose variance $\sigma_\theta^2(T)$

[2] For comparison, in the corporate finance and accounting literatures where widely cited papers have adopted the methodology of winsorization of financial variables, thresholds have varied widely. Malmendier and Tate (2005) use 1st and 99th percentile thresholds, Sufi (2009) uses 5th and 95th percentile as the cutoff values, while Sharpe and Suarez (2014) use the 10th and 90th percentiles for censoring. Similarly, in the labor literature, Angrist and Krueger (1999) recommend systematically winsorizing (censoring) earnings data and they propose thresholds ranging from 1 percent to 10 percent.

increases in T. There is a simple test to distinguish between the two theories in our context.

Suppose that true posted (π^{p^*}) and effective (π^f) price inflation are each related to unemployment u as follows:

$$\pi^f = \beta_f u + \varepsilon^f,$$

$$\pi^{p^*} = \beta_{p^*} u + \varepsilon^{p^*}.$$

The ε are shocks to each process and are possibly correlated. Under the measurement error explanation of CGH in which $\pi^p = \pi^{p^*} + \theta$ and $\beta_f < \beta_{p^*} \leq 0$ (so effective prices are more sensitive to unemployment than posted prices), running the following regression:

(1) $\quad \pi^f - \pi^p = \beta u + \varepsilon$

will yield estimates of β which are independent of the censoring threshold and which consistently recover the true difference in sensitivity to unemployment: $\hat{\beta} = \beta_f - \beta_{p^*}$.

Intuitively, we are introducing measurement error into the left-hand side of equation (1) which does not affect the properties of $\hat{\beta}$. However, the variance of the residuals of the regression will be increasing in the censoring threshold since $var(\varepsilon) = var\left(\varepsilon^f - \varepsilon^{p^*}\right) + var(\theta)$.[3]

The attenuation interpretation of GLSS makes a different prediction with respect to the variance of these residuals. When $\beta_f = \beta_{p^*}$ as GLSS argue, one can

[3] If prices are increasing over time or if there is positive skewness in price changes, censoring may disproportionately affect price increases, introducing a bias between π^p and π^{p^*}. However, this does not change the subsequent results.

show that $\hat{\beta}$ now depends on the censoring threshold: $\hat{\beta} = (1-\lambda)\beta_{p^*}$ so that higher thresholds should be associated with smaller differences in estimated sensitivities to inflation. Taking the estimates of β_f and β_{p^*} using the CGH threshold to get the "attenuated" $\hat{\beta}$ yields $\hat{\beta} = -0.084$ and taking the corresponding estimate of β_{p^*} using GLSS threshold as the true β_{p^*} (so $\beta_{p^*} = -0.127$) implies an attenuation value of $\lambda = 0.34$ for a censoring threshold of 1.[4]

Under the attenuation interpretation asserted by GLSS, one can show that the variance of the residual of the regression will be $var(\varepsilon) = \sigma_f^2 + \lambda^2\sigma_{p^*}^2 - 2\lambda\sigma_{fp^*}$ where $\sigma_f^2 = var(\varepsilon^f)$, $\sigma_{p^*}^2 = var(\varepsilon^{p^*})$, $\sigma_{fp^*} = cov(\varepsilon^f, \varepsilon^{p^*})$. When $\lambda > \sigma_{fp^*}/\sigma_{p^*}^2$, the $var(\varepsilon)$ will be increasing in λ and therefore in the censoring threshold *T*. Empirically, $\sigma_{fp^*}/\sigma_{p^*}^2$ is the slope in the regression of the residual $\hat{\varepsilon}^f$ on the residual $\hat{\varepsilon}^p$. Across different weighting specifications and weighing schemes, we find that $\sigma_{fp^*}/\sigma_{p^*}^2$ is at least 0.6 and in many cases well above 0.6. Since $\lambda = 0.34$ for the CGH truncation threshold, the attenuation interpretation makes the prediction that *the variance of the residuals should be decreasing* when we raise the censoring threshold above the CGH value (at some point, this will reverse for high enough λ but if λ is high then changing the threshold point should make little difference for the estimates), whereas the measurement error interpretation implies that it should be increasing. Hence, we can differentiate between the two potential explanations by assessing whether the variance of the regression residuals are increasing or decreasing with censoring thresholds.

We implement this simple test and present the results in Table 1. Across weighting specifications and censoring thresholds ranging from 1 (CGH baseline) through 12 (GLSS baseline), we find that the standard deviation of the residual (root

---

[4] These estimates are from Table 3 in GLSS, column (7), using market-specific expenditure-weights. Alternative weightings give similar results: implied attenuation values from columns (5)-(8) in GLSS range from 0.29 to 0.37.

mean squared error, RMSE) increases considerably in the censoring point. For example, in the equally weighted specification (row 1), the RMSE is 0.0179 for the CGH threshold and 0.0342 for censoring point of 5. For all specifications, the RMSE rises sharply (all differences from T=1 are statistically significant at the 1 percent level) when the censoring threshold goes above the CGH threshold, a finding at odds with the attenuation interpretation but entirely consistent with the measurement error interpretation.[5]

B. Weights and Outliers

One of the striking features of Table 3 in GLSS is that they find qualitatively different results than CGH *only* for some of the specifications, namely weighted regressions when categories are aggregated using expenditure weights. Some other specifications, especially unweighted regressions, yield even stronger results than originally found by CGH once GLSS apply their alternative censoring threshold. GLSS provide no explanation for these differences and argue that we should care *only* about the weighted regressions. We disagree and believe that a sensible explanation must account for all of these results. In this section, we argue that GLSS's logic for focusing only on weighted regressions is incorrect and that the increased measurement error introduced by their approach can account for the patterns in their results.

GLSS argue that only weighted regressions with expenditure-weights used to create category level inflation rates are informative. Since the CPI is itself weighted, their logic is that one must use these same weights in regressions to make explicit

[5] Relatedly, GLSS argue that dropping outliers (instead of censoring them) reverses the findings in CGH. But they again first apply a very high censoring threshold, so that almost no observations are identified as outliers to be dropped. Instead, their approach introduces measurement error and many outliers into the data relative to the specification used in CGH and therefore yields nearly identical estimates as those with censoring at their high thresholds.

statements about differences between aggregate CPI and aggregate effective inflation. We disagree. We are interested in measuring the sensitivity of posted and effective price inflation, using local variation to provide identification. Once this sensitivity has been estimated, aggregate measures of posted and effective price inflation can readily be constructed using expenditure weights. These weights are not necessary for identification of the sensitivities and, in the case of local identification, are if anything likely to be problematic. Consider the analogy of a country with one metropolis and ten much smaller cities. The logic of GLSS is to put almost all the weight on the metropolis in estimating the sensitivity of posted and effective price inflation to unemployment since it accounts for most of the expenditures in that country. But this is not productive. First, one can get better identification of the sensitivity by viewing each area as an equally valid source of information given that we use *local* variation in economic conditions. Then, once one has a good estimate of β that exploits all this information, one can construct aggregate expenditure-weighted measures that reflect the disproportionate influence of the metropolis. Second, economic conditions in the large metropolis will tend to be highly correlated with aggregate conditions, so it will provide little independent variation to identify β (since our specifications include time fixed effects to control for aggregate conditions). In contrast, smaller cities will be much more useful in this regard since they are more likely to experience local shocks that are not reflected in aggregate statistics.[6] These factors imply that, if anything, one should prefer specifications with equal weights rather than those with expenditure-weights as argued by GLSS.[7]

[6] One can readily verify that unemployment rates in large US cities are indeed more highly correlated with the aggregate unemployment rate than those of smaller cities.
[7] One exception to this argument would be if the β's were different across locations. In that case, weighted regressions might be preferable to recover the average β that applies at the aggregate level. However, GLSS provide no evidence of this kind of heterogeneity or reasons to expect this heterogeneity to be present or important.

More importantly, a satisfactory interpretation of the data should account for *why* different weighting schemes matter. In the baseline results of CGH, the specific weighting scheme used to aggregate or to estimate regressions has no effect on the results: they all yield the same qualitative (and quantitative) conclusions about the sensitivity of posted and effective price inflation to economic conditions. The sensitivity to the weighting scheme arises *only* as the censoring threshold is raised to the very high levels advocated by GLSS. But this type of sensitivity is exactly what one should expect if raising censoring thresholds serves mainly to introduce measurement error and outliers into the data, as argued in the previous section. Once large outliers are introduced, we would expect changing the weights on observations to lead to very different results depending on how much weight is assigned to specific values. Consistent with this logic, when the threshold is raised to the value suggested by GLSS, the estimates of the difference in sensitivity between posted- and effective-price inflation go up under some weighting classifications and down under others, even though the different weighting schemes yield nearly identical results under our baseline threshold.

One way to assess the extent to which these differences across weighting schemes reflect outliers introduced by the higher thresholds of GLSS is to employ methods that automatically identify and control for such outliers. In Table 2, we reproduce the baseline results of CGH and GLSS and re-estimate the specification using Huber robust regressions, which identify outliers and remove them from the estimation. When using the baseline threshold of CGH, the results from Huber regressions are almost identical to those of CGH regardless of the weighting scheme, which reflects the fact that there are few outliers with the CGH threshold. As the censoring threshold is raised, the results under Huber regressions remain very close to those of CGH for all weighting schemes. Weighted regressions no longer display any evidence that the sensitivity of the difference between posted and effective price inflation to unemployment diminishes with higher thresholds. Instead, the results using Huber robust regressions with the GLSS thresholds confirm the basic findings of CGH, indicating once again that outliers are driving the findings of GLSS.

II. BLS Price Imputation

Micro-level data sets are often rife with missing observations. The IRI scanner data are no exception, with nearly 40 percent of raw observations missing. How one addresses these observations can therefore affect results. CGH effectively impute zero inflation when price data are missing, which is a reasonable benchmark since the vast majority of prices do not change on a week-to-week basis (the probability of a regular price change on a typical week is about 0.05 in these data). If all missing observations are simply dropped, our empirical results are unchanged (see Appendix Table 3).

GLSS instead propose an imputation procedure like that used by the BLS in constructing the CPI and find that, with this imputation procedure, the response of posted price inflation to unemployment rises by a factor of three to four while the sensitivity of effective prices is unchanged. The GLSS imputation procedure works as follows. When an item has a missing weekly observation that is preceded by an observed price, GLSS compute the inflation rate for other goods in this category and location and apply it to the item for that missing week to impute its price. So if the price of a Gillette razor is missing at Target in Cleveland one week, they compute the average price change of razors in all Cleveland stores that week and adjust the price of the Gillette razor at Target by that same percentage. If observations are missing at random, this imputation procedure will serve mainly to smooth out measured price changes but should not otherwise affect the estimates. However, if the price changes of other goods are not a good proxy for those of the item whose price is missing because of systematic patterns in terms of which observations go missing, then this procedure could induce biases in the estimation.

To see how this can be, Table 3 presents a hypothetical example. Consider a good sold in two stores, A and B, where A is cheaper than B, such that in normal times store A charges a price of $99 while store B charges a price of $100. Consumers purchase 20 units of each good in each store in normal times, such that expenditure shares are 50-50. In a downturn, store A reduces its price to $90 while store B leaves its price unchanged, and households switch all of their expenditures to store A. As a result, the price at store B will not be recorded that period because of the store-switching behavior of households. In subsequent periods, prices and quantities return to their previous values. During the recession, true posted price inflation is -5 percent, a simple average of the 10 percent price decline in store A and zero percent change in store B. Effective price inflation would be -10 percent, reflecting not just the price decline in store A but also the switching of expenditures to the low-price retailer.

Now consider what this example implies when applying different imputation procedures. Following CGH, one would impute no change in price to store B, so posted price inflation would be correctly measured. With the imputation procedure of GLSS, one would use the inflation rate between the first two periods at store A to infer the recession price at store B, leading to an estimate of $91. Measured posted price inflation would be -10 percent, *the same as the effective rate of inflation and twice the true posted price inflation rate of -5 percent*. The attenuation of the difference between posted and effective price inflation due to the imputation procedure reflects two features of the hypothetical example: a) the expenditure switching toward the low-price store which generates a missing value in the high-price store during the downturn and b) a larger price decline in the low-price store during the downturn. Both of these features are present in the data.

To show the former, we calculate the incidence of non-missing values for a given month/store[8] and then regress this incidence on the local unemployment rate,

[8] For a given UPC/category/store/month, we calculate the share of weeks with non-missing observations. Then we aggregate this fraction across UPC/category to the store level.

the rank of the store (cheap vs. expensive) and the interaction of the store rank and local unemployment rate:

$$(2)\ M_{mst} = a_1 UR_{mt} + a_2 UR_{mt} \times \bar{R}_{mst,\Omega} + \alpha_3 \bar{R}_{mst,\Omega} + q_m + \lambda_t + error$$

where M_{mst} is the share of non-missing values for store *s* in market *m* at month *t*, q_m is the market *m* fixed effect, and λ_t is the time fixed effect. The rank of the store $\bar{R}_{mst,\Omega}$ is defined as in CGH and measures how far a store's average price level is from the median price level in a given market and month. That is, a positive value of $\bar{R}_{mst,\Omega}$ means that store s is relatively expensive.

Columns (1)-(3) in Table 4 show estimates of a_1 and a_2 for different weighting schemes. The key coefficient is a_2. Across all specifications it is strongly negative. These estimates suggest that when the unemployment rate is rising, prices are more likely to be missing in expensive stores relative to cheap stores. This result indicates that missing values are *not* missing at random but instead follow the predictions from systematic store-switching behavior by households.

The second feature of the data assumed in the hypothetical example, namely the differential pricing behavior of low- and high-price stores, has already been partly documented in CGH, where we emphasize how sales disappear in times of high unemployment primarily at high-price stores. Here, we show that prices more broadly behave differently at low- and high-price stores. Specifically, we regress store-level inflation rates on the local unemployment rate, the rank of the store (cheap vs. expensive) and the interaction of the store rank and local unemployment rate:[9]

[9] The store-level inflation at the monthly frequency is calculated as follows. First, we compute the average unit price for a given good/category/city/month and then calculate the percent change in the monthly prices. We winsorize price changes as in CGH. Second, we aggregate percent changes across goods using equal weights or expenditure shares to the category level. Third, we use expenditure shares to aggregate inflation rates at the category level to the store level.

(3) $\pi_{mst} = a_1 UR_{mt} + a_2 UR_{mt} \times \bar{R}_{mst,\Omega} + \alpha_3 \bar{R}_{mst,\Omega} + q_m + \lambda_t + error$

The key coefficient is again a_2 and across all specifications it is strongly positive, see columns (4)-(6) in Table 4. This estimate implies that, precisely as presented in our hypothetical example, prices decline more in low-price stores than in high-price stores during downturns.

Thus, when unemployment rates rise, expenditure-switching by households toward low-price stores leads to more missing observations at high-price stores. The imputation procedure of GLSS will then assign too large a decline in prices to these missing observations, pushing measured posted price inflation rates closer to effective price inflation rates. Store-switching behavior will therefore introduce a cyclical bias in the imputation that will reduce the difference in sensitivity of measured posted and effective price inflation to unemployment.

The fact that the imputation procedure introduces a cyclical bias into measures of posted price inflation has several implications. The first is that estimates using this imputation tell us little about the importance of store-switching behavior. If there is no store-switching behavior, then posted and effective price inflation will have the same cyclical sensitivity to unemployment, regardless of the imputation procedure. But if there is store-switching behavior, then the cyclical bias induced by the imputation procedure will again push estimates of the sensitivity of posted and effective price inflation toward the same values. Thus, finding the latter as in GLSS is ultimately uninformative about the presence of store-switching behavior.

A second implication is that one should be very careful with imputation procedures. When as much as 40 percent of the data is imputed, great care must be taken to understand what properties missing values might have and how the imputation might alter the qualitative and quantitative features of the data. As illustrated through our example in Table 3 and the empirical results in Table 4, the properties of the specific imputation undertaken by the BLS and GLSS can have pronounced implications for the cyclical behavior of measured inflation.

A third implication is that the BLS practice of imputing missing price data through the price dynamics of goods at other retailers entails that official measures of inflation already embody some elements of store-switching. In other words, official measures of inflation are already closer to effective price inflation than one might expect. Determining how large an effect this has on official statistics is an exercise that should be explored in future research.

III. Conclusion

While we do not agree with the conclusions of GLSS, we appreciate their effort at reproducing our results and considering alternative empirical choices. There are useful takeaways from this discussion that provide direction for future research. One is the paucity of guidance on how to deal with outliers. While it has become common to winsorize outliers, there is little consensus on how to balance the tradeoffs implied by different thresholds. We provide one such metric here, to compare the magnitudes of the attenuation bias emphasized by GLSS to the increased outliers introduced by higher thresholds. But more systematic tools to select thresholds optimally would be useful. The development and use of empirical methodologies that systematically address outliers, like Huber regressions, should also be encouraged.

References

Angrist, Joshua D., and Alan B. Krueger (1999) "Empirical strategies in labor economics," in: O. Ashenfelter & D. Card (ed.) *Handbook of Labor Economics*, vol. 3, ch. 23, pages 1277-1366.

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* 105(3), 993-1029.

Driscoll, John C. and Aart C. Kraay (1998) "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data," *Review of Economics and Statistics* 80(4), 549-560.

Gagnon, Etienne, David Lopez-Salido, and Jason Sockin (2017) "The Cyclicality of Sales, Regular and Effective Prices: Comment," *American Economic Review* 107(10): 3229-3242.

Malmendier, Ulrike, and Geoffrey Tate (2005) "CEO Overconfidence and Corporate Investment," *Journal of Finance* 60(6), 2661-2700.

Sharpe, Steven A., and Gustavo A. Suarez (2013) "The insensitivity of investment to interest rates: Evidence from a survey of CFOs," Finance and Economics Discussion Series 2014-2, Board of Governors of the Federal Reserve System.

Sufi, Amir (2009) "Bank Lines of Credit in Corporate Finance: An Empirical Analysis," *Review of Financial Studies* 22(3), 1057–1088.

Table 1. Root mean squared error of regressions with different censoring thresholds.

Weight used in aggregation across stores and UPCs	Weighted regression	Weights to cities	Censoring point, T				
			1 (baseline)	3	5	8	12 (GLSS)
unweighted	No	Equal	0.0179	0.0293	0.0342	0.0367	0.0381
city specific weights	No	Equal	0.0253	0.0371	0.0410	0.0429	0.0439
country weights	No	Equal	0.0252	0.0377	0.0424	0.0447	0.0458
city specific weights	Yes	Expend. share	0.0243	0.0335	0.0354	0.0361	0.0366
country weights	Yes	Expend. share	0.0239	0.0337	0.0361	0.0371	0.0377

Notes: The table reports root mean squared error in the following regression $\pi_{mct}^{eff} - \pi_{mct}^{posted} = \beta UR_{mt} + \lambda_t + \theta_{mc} + error$ where π_{mct}^{eff}, π_{mct}^{posted} are "effective price" and "posted price" inflation rates, m, c, and t index markets (e.g., Atlanta, Detroit), the category of the good (e.g., beer, coffee), calendar time (i.e., month); UR_{mt} is the local seasonally-adjusted unemployment rate; θ_{mc} denotes the fixed effect for each market and category of good while λ_t denotes time fixed effects. Columns 1 through 12 show the censoring point used to calculate "posted price" inflation rates. Rows (1)-(3) use equal weights for categories/cities. Rows (4) and (5) use the following weights across categories/cities: $\omega_{cmt} = \frac{TS_{cmt}}{\sum_c \sum_m TS_{cmt}}$ where m, c, t index market (city), category, and time (month), TS is the volume of sales. The root mean squared errors for T=2 through T=12 are statistically different from the root mean squared errors for T=1 at 1 percent level.

Table 2. Estimates of the difference in the sensitivity of posted and effective price inflation to unemployment rate

Weight used in aggregation across stores and UPCs	Weighted regression	Weights to cities	Statistic	OLS		Huber Robust Regressions	
				T=1 (CGH baseline)	T=12 (GLSS baseline)	T=1 (CGH baseline)	T=12 (GLSS baseline)
				(1)	(2)	(3)	(4)
(1) unweighted	No	Equal	difference	-0.158	-0.283	-0.178	-0.339
			(s.e.)	(0.022)	(0.043)	(0.012)	(0.020)
			p-val (diff. = zero)	0.000	0.000	0.000	0.000
			p-val (T=1 eq. T=X)		0.000		0.000
(2) city specific weights	No	Equal	difference	-0.123	-0.073	-0.114	-0.124
			(s.e.)	(0.033)	(0.053)	(0.020)	(0.029)
			p-val (diff. = zero)	0.000	0.165	0.000	0.000
			p-val (T=1 eq. T=X)		0.207		0.703
(3) country weights	No	Equal	difference	-0.129	-0.085	-0.132	-0.121
			(s.e.)	(0.033)	(0.054)	(0.020)	(0.030)
			p-val (diff. = zero)	0.000	0.117	0.000	0.000
			p-val (T=1 eq. T=X)		0.256		0.607
(4) city specific weights	Yes	Expend. share	difference	-0.083	0.047	-0.067	-0.069
			(s.e.)	(0.036)	(0.053)	(0.013)	(0.017)
			p-val (diff. = zero)	0.019	0.374	0.000	0.000
			p-val (T=1 eq. T=X)		0.002		0.910
(5) country weights	Yes	Expend. share	difference	-0.087	0.011	-0.088	-0.113
			(s.e.)	(0.039)	(0.064)	(0.014)	(0.015)
			p-val (diff. = zero)	0.025	0.865	0.000	0.000
			p-val (T=1 eq. T=X)		0.045		0.023

Notes: The table reports results for the difference in the sensitivity of effective and posted price inflation to unemployment rate using the OLS regressions (columns 1 and 2) and the Huber robust regression (robust to outliers; columns 3 and 4). Driscoll-Kraay (1998) standard errors are reported in parentheses. *difference* shows the difference in the estimated sensitivity. p-val (diff. = zero) is the p-value for the test that the difference is equal to zero. p-val (T=1 eq. T=X) shows p-value for the test that the differences estimated for censoring threshold T=1 is equal to the difference estimated for censoring threshold T=X, where X={3,5,8,12}. Rows (1)-(3) use equal weights for categories/cities. Rows (4) and (5) use the following weights across categories/cities: $\omega_{cmt} = \frac{TS_{cmt}}{\sum_c \sum_m TS_{cmt}}$ where m, c, t index market (city), category, and time (month), TS is the volume of sales.

Table 3. Fictitious case of BLS-style imputation

	Period		
	1 (expansion)	2 (recession)	3 (expansion)
Actual posted prices			
outlet A	99	90	99
outlet B	100	100	100
Quantities bought			
outlet A	20	40	20
outlet B	20	0	20
Recorded prices in scanner data			
outlet A	99	90	99
outlet B	100	X	100
Inflation (log) in actual posted prices			
outlet A (percent)		-10	10
outlet B (percent)		0	0
Imputed price in outlet B			
GLS imputation	100	91	100
CGH imputation	100	100	100
Effective price	99.5	90	99.5
Effective price inflation (log, in percent)		-10	10
Posted price inflation (log, in percent)			
Actual		-5	5
CGH imputation		-5	5
GLS imputation		-10	10

Table 4. Sensitivity of incidence of store-level non-missing values and inflation to local unemployment rate by store relative prices

Dependent variable	Share of non-missing values		Inflation	
	UR	UR×$\bar{R}_{mst,\Omega}$	UR	UR×$\bar{R}_{mst,\Omega}$
	(1)	(2)	(3)	(4)
Panel A. Unweighted				
Unweighted	0.230	-6.836	-0.117	0.222
	(0.044)	(0.516)	(0.017)	(0.110)
City specific weights	0.362	-8.758	-0.125	0.144
	(0.038)	(0.606)	(0.016)	(0.129)
Country weights	0.229	-8.472	-0.134	0.225
	(0.044)	(0.608)	(0.018)	(0.155)
Panel B. Expenditure weighted (a store's weight is relative to city-level expenditures)				
Unweighted	0.118	-5.422	-0.139	0.282
	(0.031)	(0.459)	(0.018)	(0.068)
City specific weights	0.072	-5.151	-0.155	0.233
	(0.039)	(0.480)	(0.016)	(0.064)
Country weights	-0.054	-5.042	-0.165	0.340
	(0.038)	(0.468)	(0.022)	(0.084)
Panel C. Expenditure weighted (a store's weight is relative to national expenditures)				
Unweighted	0.118	-5.422	-0.139	0.282
	(0.031)	(0.459)	(0.018)	(0.068)
City specific weights	0.072	-5.151	-0.155	0.233
	(0.039)	(0.480)	(0.016)	(0.064)
Country weights	-0.054	-5.042	-0.165	0.340
	(0.038)	(0.468)	(0.022)	(0.084)

Notes: The table reports Huber-robust estimates for specifications (2) and (3) in columns (1)-(3) and (4)-(6) respectively. The dependent variable in columns (1)-(3) is the share of non-missing values in a month in a given store/market. The dependent variable in columns (4)-(6) is the annual inflation rate in a given store/market. All dependent variables are winsorized at top and bottom one percent. The set of goods used for ranking is $\Omega_{\max}$, that is goods that are sold in all stores in a given metropolitan area. Driscoll-Kraay (1998) standard errors are reported in parentheses.

Online Appendix Tables

Appendix Table 1. Share of censored price changes out of price changes by type of price change and threshold.

row	Weight used in aggregation across stores and UPCs	Weighted regression	Weights to cities	Price changes	Censoring point, T											
					1	2	3	4	5	6	7	8	9	10	11	12
(1)	unweighted	No	Equal	sales	0.827	0.601	0.428	0.299	0.191	0.136	0.093	0.066	0.023	0.015	0.010	0.007
				regular price	0.359	0.184	0.109	0.069	0.043	0.032	0.024	0.019	0.008	0.006	0.005	0.004
				all	0.782	0.562	0.398	0.277	0.177	0.126	0.087	0.062	0.022	0.014	0.009	0.007
(2)	city specific weights	No	Equal	sales	0.819	0.602	0.432	0.304	0.194	0.135	0.090	0.060	0.020	0.012	0.008	0.005
				regular price	0.295	0.140	0.078	0.047	0.027	0.018	0.012	0.009	0.004	0.003	0.002	0.002
				all	0.773	0.561	0.401	0.281	0.179	0.125	0.083	0.056	0.019	0.012	0.007	0.005
(3)	country weights	No	Equal	sales	0.823	0.607	0.438	0.309	0.198	0.137	0.091	0.061	0.020	0.012	0.007	0.005
				regular price	0.300	0.144	0.081	0.049	0.028	0.019	0.013	0.009	0.004	0.003	0.002	0.002
				all	0.778	0.567	0.407	0.286	0.183	0.127	0.084	0.056	0.019	0.011	0.007	0.005
(4)	city specific weights	Yes	Expend. share	sales	0.812	0.603	0.436	0.313	0.204	0.141	0.092	0.060	0.021	0.012	0.007	0.004
				regular price	0.266	0.121	0.063	0.036	0.019	0.012	0.007	0.005	0.002	0.002	0.001	0.001
				all	0.765	0.562	0.404	0.290	0.188	0.130	0.085	0.056	0.019	0.011	0.006	0.004
(5)	country weights	Yes	Expend. share	sales	0.815	0.607	0.441	0.318	0.207	0.143	0.093	0.061	0.020	0.011	0.006	0.004
				regular price	0.269	0.123	0.065	0.038	0.020	0.013	0.008	0.005	0.002	0.002	0.001	0.001
				all	0.770	0.567	0.410	0.294	0.192	0.132	0.086	0.056	0.019	0.010	0.006	0.004

Notes: The table reports the share of censored price changes in all price changes. The censoring point X sets $(dlogP)*12=-X$ if $(dlogP)*12<-X$ and $(dlogP)*12=X$ if $(dlogP)*12>X$. Rows (1)-(3) use equal weights for categories/cities. Rows (4) and (5) use the following weights across categories/cities: $\omega_{cmt}=\frac{TS_{cmt}}{\sum_c\sum_m TS_{cmt}}$ where m, c, t index market (city), category, and time (month), TS is the volume of sales.

Appendix Table 2. Share of censored price changes out of non-missing price quote observations by type of price change and threshold.

row	Weight used in aggregation across stores and UPCs	Weighted regression	Weights to cities	Price changes	Censoring point, T											
					1	2	3	4	5	6	7	8	9	10	11	12
(1)	unweighted	No	Equal	sales	0.141	0.105	0.076	0.053	0.034	0.024	0.016	0.012	0.004	0.003	0.002	0.001
				regular price	0.015	0.007	0.004	0.003	0.002	0.001	0.001	0.001	0.000	0.000	0.000	0.000
				all	0.129	0.096	0.069	0.049	0.031	0.022	0.015	0.011	0.004	0.002	0.002	0.001
(2)	city specific weights	No	Equal	sales	0.157	0.119	0.086	0.061	0.039	0.027	0.018	0.012	0.004	0.002	0.002	0.001
				regular price	0.013	0.005	0.003	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
				all	0.144	0.109	0.079	0.056	0.036	0.025	0.016	0.011	0.004	0.002	0.001	0.001
(3)	country weights	No	Equal	sales	0.160	0.121	0.089	0.063	0.040	0.028	0.018	0.012	0.004	0.002	0.001	0.001
				regular price	0.013	0.005	0.003	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
				all	0.147	0.111	0.081	0.058	0.037	0.025	0.017	0.011	0.004	0.002	0.001	0.001
(4)	city specific weights	Yes	Expend. share	sales	0.180	0.138	0.101	0.073	0.047	0.033	0.021	0.014	0.005	0.003	0.001	0.001
				regular price	0.013	0.005	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
				all	0.165	0.126	0.093	0.067	0.043	0.030	0.019	0.012	0.004	0.002	0.001	0.001
(5)	country weights	Yes	Expend. share	sales	0.182	0.140	0.103	0.075	0.049	0.033	0.021	0.014	0.005	0.002	0.001	0.001
				regular price	0.013	0.005	0.003	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
				all	0.167	0.129	0.095	0.069	0.045	0.031	0.020	0.013	0.004	0.002	0.001	0.001

Notes: The table reports the share of censored price changes in all non-missing price quotes. The censoring point X sets $(dlogP)*12=-X$ if $(dlogP)*12<-X$ and $(dlogP)*12=X$ if $(dlogP)*12>X$. Rows (1)-(3) use equal weights for categories/cities. Rows (4) and (5) use the following weights across categories/cities: $\omega_{cmt}=\frac{TS_{cmt}}{\sum_c\sum_m TS_{cmt}}$ where *m*, *c*, *t* index market (city), category, and time (month), TS is the volume of sales.

Appendix Table 3. Effect of CGH imputation for the sensitivity of posted-price inflation to local unemployment rate.

Weight used in aggregation across stores and UPCs	Censoring point				
	1 (baseline)	3	5	8	12
	(1)	(2)	(3)	(4)	(5)
Panel A: No imputation					
Unweighted	-0.0763	-0.119	-0.148	-0.179	-0.183
	(0.0218)	(0.0450)	(0.0580)	(0.0681)	(0.0701)
City specific weights	-0.0853	-0.129	-0.163	-0.201	-0.207
	(0.0242)	(0.0425)	(0.0506)	(0.0579)	(0.0591)
Country weights	-0.0850	-0.127	-0.158	-0.198	-0.208
	(0.0272)	(0.0481)	(0.0573)	(0.0652)	(0.0665)
Panel B: Imputation (Baseline)					
Unweighted	-0.061	-0.098	-0.123	-0.150	-0.155
	(0.017)	(0.034)	(0.043)	(0.051)	(0.053)
City specific weights	-0.077	-0.118	-0.149	-0.185	-0.192
	(0.021)	(0.035)	(0.042)	(0.048)	(0.049)
Country weights	-0.075	-0.114	-0.142	-0.180	-0.191
	(0.023)	(0.040)	(0.048)	(0.055)	(0.056)
Observations	187,426	187,426	187,426	187,426	187,426
Number of groups	1,550	1,550	1,550	1,550	1,550

Notes: The table reproduces table 1 in CGH (2015) for different values of the truncation point. The table reports estimated coefficients on local unemployment rate when we regress a measure of posted-price city/category inflation on local unemployment rate after controlling for city/category and month fixed effects. The truncation point X sets $(dlogP) * 12 = -X$ if $(dlogP) * 12 < -X$ and $(dlogP) * 12 = X$ if $(dlogP) * 12 > X$ for a price change at the level of good/store/category/city. Panel A shows results when no imputations are used. Panel B shows results for the approach used in CGH (2015). Driscoll-Kraay (1998) standard errors are reported in parentheses.