Trading Down and Inflation*

Omiros Kouvavas†
University of Warwick
& European Central Bank

February 10, 2020

Job Market Paper

Abstract

Since the great recession the puzzle of missing disinflation (inflation) during downturns (and the recovery phases) of the business cycle has been at the central stage of academic and policy research. In this paper we suggest that part of the explanation is related to trading down during the business cycle. This consumption choice of quality, endogenous to the cycle, might bias the inflation profile leading to less responsive reported prices. The switch of individuals from lower quality goods, during downturns, and higher quality goods, during upturns, leads to changes of the composition of consumption correlated to the business cycle. In this paper we document the extent of the shift in consumption habits using detailed supermarket scanner data from the United States spanning from before the Great Recession until the recovery and estimate the effect of observed aggregate inflation rates. The use of machine learning allows us to classify a large scale dataset of products to different product qualities and track consumption patterns through time. The results suggest that accounting for quality prices demonstrates significantly higher cyclicality.

JEL codes: E2, E3, E4
Keywords: Inflation bias, quality choice, business cycles

***Please do not quote without permission. The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen.**

†University of Warwick, Department of Economics, email: o.kouvavas@warwick.ac.uk

1 Introduction

Since the beginning of the Great Recession the observed responsiveness of inflation to the business cycle has been greatly diminished. Both during the downturn and the recovery years we have observed the puzzle of missing disinflation and inflation. Several efforts have been made in the literature to address this puzzle. Part of the literature has focused on the "anchored expectation" hypothesis of of Bernanke (2010) where inflation is stabilized by the credibility of modern central banks and the anchoring of expectations as an explanation. Another part of the literature has focused on developments on the labour markets and the missing pressure to wages. Finally, there is a part of the literature that focuses on the flattening of the Philips curve which was pointed out by the International Monetary Fund (IMF 2013), however many have questioned the lack of structural changes in the economy that would justify such a result (e.g. Coibion and Gorodnichenko (2015)).

We are proposing that part of the explanation is related to consumption behaviour changes that are correlated with the business cycle leading to significant aggregation bias. Trading down describes the phenomenon that consumers tend to choose lower quality products during economic downturns and vice versa. We argue that trading down on product quality leads to an aggregate inflation bias that flattens the inflation profile. Furthermore, the nature of the bias is different from a simple composition bias, as there is an endogenous effect to the price of the different quality products due to the cyclical changes in consumption.

The literature on inflation bias has mainly focused on the effect of **i)** sales, **ii)** store switching and **iii)** unaccounted technological change in quality to explain the puzzle. Sales have been considered early on as a potential source of bias to the measurement of inflation. Eichenbaum et al. (2011), Guimaraes and Sheedy (2011) and Kehoe and Midrigan (2015) focus on sales as a source of effective price flexibility, however there is lack of empirical evidence that would justify the sales effects as being the driving force behind the limited responsiveness of inflation to the business cylce. On the other hand, Coibion et al. (2015) focus on store switching which seems a more promising approach, albeit it tends to have short lived effects around the recession and cannot explain the low inflation on the upturn[1]. Furthermore, store switching as a measurement error has been disused by the Boskin Commission Report Boskin et al. (1996); also Shapiro and Wilcox (1996) mentions store switching as one of possible sources of measurement bias to the Consumer Price Index (CPI). In this paper we show that when even accounting for sales and store switching, trading down increases price cyclicality.

Until now the literature has considered the quality of products as a possible bias to inflation, however not endogenous to the business cycle. Bils (2009) describes how improvements in newer product models might reflect a significant part of the price increase that we miss-

[1] As consumers tent to return fairly quickly to the original store, thus ignoring this effect might justify the missing dis-inflation in the beginning of the crisis but not the missing inflation in the growth years. In essence we need something more slow moving and symmetric across the business cycle.

attribute to inflation. Crucially, this is a long-term trend of technological improvement and not a choice of quality of the product to consume due to short term constrains correlated with income. Thus, Bils (2009) argues that inflation is lower than currently measured as the products of one year are not comparable to the previous due to improvements. In a parallel vein Jaimovich et al. (2015) partially document the phenomenon of trading down during the great recession, finding a significant change in consumer choices. However, they do not focus on possible effects on measured inflation as they argue that it might lead to an amplification of the business cycle due to labour intensiveness of different product qualities. Hence our goal in this paper is to combine the literature on missing inflation and the phenomenon of trading down in qualities to argue that this partially explains the development of the inflation schedule since the Great Recession.

An obvious question is: Why is this bias significant now and not in the past? This question is common for the literature involving compositional changes of any sort. The reasoning goes as follows: if trading down is a structural phenomenon it must have occurred in previous cycles hence it cannot explain the difference in the current cycle. However this argument relies on the assumption that the magnitude of the phenomenon remains the same over time. In reality, in the product quality space we have observed a rapid expansion of offered qualities and varieties which in turn can accommodate incremental shifts across quality types. In a nutshell, during previous business cycles the available product quality varieties were significantly less, leading to a smaller bias. Furthermore what is important is the distribution of consumer preferences across the qualities, as the quality switching might be due to constraints. As such it is important that a significant part of consumers becomes constrained and thus switches in combination with them being on higher qualities(in a down-turn). However, if part of the population already consumed lower qualities then the only adjustment margin left is a reduction in quantity. Thus, trading down has become more important in the recent years as both the amount of available qualities has increased as well as the portion of consumers who consume higher qualities (before the downturn).

The rest of the paper will be structured as follows: in section 2 we will describe the trading down effect and the arising inflation bias. In section 3 we will describe the data (subsection 3.1), methods that will be used to document the phenomenon along with a discussion on price censoring which recently became very prominent. In section 4 we discuss the results. In detail subsection 4.2 discusses the cyclicality of prices looking at posted prices, effective prices and sales. In subsection 4.3 we introduce the quality choice to the cyclicality of prices and contrast it with the store switching case. section 5 concludes the paper.

2 Tracing the Trading Down Effect

Consumer consumption choices and the impact on the economy is one of the classic research areas in macroeconomics. Following Burns and Mitchell (1946) an extensive empirical literature focused on the impact of consumption choices on the business cycle. Trading down on the other hand has only been recently brought into focus by Jaimovich et al. (2015) when they showed that during the last recession there was significant trading down of qualities of products and services consumed[2]. However, for the most part they used sector level data. Furthermore, they only focused on the effect of trading down on the business cycle as a two stages process. Firstly, they documented the effect of the recession on trading down and secondly they argued that products of lower quality use less labour to be produced thus trading down leads to less labour needed in equilibrium and the amplification of the business cycle. On the contrary we will simply argue that if the business cycle leads to trading down in quality, this in turn will bias the measurement of observed aggregate inflation. Furthermore, their classification of qualities heavily relies on the prices of the goods which are eventually correlated with their outcome variable. We will abstract from the price as a classifier and use supervised machine learning to classify product into different quality categories.

We are contributing to the existing literature in two main ways: first, we use detailed scanner data ,that cover over half of total sales volume and over 30% of mass merchandise sales volume, to document the trading down phenomenon for the period 2006-2017. Second, this is the first study to consider implications of this endogenous consumption quality choice for aggregate inflation and possible measurement issues that arise.

2.1 Trading Down and Inflation

The idea of trading down boils down to adding an extra margin to quantity on the consumption adjustment due to budget constrains. This extra margin is quality. Thus the consumer can choose either to adjust the quantity or the quality of the product they consume. Given that the adjustment of quality and quantity will be correlated with income that will results to readjustments of the quantities of different products demanded belonging to different quality categories that will also be correlated with the business cycle.[3]

In a nutshell the re-adjustment of demand between products of different quality categories will lead to an inflation bias correlated with the business cycle. As such, measured inflation will tend to decrease less during downturns and increase less during upturns leading to a dampened inflation profile.

3 Documenting Trading Down

3.1 Data Description

The dataset used to document and estimate the trading down and its effect on inflation is the Retail Scanner Data collected and provided by the Nielsen marketing group and managed

[2] Table 1 shows the results per category

[3] This will also affect the prices of the respective categories leading to a secondary bias.

by the Kilts center for marketing at the University of Chicago[4]. The data covers 35000 stores including drug and grocery stores along with mass merchandisers that belong to over 90 retail chains spanning across 55 metropolitan statistical areas in the U.S. and range from 2006 to 2017. The detailed micro-data include both prices and quantities of purchases for 2.6 million unique UPC[5] codes. The availability of both prices and quantities is crucial if we want to document changes in quantities of qualities sold and as such scanner data are the only viable source of data in direct contrast with CPI data that include only prices or quantities on the aggregate level (e.g. sectors). There are 1100 products categories classified into 10 general groups: alcoholic beverages, dairy, deli, dry groceries, fresh food, frozen, general merchandise, health and beauty, meat and non-food. The variables we will be focusing on are units of packages sold and the overall price of the package, allowing us to construct the price per unit. The coverage of the dataset is around half of the sales volume of drug and grocery stores and higher than 30% of all US merchandiser sales volume. An additional advantage is the frequency of the data, which is recorded on a weekly basis. This is very rare for economic indicators especially if we factor in the location availability. In detail, additional information includes location data such as zip and fips county codes as well as the retailer id. However, the retailer name is not disclosed. Furthermore, the data do not include services and durables which is a disadvantage as they tend to be more cyclical. Finally, the data display strong seasonality patterns which, however, are minimised at higher levels of aggregation.

Although the Nielsen data have been used widely in marketing research of specific products, only recently efforts have been made to use them to answer more macro oriented questions. Even when scanner data are used, there are very few studies that use the native weekly frequency of the data or analyse both price and quantity (e.g. **?** use only a random sample of 30 product categories). The reason is mainly that the size of the data which exceeds 100 billion observations complicates the use and the analysis, due to very high computational cost, at least when using all the information contained in the dataset. As such to use the Nielsen scanner data was a challenge as well as an opportunity to check for evidence on a more granular level in order to show how the aggregate biases arise.

Sales are partially flagged in the dataset but not consistently. Thus, we use the algorithm of **?** and **?** to compute regular prices as a cross check. Our primary unit of analysis is defined as a product, in detail the unique store-UPC combination for a given week. Retailers report the dollar price of weekly sales (TR) as well as the total quantity sold (TQ). The combination results to the average retail price for that week for a product:

$$P_{asctu} = \frac{TR_{asctu}}{TQ_{asctu}} \tag{1}$$

[4]For previous research using the same data please refer to: https://research.chicagobooth.edu/nielsen/working-papers

[5]A UPC is a universal product code. A unique identifier assigned to retail item.

where a, s, c, t and u index areas, stores, categories of products, time and the UPC codes. We follow the literature and refer to this measure as posted prices. Furthermore, we compute product specific monthly inflation rates defined as:

$$\pi_{asctu} = log(\frac{P_{ascut}}{P_{ascu,t-1}}) \tag{2}$$

where t now denotes the month. For both prices and price inflation we also construct the aggregates using **(i)** equal weights, expenditure shares for each market and finally cumulate montly inflation rated into annual inflation rates.[6]Hence, the effective prices taking into account quantities sold are:

$$P^{eff}_{acut} = \frac{\sum_{s\in m} TR_{actu}}{\sum_{s\in m} TQ_{actu}} \tag{3}$$

This measure can change if individual prices change, if consumer reallocate consumption across stores and it also takes into account quality changes. In a nutshell, the difference between effective prices and posted prices will indicate the direction of the bias with respect to the cyclicality, but will not effectively explain its cause. In contrast with Coibion et al. (2015) we will show that the bias mainly comes from switching qualities of goods consumed both within stores and across stores. Hence, they only partially capture the effect by only focusing on store switching. Or, to put it differently, most of the reallocation observed in the literature is not within the same UPC, as it was hypothesized, but mainly reallocation across different qualities. We also construct a measure that does not include substitution across goods ($P^{eff_NS}_{acut}$) to be able to directly compare it with previous results. Finally we aggregate monthly to annual inflation rates $\bar{\pi}^{eff}_{act}$.

One important issue is the case of sales. We decompose the prices into regular price changes and sales. To do this we follow two different approaches: First, we utilise the information on sales provided by the retailer. Second, we use the algorithm along the lines of **?** and **?** to decompose to regular and sales prices. Both methods lead to similar results.

3.1.1 Censoring and Price Changes

An important issue discussed in the literature is whether or not to censor price changes above a percentage as they might effectively introduce noise. In Coibion et al. (2015) they censor all annual price adjustment for values that the log price movements exceed 1 on an annualized basis. This implies censoring values of price increases/decreases of above 100% on an annual basis or around 8.3% on a monthly basis.

[6] We opted to use similar aggregation with Coibion et al. (2015) to replicate their results on effective prices and for comparability reasons.

According to Gagnon, Lopez and Sockin this aggressive censoring is mainly driving the differences in the price cyclicality and they propose a different weighting scheme along with minor to no censoring following their preferred methodology. They show that avoiding censoring for what CGH (Coibion et al. (2015), Coibion et al. (2019)) call outliers reduces the difference between spot and effective price inflation arguing that store switching is neither statistically nor economically significant.

On the contrary Gagnon et al. (2017) argue that the standard argument for the need of censoring is simple due to 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 σ increases in T where T is the censoring threshold.

There is a simple test proposed by CGH in order to support lower censoring threshold. Given the fact that the scanner data that we use come from a different source than CGH that use the IRI data, we have implemented the test to see if the same augment can be applied. In detail suppose that the true posted inflation (π^{p^*}) and effective are each related to unemployment as follows:

$$\pi^f = \beta_f u + \epsilon^f \tag{4}$$

$$\pi^{p^*} = \beta_p u + \epsilon^p \tag{5}$$

where ϵ are the shocks to each process and could be correlated. Following the measurement error explanation, effective prices should be more sensitive to unemployment which implies that $\beta_f < \beta_{p^*} < 0$. Thus the following regression will yield estimates of β and recover the sensitivity to unemployment ($\hat{\beta} = \beta_f - \beta_{p^*}$):

$$\pi^f - \pi^p = \beta u + \epsilon \tag{6}$$

The argument goes as follows: the measurement error on the left hand side of equation (6) should intuitively not affect the properties of $\hat{\beta}$, however the residuals will be increasing in the censoring threshold ($var(\epsilon) = var(\epsilon^f - \epsilon^{p^*}) + var(\theta))$). On the contrary the prediction of the attenuation interpretation implies that the variance of the residuals will be decreasing when we raise the censoring threshold. Hence, this provides us with a simple test that might clarify if censoring reduces noise or introduces attenuation bias.

Table 1 reports the results of this exercise pointing to a similar conclusion as in CGH. Columns (2) through (12) represent the censoring point used to calculate the posted prices inflation rate. The root mean squared error results from the following regression:

$$\pi^{eff}_{act} - \pi^{posted}_{act} = \beta UR_{at} + \lambda_t + \theta_{ac} + error \tag{7}$$

Table 1: Root Mean Squared Error Using Different Censoring Thresholds

			Cencoring point, T				
Weights in aggregation across stores and UPCs	Weighted regression	Weights to cities	CGH (1)	(3)	(5)	(8)	GLSS (12)
1. Unweighted	No	Equal	0.0183	0.0306	0.0338	0.0371	0.0389
2. City-specific weights	No	Equal	0.0223	0.0365	0.0431	0.0442	0.0448
3. County weights	No	Equal	0.0226	0.0376	0.0430	0.0447	0.0451
4. City-specific weights	No	Equal	0.0219	0.0327	0.0346	0.0348	0.0412
5. County weigths	No	Equal	0.0216	0.0318	0.0321	0.0335	0.0375

Notes: The table is reporting the root mean squared error of the regression $\pi^{eff}_{act} - \pi^{posted}_{act} = \beta UR_{at} + \lambda_t + \theta_{ac} + error$ are the effective and posted inflation rates similar to CGH. The indexation a, c and t refer to area, the category of the good and calendar time; UR refers to the local unemployment rate in that area, seasonally adjusted. θ_{ac} denotes the fixed effect for each market and category of the good and λ_t denotes time fixed effects.

where the indexation a, c and t refer to area, the category of the good and calendar time; UR refers to the local unemployment rate in that area, seasonally adjusted. θ_{ac} denotes the fixed effect for each market and category of the good and λ_t denotes time fixed effects.

The results, with respect to the root mean squared error estimation, show that the errors increase with the increase of the censoring point. All estimated root mean squared errors reported are statistically significant from the base case. Hence, using this simple test to evaluate the need of censoring and avoid noise such as measurement error leads us to the conclusion that a low censoring point is appropriate. Following this we will also follow GCH and censor price changes that the log price movement exceeds 1. This also allows us to directly compare the results to the store switching case of Coibion et al. (2015).

3.2 Quality Classification

In this section we will focus on methods used to classify products into qualities with special focus on the challenges and the current state of the literature. Quality has been recently considered important in understanding consumption choices. However, there is no literature that classifies products into quality categories on a large scale. Previous literature has focused mainly on classifying retailers into discount and not, or different types of services into low and high quality. The challenge is that is impossible to manually classify over 2 million unique UPCs referring to different products. Thus the difficulty of credibly identifying quality is likely one of the reasons why the literature focussing on scanner data on the topic is so scarce. Furthermore, in several cases the literature has tried to make use of *coresets* or *sketches* of the original data that preserve the information of the bigger dataset (Agarwal et al., 2004)[7]. However, a random sample, even if we are really confident that it is representative, significantly

[7] We will refrain form elaborating here as such issues and methods along with there application on scanner data are described in a very analytical way by Ng (2017).

Table 2: Ratio of High to Low Quality Goods Sold for Broad Product Categories

	Ratio H/L quality goods for year periods			
Product Categories	2006-2008	2009-2011	2012-2014	2015-2017
Alcoholic Beverages	0.431	0.217	0.334	0.463
Dairy	0.216	0.101	0.213	0.327
Deli	0.432	0.234	0.287	0.441
Dry Groceries	0.324	0.308	0.327	0.458
Fresh Food	0.192	0.081	0.237	0.289
Frozen	0.168	0.038	0.089	0.194
General Merchandise	0.327	0.126	0.179	0.385
Health and Beauty	0.375	0.183	0.189	0.468
Meat	0.261	0.094	0.158	0.227
Non-Food	0.371	0.132	0.156	0.417

Notes: Products were classified into categories using a trained CNN using all product characteristics excluding price as a feature set complemented by text analysis of the web scraped product description. The classification was done within very narrow product categories and then results were aggregated up using the product/category revenue share in total revenue.

reduces the power of the estimation. Furthermore, in most cases there is a restriction on the number of product categories and that might bias the results. Hence, in our case we have opted to use the full data and utilise methods that effectively automatize the classification of products into quality categories.

3.2.1 The Importance of Identifying Quality

The difficulty of credibly identifying quality is probably one of the reasons why the literature of the effects of changes of consumption with respect to quality is relatively small. In this section we will review possible alternative methods tackling the issue. Namely we will focus on the most common case of price as a proxy for quality, the use of external sources and finally a classifier using machine learning methods.

Price as a proxy of quality. This is probably the most commonly used method in the literature, as it is easy to implement and relies on readily available information included in the data. Jaimovich et al. (2015) use this method to identify different price ranges that indicate different quality products. The fundamental problem with this method is that price is endogenous to both quality and demand and as such makes it an imperfect proxy. Furthermore, price and price changes are eventually our target variable, making it even less credible to follow that path.

External Quality Measures. Such measures of quality include product reviews or ratings implemented in a standardized way. One possible example is reviews presented on websites like Yelp!, which can be attained by web scraping and mostly refer to the quality of a specific service provider. A similar logic could be implemented to quality of products. This measure obviously reflects costumer perceived product quality, however this might be biased by the selected sample that provides a review. In our case we will only partially utilize this technique to create one of our features used for machine learning. In practice, we have web scraped UPC look up websites to retrieve the product description text, which in many cases contains information useful for classifying the quality of the products.

Machine Learning Methods. Probably one of the most promising techniques is to use machine learning methods to classify the products into different qualities. The reason that machine learning methods become more and more popular in economics is that they allow for easy and credible classification of data into categories that can then be used for further analyses. Such methods have been widely used to retrieve information from texts analyses other types of Big Data that does not allow for the classification to be done manually. Furthermore, allowing the machine to do the classification based on certain features of the data provides a degree of credibility that is not attainable with ad hoc classification by the researcher. Classification of qualities using machine learning can be done using either supervised or unsupervised methods[8].

3.2.2 Classifying Quality using Convolutional Neural Networks (CNNs)

In this part we will focus on the application of machine learning using Convolutional Neural Networks to classify the products into qualities. We will first briefly introduce the concept of supervised learning and its benefits. Then we elaborate on the algorithm used, performance evaluation and reproducibility. Finally we will show some resulting aggregates with respect to percentages of high/low quality products sold before the crisis, during the crisis and in the recovery period.

Convolutional neural networks are based on back-propagation in a feed forward network with many hidden layers. Given the structure of our data using CNNs is an ideal approach. The main reasoning is that we have opted for the use of supervised methods where the researcher has to label manually a small random part of the data and use that for training and testing purposes.[9] The feature set included all available product characteristics, the actual UPC in partial abstracts as it contains information on the company. In some cases we complemented the feature set with a web scraped description of the product marked for quality. The resulting classifier performed well over 90% for the test set.

[8] Supervised methods refer to the fact that the researcher could classify a small portion of the data and then allow a trained algorithm to classify the rest, which is particularly useful when we are dealing with very large scale data as in our case. Unsupervised methods allow the algorithm to decide the classification on each own, and the researcher can impose the number of categories or even allow this to be automatically identified

[9] In early stages of the project we have experimented with unsupervised methods, however, they seemed to only be accurate when taking price into account and when the narrow product category that they were used in was very well defined and contained homogeneous products.

Table 3: Cyclical Properties of Price Changes

Area x cat. fixed eff.	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed eff.	No	No	Yes	trend	Yes	Yes	Yes	Yes
Weigted regression	No	No	No	No	No	No	Yes	Yes
					Aggregation using expenditure shares			
Depend. variable	Equal weight in UPCs aggregation				Market Specific	Common	Market Specific	Common
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Inflation:								
Posted	$-.097^{***}$	$-.0963^{***}$	$-.074^{***}$	$-.117^{***}$	$-.086^{***}$	$-.085^{***}$	$-.084^{***}$	$-.091^{***}$
	(.018)	(.019)	(.013)	(.022)	(.012)	(.013)	(.011)	(.016)
p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Effective	$-.145^{***}$	$-.152^{***}$	$-.305^{***}$	$-.331^{***}$	$-.279^{***}$	$-.285^{***}$	$-.121^{***}$	$-.164^{***}$
	(.033)	(.042)	(.031)	(.035)	(.019)	(.032)	(.028)	(.026)
p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Notes: The table reports the estimated coefficients on seasonally adjusted unemployment. Column 4 controls for linear trends. Number of observations is 465402 . The last two columns report the regression results where observations are weighted by the average expenditure share of a given area/category cell in total spending across all cities and categories, taking into account both the price level and quantity sold. Following CGH we report the Driscoll and Kraay (1998) standard errors in parentheses.
*** Significant at 1 percent level
** Significant at 5 percent level
* Significant at 10 percent level

Table 2 reports the resulting ratios of high to low quality goods sold for specific time periods aggregated for 10 broad categories. For the aggregation we have used the revenue share of this products. The results are very similar across different aggregation schemes. In all categories we observe a clear cyclicality. The categories demonstrating the most cyclicality are general merchandise, health and beauty, non-food and the least in dry groceries where there seems to be a trend towards high quality good.

4 The Cyclicality of Prices: Posted Prices, Effective Prices, Sales Prices and the Quality Choice Effect

As previous literature has already uncovered that effective prices tend to be more cyclical, we will start this section trying to verify this result using our data. Contrary to previous papers that focus only on the crisis years we include all years up to 2017. The second part of this section will focus on sales size cyclicality and sales frequency. Finally,the last part of the section will introduce the quality choice as one of the main drivers of this discrepancy between posted and effective prices and will be contrasted with the previous literature on store switching.

4.1 The Cyclicality of Prices: Posted Prices, Effective Prices

To assess the cyclical behavior of posted and effecive prices we run the following regression:

$$Y_{act} = \beta UR_{at} + \lambda_t + \theta_{ac} + error \tag{8}$$

Y_{act} refers either to posted or effective prices as defined in section 2. UR_{at} is the seasonally adjusted unemployment rate, θ_{ac} denotes the area fixed effect and λ_t is the time fixed effect. As in CGH we estimate the regression on a monthly frequency, as unemployment of metropolitan areas is only available monthly. Also we control for serial correlation of the error using Driscoll and Kraay (1998) standard errors.

Turning to the results on Table 3 we observe that for all weighting choices and additional checks effective prices are significantly more cyclical than posted prices. This difference becomes even bigger when we include the monthly fixed effects or a linear trend.

Although this estimation by no means captures the causal effect of unemployment to prices CGH provide some arguments why the endogeneity of the estimates might not be as bad. In detail, most products are not produced locally and additionally we control for time fixed effects. In any case, there is no reason to believe that the endogeneity issue is affecting posted and effective prices differentially.

4.2 The Cyclicality of Prices: Sales Frequency and Size

In the previous section we demonstrated that, similar to previous the literature, effective prices, i.e. prices taking into account changes in quantities sold, tend to be significantly more cyclical that posted prices. This is supported by most of the literature, however, it is unclear what is driving these differences. One of the main explanations considered are sales and sales frequency. In this section we are going to focus on sales and investigate their cyclicality.

To achieve this we estimate again the following regression:

$$Y_{act} = \beta UR_{at} + \lambda_t + \theta_{ac} + error \tag{9}$$

Y_{act} refers to the size, frequency and share of sales. UR_{at} is the seasonally adjusted unemployment rate, θ_{ac} denotes the area fixed effect and λ_t is the time fixed effect. As in CGH we estimate the regression on a monthly frequency, as unemployment of metropolitan areas is only available monthly. Also we control for serial correlation of the error using Driscoll and Kraay (1998) standard errors.

In Table 4 we report the results with respect to sale cyclicality. Initially, we check to see if the frequency of sales demonstrates cyclical characteristics. Here the results tend to

Table 4: Cyclical Properties of Sales

Area x cat. fixed eff.	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed eff.	No	No	Yes	trend	Yes	Yes	Yes	Yes
Weigted regression	No	No	No	No	No	No	Yes	Yes
					Aggregation using expenditure shares			
Depend. variable	Equal weight in UPCs aggregation				Market Specific	Common	Market Specific	Common
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sales:								
Frequency	.821***	.668***	.151**	.003	.082	.085	.001	.003
	(.095)	(.067)	(.081)	(.079)	(.121)	(.132)	(.118)	(.136)
Size	.167***	.147***	.311***	.287***	.245***	.274***	.234***	.312***
	(.023)	(.052)	(.043)	(.024)	(.026)	(.023)	(.019)	(.012)
Share	.835***	.533***	.005	.001	.019	.032	.012	.031
	(.023)	(.052)	(.033)	(.027)	(.023)	(.019)	(.028)	(.020)

Notes: The table reports the estimated coefficients on seasonally adjusted unemployment. Column 4 control for linear trends. Number of observations is 405402 . The last two columns report the regression results where observations are weighted by the average expenditure share of a given area/category cell in total spending across all cities and categories, effectively taking into account both the price level and quantity sold. Following CGH we report the Driscoll and Kraay (1998) standard errors in parentheses.
*** Significant at 1 percent level
** Significant at 5 percent level
* Significant at 10 percent level

be mixed. In detail most of the cyclicality appears in the first two columns where we do not control for monthly fixed effects. For columns 3-8, when controlling either for time fixed effects or including a linear trend, the coefficients become insignificant. This result is also supported by the literature. Furthermore, the cyclicality of the share of sales to non-sales is insignificantly different from zero for most specifications. Finally, the coefficient on price changes due to sales demonstrates some pro-cyclicality with the size of sales reducing when unemployment increases. This is a similar result to CGH but at odds with Kryvtsov and Vincent (2014). Hence, sales do seem to have a very strong cyclical component that will help us understand the cyclicality of effective prices.

4.3 The Cyclicality of Prices: Quality Choice and Store Switching

Up to this point we have demonstrated that effective prices are indeed more cyclical than posted prices, an estimate that would account for an overestimation of inflation by 0.3-0.4 p.p. during downturns and similar magnitude underestimation during expansion periods. This is an economically significant finding that points to effective cyclicality driven by household choices. The previous literature has tried to attribute this difference to store switching which leads to households paying a lower price after searching for a better price of the same product. However, this might be asymmetric as there is no reason to assume that households would

Table 5: Cyclical Properties of Store Switching and Quality Choice

	Base	Comparison	
Dependent variable	UR (1)	$UR \times \bar{R}_{ast}$ (2)	$UR \times \bar{Q}_{cst}$ (3)
Share of revenues	-0.101^{***}	-0.067	-0.265^{***}
	(0.011)	(0.079)	(0.061)

Notes: The table reports the estimated coefficients on seasonally adjusted unemployment. Column 4 controls for linear trends. Number of observations is 441743 . The last two columns report the regression results where observations are weighted by the average expenditure share of a given area/category cell in total spending across all cities and categories, taking into account both the price level and quantity sold. Following CGH we report the Driscoll and Kraay (1998) standard errors in parentheses.
*** Significant at 1 percent level
** Significant at 5 percent level
* Significant at 10 percent level

search again at period of expansion to find the same product with a higher price. The obvious counter argument is that they have to search in each period and they only do that in downturns, however this is only partially what we observe in the data. In this section we will try to demonstrate that the effective prices are mainly driven by households switching to lower quality goods and not per se store switching, which also explains why when households become unconstrained and switch back to higher quality goods we have the reverse bias.

To test for this we calculated a measure of stores relative price per market denoted as R_{acstj}. Which then is used to calculate the relative price of a store defined as:

$$\bar{R}_{ast} = \sum_{c} \sum_{j} \omega_{acstj} R_{acstj} \tag{10}$$

where ω is the weight. This measure captures a store price measure in a given area in a given month. Similarly we calculate Q_{cst} as the average quality ratio consumed for a given product category for a given area and a given month.

To test the relative importance of store or quality switching with respect to cyclicality we are using the following regression:

$$Y_{ast} = q_{sa} + \alpha_1 UR_{at} + \alpha_2 UR_{at} \times \bar{R}_{ast} + \alpha_3 \bar{R}_{cst} + \alpha_4 UR_{at} \times \bar{Q}_{cst} + \alpha_5 \bar{Q}_{act} + \lambda_t + error \tag{11}$$

where Y_{ast}, thus we effectively identify the effect of within store product quality consumed changes as well as store switching as in CGH.

Table 5 presents the results of this exercise. When we include the interaction term, with shares of product quality within store, the term of store switching which becomes insignificantly different from zero. This implies that effective prices are mainly driven by the within and across stores quality shifts and not lower price of the same product.

5 Conclusions

This paper examines how consumer choices, with respect to quality of goods consumed, could, if not accounted for properly, bias the inflation profile in a counter-cyclical manner. This constitutes a possible explanation for the missing inflation/deflation puzzle.

This study is the first to document that quality choices are correlated with the business cycle both in the downturn and during the recovery, utilizing the complete panel of Nielsen Scanner Data for the period 2006-2017. Machine learning was used to classify all products to quality categories. We classify products into low and high quality using Convoluted Neural Networks and supervised learning. We show that for whole time period effective prices are more cyclical than posted prices.

Furthermore, we compare our result with the previous literature and demonstrate that most of the difference is effectively explained by differences in the quality of goods consumed across the business cycle. In contrast to previous literature we show that store switching is confounded with quality switching, however quality seems be able to explain a significant part of the revenue shift even when accounting for store switching.

References

Agarwal, P. K., Har-Peled, S., and Varadarajan, K. R. (2004). Approximating Extent Measures of Points. *Journal of the ACM*, 51(4):606–635.

Bils, M. (2009). Do Higher Prices for New Goods Reflect Quality Growth or Inflation? *Quarterly Journal of Economics*, 124(2):637–675.

Boskin, M. J., Dullberger, E. R., and Gordon, R. J. (1996). Toward a More Accurate Measure of the Cost of Living: Final Report to the Senate Finance Commitee from the Advisory Commitee to study the Consumer Price Index. Technical report.

Burns, A. F. and Mitchell, W. C. (1946). *Measuring Business Cycles*, volume I. National Bureau of Economic Research, New York.

Coibion, O. and Gorodnichenko, Y. (2015). Is the Phillips Curve Alive and Well After All ? Inflation Expectations. *Americal Economic Journal: Macroeconomics*, 7:197–232.

Coibion, O., Gorodnichenko, Y., and Hong, G. H. (2015). The Cyclicality of Sales , Regular and Effective Prices. *The American Economic Review*, 105:993–1029.

Coibion, O., Gorodnichenko, Y., and Hong, G. H. (2019). The cyclicality of sales, regular and effective prices: Business cycle and policy implications: Reply†. *American Economic Review*, 109(1):314–324.

Eichenbaum, B. M., Jaimovich, N., and Rebelo, S. (2011). Reference Prices , Costs , and Nominal Rigidities. *American Economic Review*, 101(February 2011):234–262.

Gagnon, E., Lopez-Salido, D., and Sockin, J. (2017). The cyclicality of sales, regular and effective prices: Business cycle and policy implications: Reply†. *American Economic Review*, 107(1):3229–3242.

Guimaraes, B. and Sheedy, K. D. (2011). Sales and monetary policy. *American Economic Review*, 101(2):844–876.

Jaimovich, N., Rebelo, S., and Wong, A. (2015). Trading Down and the Business Cycle. *NBER Working Papers*, September(21539).

Kehoe, P. and Midrigan, V. (2015). Prices are sticky after all. *Journal of Monetary Economics*, 75:35–53.

Ng, S. (2017). Opportunities and Challenges: Lessons from Analyzing Terabytes of Scanner Data. *NBER Working Paper Series*.

Shapiro, M. D. and Wilcox, D. W. (1996). Mismeasurement in the Consumer Price Index : An Evaluation. *NBER macroeconomics annual 1996*, 11(1996):93–142.