Author(s): Ratula Chakraborty, Paul Dobson, Jonathan S. Seaton and Michael Waterson

Article Title: Pricing in inflationary times - the penny drops

Year of publication: 2011

Link to published article: http://www2.warwick.ac.uk/fac/soc/economics/research/workingpapers/2011/twerp_975.pdf

Publisher statement: None
Pricing in inflationary times- the penny drops*

Ratula Chakraborty\textsuperscript{a}, Paul Dobson\textsuperscript{a}, Jonathan S. Seaton\textsuperscript{b} and Michael Waterson\textsuperscript{c}

Draft Version- December 2011

Abstract

We investigate micro pricing behaviour in groceries (the UK’s most important consumer market) over eight years including the inflationary period of early 2008. We find behaviour sharply distinguished from most previous work, namely that overall basket prices rise but more individual prices fall than rise! This is consistent with retailers obscuring the fact of rising basket prices. We employ a significant new source of data that captures cross-competitor interplay in prices at a very detailed level. Unusually but importantly, our work takes into account that consumers buy baskets of goods, rather than individual products, when shopping at supermarkets.

JEL numbers: L16, L81, E31

Keywords: Pricing behaviour; supermarket prices; inflationary behaviour; price indexes

*We would like to thank seminar audiences at Toulouse, EARIE Stockholm and Warwick, particularly Michael McMahon, Carlos Noton and Howard Smith, for helpful comments and suggestions. We would also like to acknowledge ESRC funding for this project (Grant RES-062-23-1962)

a. Norwich Business School, University of East Anglia
b. School of Business and Economics, Loughborough University
c. Corresponding Author, Department of Economics, University of Warwick, Coventry, CV4 7AL, michael.waterson@warwick.ac.uk
1. Introduction

Movements in relative prices as between firms are of interest in revealing how demand effects impact on behaviour (Chevalier et al, 2003). Equally, even at micro level, price movements are of interest to macroeconomists (Chen et al, 2008). In setting monetary policy to meet an inflation target, a key question facing policymakers is how prices behave, in particular how they respond to shocks (Eichenbaum et al, 2011). Macroeconomists from at least Mankiw (1985) onwards have recognized that price setting at the micro level can have a significant influence on the performance of policies to target inflation, due for example to sluggishness in price movements. Hence the importance for monetary policy and the substantial literature examining the behaviour of consumer prices in micro detail following Bils and Klenow’s (2004) seminal paper (e.g. Abe and Tonogi, 2010; Bunn and Ellis, 2009; Ellis, 2009; Midrigan, 2011; Nakamura, 2008, Nakamura and Steinsson, 2008; Chen et al., 2008; Kehoe and Midrigan, 2007; Berka, Devereux and Rudolf, 2011).

We investigate micro pricing behaviour over a period of time that includes a significant recent inflationary period in the most important consumer market in the UK, but document findings sharply distinguished from most previous work. In doing this, we refocus attention on a key factor that appears to have been almost entirely ignored in the analysis of supermarket pricing to date, namely that consumers buy baskets of goods, rather than individual products, when shopping at supermarkets. Specifically, we find that overall basket price rises are disguised in a very subtle manner, whereby individual prices appear to fall yet (most) consumers actually pay a higher basket price. In other words, the episode we investigate reveals and subsequently resolves a puzzle, most keenly seen in the inflationary period of early 2008, where overall basket prices rise but more individual prices fall than rise! We argue that this behaviour is consistent with

---

1 Midrigan (2011) does take full account of the fact that firms are multiproduct, but does not relate this specifically to the consumer side of the equation. An earlier empirical paper, Lach and Tsiddon (1996) also considers price setting by multiproduct forms, but does not link this directly to consumer purchase baskets.
retailers aiming to obscure the fact that basket prices are rising. Our analysis involves a significant new source of data that captures price behaviour at a very detailed level, including crucially interplay across market competitors over period of several years.

Modern consumers face a formidable problem (admittedly, partly of their own making) even when purchasing everyday products. In a large supermarket expedition, they are confronted with tens of thousands of lines. Hence in searching for grocery products they may rationally be inattentive (Reis, 2006; Mackowiak and Wiederholt, 2009) to much of the detail on prices. As Chen et al. (2008) show, rational inattention is one explanation for asymmetric price adjustment, although we would argue that it is difficult to be precise concerning the prediction regarding pricing based upon consumers’ rational inattention.

At the same time, it must be recognized that prices are influenced by both sides of the market, consumers and sellers. Sellers potentially face menu costs (Mankiw, 1985) in resetting prices, although these costs have undoubtedly been reduced over the last couple of decades by improvements in information technology, connecting cash tills to central computers, to ordering and to shelf pricing mechanisms, etc. Sellers potentially also expect fierce competition on pricing from other firms, so are influenced by things other than costs. Whether consumers are rationally or irrationally inattentive, we may expect sellers at least to take their behaviour into consideration in their price-setting actions. Our argument is that at times of inflation, sellers can take advantage of this inattention by flexing multiple prices, so reducing resistance to price rises.

Given this, it is surprising that there has been comparatively little examination of the influence of interplay between sellers in determining price patterns over time, outside the area of energy prices (where there is a significant literature on “rockets and feathers” pricing mechanisms- Bacon, 1991; Bils and Klenow 2004; Borenstein et al., 1997). This is particularly important since it is well-known (Anderson et al., 2001) that in

---

2 See Frey and Manera (2007) and Meyer and von Crammon-Taudabel (2004) for useful surveys and Geweke (2004) for a more sceptical view. There is also a significant literature on agricultural commodities, which these surveys cover. However, this agricultural literature tends not to have any focus on interplay between sellers.
oligopolistic market situations, “overshifting” of marginal cost increases can occur. Moreover, in food and associated household good purchases, the consumer is buying a multi-product bundle, so there are additional considerations (which turn out here to be of crucial importance) of behaviour within the menu of goods on offer.

In this paper, we make use of a very significant source of data on prices of precisely defined lines of goods across major supermarket chains in the UK. It is important to recognise the significance of supermarket purchases to UK consumers, the professionalism of the companies involved, and the particular characteristics of the market that render it so useful in analysing the issue of interplay in firms’ pricing behaviour. This enhances our ability to explain pricing behaviour as compared with some other studies.

The four largest supermarket chains whose data we have account for around three-quarters of all supermarket sales in Britain and the majority of all food and non-alcoholic drink purchases for the grocery sector as a whole (Competition Commission 2008). It has been estimated that the largest, the Tesco chain which has roughly 30% market share, accounts for 12% of British consumers’ current expenditure. Tesco is the world’s third or fourth largest retailer by revenues\(^3\); its rival Asda with a 17% share of the British market, is a Walmart subsidiary. Competition on pricing between them is significant, but in its most recent and comprehensive investigation of the market, the UK Competition Commission (CC) took the view that “there would obviously be cause for concern if one retailer were able to achieve and exploit market power. We are not convinced Tesco is in that position.” (CC, 2008, p5)\(^4\)

Our data have several major advantages compared with much of the data used in the existing academic research. First, we have a relatively long time period of detailed weekly price data, with up to eight years of weekly data available for our analyses. Second, as Nakamura (2008) points out, a good many of the existing studies are based on

\(^3\) Deloitte (2010). This depends on whether Metro, which is slightly bigger, is considered a retailer.
\(^4\) Michael Waterson is a Member of the Competition Commission but was not involved in the production of this report. He has seen only versions of the report that are in the public domain.
data from one supermarket chain, the Dominick’s Finer Foods database, whereas her data, and ours, cover several store chains at a detailed level. In our case the major national chains are represented, enabling cross-chain comparisons of price movements.

Our data should be contrasted with retail scanner data and consumer panel data which are common alternative sources available on a commercial basis (e.g. from companies such as Nielsen and Kantar). Retail scanner and consumer panel data have many strengths in recording retail sales and consumer purchases respectively, but matching products across time and across supermarkets can be problematic, and the prices reported usually reflect average unit prices rather than individual item prices. The resultant averaging effect can overstate the amount of week-to-week price variability for individual items (e.g. as found by Ellis, 2009) compared with the focus here on ticket prices of individual items.\(^5\)

One key data resource we have is the body of pricing data that has been published by Tesco on its website (http://www.tesco.com/todayattesco/pricecheck.shtml) listing prices in each of the “Big 4” supermarket chains (Asda, Morrisons, Sainsbury’s and Tesco), downloaded and collated by us every week from November 2003 to December 2008.\(^6\) The Tesco website explains the methodology used in drawing the price comparisons, but a key feature (and major research advantage for us) of the Big-4 pricing strategy is that, within their large supermarket outlets, uniform national pricing is practised across Great Britain (CC 2008). Thus the prices really do represent those any consumer will face across the country, rendering them equivalent to scanner prices.\(^7\) The level of detail in our database is impressive - up to 16,000 individual lines for each chain, gathered weekly, amounting to over 5 million observations in total. It is notable that although

---

\(^5\) A further popular data source is BLS monthly data, also based on a consumer panel but which is more aggregate than commercial sources such as Nielsen's Homescan.

\(^6\) The impact of a competitor price discovery site (mysupermarket.co.uk) led to changes in price reporting after about this point; on this see below.

\(^7\) The national pricing feature contrasts strikingly with local store-level pricing observed in the USA (Hosken and Reiffen, 2007; Ellickson and Misra, 2008), and “price flexing” formerly practised by Tesco and Sainsbury (CC 2008; Smith, 2004) but abandoned to join Asda and Morrison’s national pricing by 2003 (CC 2003). Note that national pricing relates only to their larger branches though. Both Tesco and Sainsbury operate under more than one format and their smaller in-town stores (Tesco Metro, for example) have prices which may diverge from this. Asda and Morrison have no such smaller stores. Most sales in Tesco and Sainsbury take place in their larger stores.
undertaken by one of the key players (though reported as “independently collected”), none of the others has challenged its impartiality.\(^8\) Moreover it covers all main grocery categories, with a full range of brands and own label products, although focussing on processed products rather than the more basic raw fruit, vegetables, meat and fish. This level of detail means that we can control, amongst other things, for product type, brand origin, package size, and number of close substitutes.

Supplementary to this, we engage in confirmatory work on a further database of the same character that brings our sample up to the end of 2010. This is a dataset downloaded over time from the site of mysupermarket.co.uk. Essentially, the structure and scope of the data in this second series is the same as that in the first, except that it does not include Morrisons, the smallest chain. The advantage is twofold, first that we have data for both sites for most of 2008, so can engage in cross-checking; second we can bring the series into a period where cost pressures and inflation moderated. The disadvantages are that creating a matched index covering the whole of the eight year period reduces the sample size and that Morrison, the smallest of the four chains, is no longer represented. Details on both data sources and our selection from them are covered in our Data Appendix.

Our data sources may seem to pale to insignificance against that of Abe and Tonogi (2010) for example, who have three billion observations. However, given that our prices are featured, in each case, across several hundred stores, and given that the firms in question normally change prices no more than weekly (hence the frequency of the data collection exercise), whilst the modal consumer shops at their particular choice amongst such stores weekly (CC, 2008),\(^9\) then the equivalence is closer than immediately apparent. Our data most directly relate to those used by Ellis (2009), in being weekly and covering major UK supermarket chains, but we have many more products and, most importantly from the viewpoint of analysing inflationary episodes, his data stop short of (almost all) the inflationary period of 2008 that is one focus of the current analysis.

\(^8\) Where challenges have been made, e.g. cases examined by the Advertising Standards Agency, these have related to marketing campaigns used by different players (e.g. Tesco and Asda) featuring selective price comparisons that are argued to be misrepresentative.

\(^9\) See paragraphs 3.48 to 3.50 of the report for a discussion on this point. The precise proportion that does a weekly shop appears not to be something that can be precisely measured, but may be as high as 70%.
The time period in 2008 where commodity price inflation was very rapid is of particular interest. Supermarket chains clearly face a problem of coping with this input price inflation in the context of a significant degree of inter-chain rivalry. In turn, because of their significance to the economy as a whole, their behaviour will impact significantly on the movement of retail price indices and so the transmission of inflationary processes nationally. Observing their detailed behaviour allows us to uncover inflation transmission mechanisms at a very micro level.

To preview results, we have two main findings. First, mean basket prices behave as expected given inflationary trends. Overlaid on that, we find a pattern of individual item price change markedly different from the regularity exhibited by Dominick’s prices in the Chen et al. (2008) study, and indeed probably inconsistent with Nakamura and Steinsson’s third of Five Facts about Prices (2008).\textsuperscript{10} In our data, most obviously during the 2008 inflationary episode, a very large number of supermarket shelf prices go down; they are anything but sticky! But at the same time, a much smaller set of prices rises. However, those prices that go down typically do so by a very small amount, commonly a penny (the smallest monetary unit), whereas those that rise do so on average by a greater amount. These results are almost the opposite of Chen et al.’s findings. In sum in our sample, a rise in a typical basket price is disguised by a fall in large numbers of individual item prices, thereby raising the variance in basket prices significantly.

The plan of our paper is as follows. In section 2 we describe the puzzle we later examine and the extent to which our data samples are representative of food pricing patterns as a whole. Then in section 3, we focus on patterns in the microstructure of price changes, relating this to the existing literature. Section 4 moves to examining the impact on baskets of products, given that these, rather than individual items, are what consumers actually purchase. It also examines subsequent experience. Section 5 makes some broader concluding remarks.

\textsuperscript{10} They do express some caution about this finding. See also their follow-up note (2010).
2. The pricing puzzle

There is no doubt that 2008 was a traumatic year for the world’s financial markets. This has overshadowed events in the early part of the year which was characterised by substantial and widespread rises in commodity prices, subsequently checked and reversed by the banking crisis. Figure 1 shows how after a long period of gentle inflation, world commodity prices took off rapidly in the first six months of 2008. Clearly, although supermarkets including those in the UK have substantial buying power, this will have impacted on UK supermarkets, as well as smaller scale retailers, and forced them to consider real price increases, because even Tesco is a small purchaser on a world scale. Our focus is on the mechanisms by which these price increases were implemented.

The analytical approach we take is to carry out our work in the main on a sample of 600 products from the Tesco website (the main, T600, sample) for which we have consistent and virtually continuous data for the 5 year period ending in late 2008 in order to gain a perspective on pricing movements over time.\[11\] Clearly this is a sample restricted amongst other things by its end date and for this reason we check elements of our analysis against an extended sample of 370 products consolidated across both the Tesco website data and data from mysupermarket.co.uk (the M370 sample) collected for the shorter period from late 2007 to 2010, hence covering the period after the inflationary episode as well as before it. In total then, this second sample of products covers an eight year period from 2003 to 2010 (albeit with a more limited set of products). This enables us to examine whether, once the main inflationary pressure ceased, the supermarkets returned to previous practice in price setting or not.\[12\]

---

\[11\] The collection of data over the Christmas/ New Year period is limited and erratic, for all chains and products. Our product sample is chosen to maximise completeness of observation given that. In facing a missing observations problem, our sample is no different from other alternatives which commonly suffer from a degree of missing observations- see e.g. Klenow and Kryvtsov, 2008, and Nakamura and Steinsson, 2008.

\[12\] Besides the problem of the Christmas/ New Year period, the data suffer from several difficulties that limit the sample given that we want to make it as compatible and complete as possible. First, coverage of products particularly from Morrison is limited even in the Tesco sample, hence constraining the whole sample. Second, gaps appear from time to time, either because the operatives collecting prices were unable to find the product or because it was sold out. Third, the product mix sold in supermarkets does naturally change over time. Fourth, because the collection process is manual, occasional mistakes are made, such as
Figures 2 and 3 below starkly illustrate the puzzle in inflation transmission mechanisms in the UK. In the first half of 2008, retail food prices in the UK (like many other countries) moved from gentle to rapid inflation, which was subsequently sharply reversed in association with the significant financial turmoil of the closing quarter of 2008. Figure 2 shows a rapid increase in UK food prices in the first half of 2008 over the previous year, whether examined broadly or focussing on processed food, as measured by the official CPI index. It also shows that this pricing event was substantially different in magnitude from any experience in the preceding five years or subsequently. At the same time, Figure 3 relates to our sample of 600 products. These data come from retailers covering over 2/3 of food expenditure (defined broadly; source TNS), and shows there to be consistently, throughout this inflationary episode in 2008, a record of very significant numbers more of prices falling rather than rising.

Clearly, although most food buying in Great Britain goes through the firms on whom we record pricing data, the first thing we need to establish is whether our sample is broadly representative of the set of products captured by general indexes such as the CPI— it could be the case that there is something odd about our particular sample, in which case the puzzle would disappear. Being able to reproduce the pattern exhibited in figure 2 provides a very good test of this. Once we have shown that the sample is not unusual, we can investigate what is happening in somewhat more detail.

---

prices which appear to be for a different quantity of the product, say. Our approach has been to drop products that do not appear initially, that leave the series before the end of 2008, or that have excessive gaps in the reporting period. To combine these various constraints, a Herfindahl-type approach was adopted. More details are available in the Data Appendix.
Figure 1: Trends in commodity prices—quarterly moving average

Figure 2: Retail price inflation in the UK, as measured by CPI indices

Source: ONS (DK9O and DK9P)
Figure 3: Quarterly moving average price falls, T600 sample

**Representativeness of our sample**

A difficulty arises because of the enormous range of products available within a typical store, of which we have a relatively small sample. Given this, there is a significant ante chance that the products we have chosen are in some way idiosyncratic. The main step in establishing the contrary is to investigate price movements within our sample by reference to the underlying position in official statistics, namely the CPI. To put it another way, the puzzle should not arise because of sample selection.

Figures 4 and 5 and table 1 examine this in various ways. Figure 4 simply takes our 600 product sample and examines simple mean price changes when compared with two alternative definitions of relevant components of the CPI, which we have labelled as CPI1, which is composed of Food, Alcohol and Tobacco (data code DK9O) and CPI2.

---

13 Abe and Tonogi (2010) take a different tack, showing that their data implies the official Japanese CPI is not accurate. We have confidence in the UK CPI, and therefore want to check whether our sample is consistent with it.
Processed food and non-alcoholic drinks (code DK9P). Clearly, whilst the supermarkets (taking this sample) appear not to exhibit quite as extensive price rises as do the more general CPI series, the underlying patterns are very similar. Figure 5 looks only at the simple mean across supermarkets and its link to CPI1, for clarity.

Table 1 focuses on the correlation between a constructed quasi- CPI index using our samples and the official CPI series. To generate our indexes, for each supermarket we employed the weights used in creating the official CPI and applied them to the geometric mean of the price observations we had within that category, subsequently aggregating up the relevant categories to the level of the CPI index.14 Given the relative sparseness beyond processed products, and the high weights for alcoholic drink, the table uses data for categories relating to the CPI for processed food and non-alcoholic drink only. Again, this confirms that our main T600 sample of prices and our supplementary narrower M370 sample over eight years both exhibit a very similar pattern of price movements, of broadly similar character and extent to that shown by the CPI. This is of some importance, because it gives credence to our sample. As Table 1 shows, the correlations between supermarket values and CPI values in each case are very high. To put it another way, someone buying the CPI basket but confined to our sample of products would face essentially the same inflationary pressures as if they were buying from all CPI basket products across the entire grocery market.

So, the puzzle remains. Prices definitely rose rapidly in 2008, particularly the earlier part, both in the CPI and our data samples. How did the firms involved manage to raise prices to consumers to such an extent whilst maintaining their competitive position? And how did they do it whilst actually engaging in a welter of price cuts?

14 This conforms closely to the approach used in the UK to calculate the official CPI.
Figure 4: Price movements in our data samples by comparison with more general CPI indices.
Figure 5: The relationship between the official CPI and our “constructed” CPI

Table 1: Correlations between the prices in supermarkets and the relevant CPI measures

<table>
<thead>
<tr>
<th></th>
<th>CPI 1</th>
<th>CPI 2</th>
<th>Tesco</th>
<th>Sainsbury</th>
<th>Morrison</th>
<th>Asda</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPI 2</td>
<td>0.998</td>
<td>0.947</td>
<td>0.937</td>
<td>0.960</td>
<td>0.936</td>
<td></td>
</tr>
<tr>
<td>2005-2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPI 2</td>
<td>0.947</td>
<td>0.947</td>
<td>0.947</td>
<td>0.957</td>
<td>0.934</td>
<td></td>
</tr>
</tbody>
</table>

Note: The correlations in the T600 product sample are for January 2005 onwards. The prior period is strongly influenced by the significant shift in what are here called Morrison prices, but in fact in 2003 start by being Safeway prices.
3. Examining the microstructure of price changes

The pricing puzzle we have identified initially deepens when we examine price changes in more detail. First, the absolute number of price falls is strongly positively correlated with the positive rate of inflation, so that price falls intensify in 2008, compared with earlier periods. This is particularly true of penny price falls- the relationship is illustrated in Figure 6 for penny price falls using the T600 sample. In line with this, Figure 7 shows that even in 2008, the cumulative distribution of price fall values is above that of price rises. In both these cases, we focus on actual money values, not percentages, given the low underlying value (under £1) of many of the goods, meaning that examining percentage changes would obscure the integer constraints involved.

So more basket component prices are falling than rising, but at the same time, we have already seen that on average, basket prices have definitely risen. How is this resolved? The answer is that although price falls outweigh price rises in absolute number terms, and also outweigh them for all low values of pence, price rises outweigh falls for almost all intermediate values, namely from around 10p onwards. Note that this is precisely the opposite of Chen et al’s (2008) finding for Dominicks’ prices, but in our case we have a much more comprehensive sample and more firms, as well of course as a different country and time period. What is remarkable is the degree of concordance between the four supermarket chains on this point, together with the extent of price movements.

Two figures, figures 8 and 9, illustrate the key features of these individual price movements again using the T600 sample. Figure 8 shows an aggregate picture across all four firms. Noting the log scale in the vertical axis, small pence price cuts dominate numerically, although there are also many small pence price rises. Round number price falls and rises are popular (10p, 20p, 30p, 100p, for example).

Figure 9 focuses on the micro movements up to 20p, looking at each chain separately. This figure is constructed as follows. The price movements are scaled such that, if for any value of price movement each chain were making only price cuts, a value of +1 would be
reached. If at that value of price movement each chain was only raising prices, not lowering them at all, then a value of -1 would be generated. Thus this figure shows both preponderance in direction and concordance amongst firms. As we see, up to 6p, the price movements are largely or predominantly falls, whereas as we move nearer to 20p, price movements are much more likely to be rises, although dominance here is not as strong as at lower levels. The other thing of significance is the extent of concordance in approach as between firms, although Tesco and Asda both engage much more extensively in penny price cuts than do Sainsbury and Morrison’s. We now turn to considering the impact on basket prices.

![Graph showing CPI and penny price changes](image)

Source: T600 data

Figure 6: Inflation as represented by the CPI index plotted alongside penny price changes.
Figure 7: The cumulative distribution of price changes by size, T600 sample.

Figure 8: Price rises and falls, 2008, summed across supermarket chains (T600).
4. Analysis of obfuscation in pricing changes

Our analysis so far suggests that the supermarket chains, particularly Tesco and Asda, were engaging in obfuscation of their true position that basket prices were rising, through the snowstorm of penny and other small value individual price falls that we have observed. Consistent with this view, Nielsen media estimate that advertising spend in 2008 was up on 2007 by 23.2% for Tesco, 47.4% for Asda, 11.4% for Sainsbury and 6.2% for Morrison. Advertising over this period focussed strongly on price promotions, such as how many prices had fallen and number of price falls relative to other players (expenditure on “pricecheck”- type promotions peaked in March- May 2008). See fig 10 for the overall picture.
There are several papers which develop theoretical models of obfuscation potentially relevant to our work. These include Gabaix and Laibson (2006), Ellison and Wolitzky (2008), and Carlin (2008); see also Piccione and Spiegler (2010). Carlin assumes that the relevant statistic is the average obfuscation level, so an increase in obfuscation by one firm has general effects. Ellison and Wolitzky assume that search costs are convex and have a component that is determined by the firm(s) engaging in obfuscation. In their basically Stahl-type model (Stahl, 1989), they show (essentially) that firms engage in obfuscation in equilibrium. Gabaix and Laibson call a similar phenomenon “shrouding”.

Much of the work in each paper is simply dedicated to showing the existence, in equilibrium, of obfuscation. However, both Ellison and Wolitzky and Carlin include potentially testable comparative static results. In the former, highest cost firms use most obfuscation, but the relationship between price level and obfuscation is non-monotonic. Carlin’s paper also has a very interesting prediction that “increased competition makes it more likely that firms make their price disclosures opaque”.

Figure 10: Recorded advertising spend by major players, 2007-2010 (source NMR)
Relating this to our findings, Tesco is known as a low, rather than high, cost firm (CC, 2008, p158.)\(^{15}\) However, it obfuscates more than say Sainsbury, which runs counter to the Ellison-Wolitzky prediction. Nevertheless, it is feasible to argue that the inflationary period intensified competition between the players, particularly Asda and Tesco, so leading to greater desire to make price disclosures opaque. Hence it is reasonable to say that Carlin’s prediction is consistent with the evidence.

**Basket prices**

Returning to the consumer choice problem, a consumer faces a significant difficulty, not so far acknowledged in any of the papers on the microstructure of supermarket pricing. In these stores, consumers are rarely shopping for only one item. Rather, consumers are purchasing a basket of goods and it is costly to split that basket (for that week) between several supermarkets, because of the fixed costs involved in travelling to them and negotiating their aisles.\(^{16}\) Suppose the consumer typically buys that basket at supermarket A. The basket involves some products that are bought every week, some that are bought only occasionally, etc. The consumer’s problem is not so much one of search as one of recall. Initially, they visited a particular supermarket because that supermarket provided them with the best deal, in some sense. Week by week, the overall bill can be thought of as the sum of a large number of separate prices and hence as a random variable, dependent partly on the particular basket purchased. If the consumer’s bill is higher than expected, the question is whether it is suspiciously high, possibly even high enough to encourage a visit to an alternative supermarket for next week’s shop.

Let us develop the idea that the basket price is a random variable, dependent both on individual price values and on the components of the basket. Note that if the variance of the basket price is large across baskets, then if individual consumers vary their basket over time, or if they talk to fellow consumers with differently constructed baskets, they may find it difficult to determine whether their basket is actually getting more expensive.

\(^{15}\) Para 9.11 reads in part “our analysis of supplier prices …showed that suppliers receive the lowest prices, on average, from Tesco …”

\(^{16}\) UK consumers usually engage in a weekly or fortnightly major shop at one of the stores we are examining, then do “top-up” shops in the intervening period, commonly at different, smaller, stores (CC, 2008; Smith, 2004)
or cheaper over time. By contrast, if a constant proportionate increase across all prices is imposed (somehow), the variance is minimal and the impact of the imposed price rise will be clear.

More formally, the higher the variance the less likely the hypothesis that the mean basket price has risen will be accepted. Consider a market in which goods are priced individually, but bought in baskets, consisting of an assortment of many goods (say around 50 different items, typically). Each consumer purchases their own basket, taking prices as given. Sometimes it will transpire that this basket costs a bit more, sometimes a bit less. The consumer is fixed in their range of goods purchased (their “shopping list”), but uncertain as to whether to visit their usual store, or a replacement. The choice will be based on price for the basket - that is the only logical thing the consumer cares about is the overall basket price.

Firms are assumed to choose prices for individual products each from a distribution of possible prices, fixed for one period. The form of this distribution is irrelevant; it is likely to be significantly non-normal, though. Let us assume, for the moment, that each mean price, \( \mu \), is identical and each variance, \( \sigma^2 \), is the same.

This makes the consumer’s basket price the sum of a large number of independent random components, each drawn from an arbitrary distribution. Define the price of a basket, size \( n \), as \( B_n \). Then, by the Central Limit Theorem,

\[
Z_n = \frac{B_n - n\mu}{\sigma \sqrt{n}} \rightarrow N(0,1)
\]

Assuming Lindeberg’s condition holds (Billingsley, 1995, p357), there is no necessity for the means (\( \mu \)) or variances (\( \sigma^2 \)) of the individual good price distributions to be identical in order for this theorem to apply or, of course, for the individual price distributions to be normal.

\[17\] Of course, for only a subset of purchases a consumer makes, have they determined their list in detail. For others, they may make choices in store. This does not affect the analysis so long as the set of products for which quantities have already been chosen is still “large”.

A key implication of the formula is that the larger is $\sigma$, more generally (Lindeberg) the larger is the average variance from which prices are drawn, the more difficult it is for a consumer to draw the implication that, across two shopping experiences, one experience has given rise to a definitely higher priced basket than another (rather than, as a null hypothesis, being unchanged in expected price).

As a firm, you may be concerned that consumers perceive prices as rising. A means of obscuring this is to increase the variance in the set of price changes that you make. One means of increasing variance is to lower some prices at the same time as increasing others. Hence as a firm, at times when you want to obscure price rises, it is logical to increase the variance in prices.

To examine this key hypothesis we select a basket size from our sample, for concreteness we select baskets of size 30 and size 50, then generate (with replacement) a random sample of 50 such baskets and track the mean and the variance of this distribution, for each week of our data. The mean price of the basket will naturally vary over time as individual prices vary. Our interest is specifically in movements of the coefficient of variation (CV) in the basket price over time. On the null hypothesis that price changes are all of the same percentage, the correlation between mean basket price and CV in this should be zero.\(^{18}\)

In fact, this is not the case. The correlation between mean and CV in basket price, for baskets of size 30 and size 50, is given in table 2 below. For Tesco, Sainsbury and Asda, these correlations are relatively high. In fact, on a normal F test for significance, in all three cases we can reject the null hypothesis of no association even at the 0.1 percent

\(^{18}\) If all mean values were to rise by a constant amount, the mean would rise but the variance would remain constant, so the coefficient of variation (CV) would fall and correlation with mean price would be negative. If the mean values all increase by a constant proportion the CV, being dimensionless, would remain unchanged, so the correlation would be zero. To the extent that there is variation in price rises across products, the CV will rise. We therefore analyse whether a rise in CV is positively correlated with the rise in the mean and whether this correlation is statistically significant.
level of significance. However, in the case of Morrison, the association, although positive, is insignificant except at the 10% level.

Our conclusion is that in the case of baskets from Tesco, Sainsbury and Asda, at times when basket prices are high, it is made more difficult for the consumer to analyse whether the basket is more expensive. This is certainly evidence consistent with obfuscation, although does not demonstrate causation. Further evidence consistent with the hypothesis is that the maximum values for CV in each firm occur in May 2008 (For Tesco, April 2008 in the 30 basket case), at the height of the inflationary period.

<table>
<thead>
<tr>
<th>Basket size</th>
<th>Tesco</th>
<th>Sainsbury</th>
<th>Morrison</th>
<th>Asda</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0.372</td>
<td>0.365</td>
<td>0.113</td>
<td>0.356</td>
</tr>
<tr>
<td>50</td>
<td>0.355</td>
<td>0.327</td>
<td>0.101</td>
<td>0.346</td>
</tr>
</tbody>
</table>

**Subsequent experience**

An alternative possibility, which we should not discount before examination, is that the large increase in price falls, particularly penny price falls, is not associated with the inflationary period in particular but is a secular trend to increased price movements in supermarket competition. Our T600 sample of products is not well suited to examining this possibility because we have few observations in the immediately following period. Hence we turn to our extended M370 sample covering the period 2003 to 2010 inclusive.

This possibility that price changes are subject to secular trend rather than tied with the inflationary period is examined in relation to price falls in Figure 11. We see a picture where for Tesco and Asda the early period in 2008, at the height of the inflationary incident, still stands out as a clear high point for price falls. There is some resumption to

---

19 The F statistic at the 0.1 percent level of significance for (1, 261) degrees of freedom is approximately 11.15. The F values for Tesco, Sainsbury and Asda for the basket size 50 are (37.75, 31.29 and 35.45) respectively.
this form of competition in the mid months of 2009, which is however a period where food prices (and incomes per head) were falling, not rising (see figures 1 and 2). Thus, unlike the period under study, this later slight rise in price falls was not against the prevailing trend in price movements overall. In order to investigate whether there are broader differences between 2008 and other years within the M370 sample, we now move to a regression analysis of factors influencing price changes within this sample.

Figure 11: Number of price falls centred to quarterly MA on weekly data (M370 sample)

**Summarising individual price change behaviour by supermarkets**

There are several possible hypotheses regarding which products the firms decide to change in what is clearly a process where there is substantial discretion. First, the consumer may recall less accurately the price of an infrequently purchased item. If so, price rises might be loaded onto such items. However, supermarkets then face a difficulty- for any given price rise (say 10p), imposing it on an infrequently purchased good yields less revenue than imposing it on a frequently purchased good. A second possibility is that the consumer notices price increases less on high value goods than on
low value goods, or on own brand rather than branded goods. More generally, there will be greater upward pressure on prices when there are cost increases than when this is not the case. The difficulty here is that most processed products are such a mixture of ingredients that it is difficult to pinpoint particular products as more affected at any one time. Moreover, all products are influenced by oil prices, because they are transported to the store. Nevertheless, when there are general increases in costs and prices, upward pricing pressure on individual goods is likely to be more apparent. Finally, we already know that individual stores exhibit somewhat different pricing practices.

It is useful to summarise the outcome of these hypotheses regarding pricing behaviour by suppliers through the medium of a descriptive regression framework which we employ on the 370 sample to capture longer term trends. Consider the following probit panel data regression framework, capturing the various hypotheses regarding the probability of a price change. The dependent variable is a dummy $\rho$ taking the value 1 when there is a price change, zero otherwise. The regression equation is then:

$$\rho_{i,j,t} = f(BROL_i, pc_i, pf_i, st_j, \Delta CPI_t)$$

Here subscript $i$ refers to the good (dimension 370), $j$ to the store (dimension 3) and $t$ to time (dimension 365 weeks after differencing). Explanation is coming from a series of product fixed effects, namely for example whether the product is branded or own-label (BROL), the broad price category of the product ($pc$) and purchase frequency ($pf$), a store fixed effect ($st$) and a measure of the (smoothed weekly) change in the overall CPI.

Although the right hand side variables are exogenous, we view this as a descriptive regression framework because it simply summarises our informal hypotheses rather than

---

20 One factor that is clearly relevant is the elasticity of demand for the individual product. Such elasticities are unavailable, but the first two factors here could be thought of as proxies for them.

21 It is true that some products are more straightforward than others; milk would be an example. But the cost of milk is the subject of a bargain between the supermarket and their supplier (Smith and Thanassoulis, 2011).

22 In order that this variable remains exogenous, it is important that the price categories are broad enough so that the penny price fall does not move the product between categories. Hence we use the median as a divider. Similarly, a product where there are many promotions (price falls) over time is perhaps more likely to be frequently purchased, so again we use a broad classification of purchase frequency. Using the same reasoning, we use the overall CPI change as something exogenous to any individual supermarket.
having a clear theoretical *causal* justification. However, given what we have found already, the expectation is that 2008 will stand out as somewhat different. So for example, there is a natural arithmetic tendency for a price rise to be more likely when the CPI increases (although a change in the CPI does not cause a price change by a supermarket). But does this tendency hold for 2008? This and other associated hypotheses are examined in a series of probit regressions reported in table 3. In each case we include a full set of interactive variables relating specifically to 2008. To avoid a plethora of asterisks, we designate specifically only variables that are *not* significant at 1% or better.

The first regression reports results based on penny falls. As we see, Asda is slightly and Sainsbury markedly less likely to engage in such falls than Tesco. In general, penny price falls are more likely on branded goods but given that, less likely on goods high-priced goods and those that are frequently purchased.\(^{23}\) They are also, unsurprisingly, less likely when CPI increases. However, there are big changes in 2008. They are even more concentrated on branded goods, they are slightly more likely on frequently purchased goods and they are so much more likely when the CPI rises that the previous effect is outweighed making price *falls* absolutely more likely when the CPI rises. This is confirmatory of our previous results, including the view that price falls in high inflationary times are used as a distraction, drawing attention away from the overall rising basket price.

These results are largely repeated when we move in column 2 to looking at all price falls, then in column 3 to all price changes and in column 4 to all price rises, in the following respects. Branded goods are always more likely to experience the change, whilst frequently purchased goods are less (or in one case, no more) likely to show changes. But there are some interesting impacts in 2008. B goods are even more likely to experience a price fall in 2008. Price changes are even less likely to occur for frequently purchased goods overall, although this is not true for penny price falls. Price rises on frequently purchased goods are absolutely less likely in 2008.

\(^{23}\) There is a positive correlation between the good being branded and it being high-priced.
So far as the CPI is concerned, in general a positive change in the CPI is less likely to “lead to” a price fall, but more likely related with a good’s price rise; this is as expected. However, the effect of changes in the CPI are moderated in both directions in 2008, so that price falls are somewhat more likely and price rises somewhat less likely than in other years. One way to interpret this is that inflationary movements lead to more variance in prices, but this effect is significantly exacerbated in 2008 (with fewer but larger price rises and more price falls).

Finally, in column 5 we move to looking at the relative impacts on price falls. That is, given that there is a price movement, what makes it more likely to be a fall than a rise? Here, the sample is around 10% of the sample size for the previous regressions, because the large numbers of “no change” options are dropped. A movement is more likely to be a price fall if it is a branded good and if it is not frequently purchased. Also, as expected, it is more likely to be a price fall if the CPI is falling rather than rising. But in 2008 it is even more likely on a branded good or a frequently purchased good, also more likely if the CPI is rising. The CPI results imply that predicting whether a good experiences a price fall or a price rise is less easy in 2008- this is an expression of greater randomness. It should also be said that, except for this final regression, the nature of the sample with many zeros and few ones in the dependent means prediction of the ones is poor. In column 5, nearly 90% of the ones (price fall) are predicted, so it is good to notice that this regression has a very similar coefficient structure to the previous cases.24

Overall, these results strongly confirm our impression based on the T600 sample that, in 2008, obfuscatory strategies were being pursued in respect of price changes within the typical consumer basket.

---

24 We also estimated a multinomial logit, with no change, fall and rise as the three cases. The results are essentially confirmatory of those reported.
Table 3: Binomial probit regression analysis on the M370 sample

<table>
<thead>
<tr>
<th>Probit Dependent (=1 if)</th>
<th>1p price fall</th>
<th>Any price fall</th>
<th>Any price change</th>
<th>Any price rise</th>
<th>Price fall (vs. rise)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Regressors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0865</td>
<td>-0.1826</td>
<td>-0.2405</td>
<td>-0.1535</td>
<td>0.1513</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.00109</td>
<td>0.00144</td>
<td>0.00168</td>
<td>0.00123</td>
<td>0.0101</td>
</tr>
<tr>
<td>Sainsbury dummy</td>
<td>-0.0119</td>
<td>-0.0282</td>
<td>-0.0402</td>
<td>-0.0115</td>
<td>-0.0614</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.00044</td>
<td>0.00085</td>
<td>0.00111</td>
<td>0.00072</td>
<td>0.00728</td>
</tr>
<tr>
<td>Asda dummy</td>
<td>-0.00366</td>
<td>-0.01403</td>
<td>-0.0251</td>
<td>-0.0105</td>
<td>0.0060**</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.00044</td>
<td>0.00087</td>
<td>0.00113</td>
<td>0.00072</td>
<td>0.0068</td>
</tr>
<tr>
<td>BROL</td>
<td>0.0169</td>
<td>0.0359</td>
<td>0.0442</td>
<td>0.00734</td>
<td>0.1446</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.00041</td>
<td>0.00082</td>
<td>0.00112</td>
<td>0.00078</td>
<td>0.00849</td>
</tr>
<tr>
<td>High-priced good</td>
<td>-0.00139</td>
<td>-0.0020*</td>
<td>0.0141</td>
<td>0.0164</td>
<td>-0.1192</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.00043</td>
<td>0.00083</td>
<td>0.00108</td>
<td>0.00070</td>
<td>0.00601</td>
</tr>
<tr>
<td>Frequently purchased</td>
<td>-0.0010</td>
<td>-0.0171</td>
<td>-0.00143</td>
<td>0.00016*</td>
<td>-0.0102</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.00007</td>
<td>0.00011</td>
<td>0.00014</td>
<td>0.00008</td>
<td>0.00084</td>
</tr>
<tr>
<td>Change in CPI (pos)</td>
<td>-0.0222</td>
<td>-0.0250</td>
<td>0.0297</td>
<td>0.0541</td>
<td>-0.4092</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.00164</td>
<td>0.00316</td>
<td>0.00408</td>
<td>0.00261</td>
<td>0.0218</td>
</tr>
<tr>
<td>2008 dummy</td>
<td>-0.00556</td>
<td>0.0274</td>
<td>0.0860</td>
<td>0.0542</td>
<td>-0.1638</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.00158</td>
<td>0.00410</td>
<td>0.00582</td>
<td>0.00449</td>
<td>0.02119</td>
</tr>
<tr>
<td>Sains *2008</td>
<td>-0.0113</td>
<td>-0.0337</td>
<td>-0.0477</td>
<td>-0.00609</td>
<td>-0.1098</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.0005</td>
<td>0.00115</td>
<td>0.00180</td>
<td>0.00157</td>
<td>0.01577</td>
</tr>
<tr>
<td>Asda *2008</td>
<td>-0.00179*</td>
<td>0.00053**</td>
<td>0.00147**</td>
<td>0.00256**</td>
<td>-0.0073**</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.00087</td>
<td>0.00198</td>
<td>0.00271</td>
<td>0.00183</td>
<td>0.0130</td>
</tr>
<tr>
<td>BROL *2008</td>
<td>0.0311</td>
<td>0.0614</td>
<td>0.0533</td>
<td>-0.00445</td>
<td>0.1723</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.00402</td>
<td>0.00424</td>
<td>0.00408</td>
<td>0.00170</td>
<td>0.0149</td>
</tr>
<tr>
<td>Hi-price *2008</td>
<td>0.0011**</td>
<td>-0.0256**</td>
<td>-0.00788</td>
<td>-0.00420</td>
<td>0.02009</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.0009</td>
<td>0.00174</td>
<td>0.00229</td>
<td>0.00144</td>
<td>0.01206</td>
</tr>
<tr>
<td>Freq purchased *2008</td>
<td>0.0005</td>
<td>-0.00134</td>
<td>-0.00333</td>
<td>-0.00167</td>
<td>0.00767</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.0001</td>
<td>0.00026</td>
<td>0.00033</td>
<td>0.00020</td>
<td>0.00172</td>
</tr>
<tr>
<td>Change CPI *2008</td>
<td>0.0445</td>
<td>0.0899</td>
<td>0.0690</td>
<td>-0.0314</td>
<td>0.4942</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.00279</td>
<td>0.00559</td>
<td>0.00748</td>
<td>0.00475</td>
<td>0.0374</td>
</tr>
<tr>
<td>n</td>
<td>405150</td>
<td>405150</td>
<td>405150</td>
<td>405150</td>
<td>40381</td>
</tr>
<tr>
<td>Goodness of fit measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efron</td>
<td>0.0203</td>
<td>0.0361</td>
<td>0.0374</td>
<td>0.0075</td>
<td>0.0513</td>
</tr>
<tr>
<td>McFadden</td>
<td>0.0740</td>
<td>0.0593</td>
<td>0.0470</td>
<td>0.0217</td>
<td>0.0392</td>
</tr>
<tr>
<td>Veall/Zim.</td>
<td>0.0880</td>
<td>0.0845</td>
<td>0.0751</td>
<td>0.0285</td>
<td>0.0868</td>
</tr>
<tr>
<td>Rsq ML</td>
<td>0.0152</td>
<td>0.0271</td>
<td>0.0300</td>
<td>0.0070</td>
<td>0.0508</td>
</tr>
<tr>
<td>Coefficient values represent marginal effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*significant only at 10% level (all other variables significant at 1% or better)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>** not significant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5. Concluding remarks

Our core finding is that supermarket basket prices move in quite different ways from the prices of their constituent products, at least within our sample that encompasses the most significant UK supermarkets. Individual prices are very flexible. But whilst most prices are falling, basket prices can and do rise. In particular, our study has shown that the period in 2008 where there was rapid cost inflation in basic food commodities was characterised by the UK supermarket chains raising basket prices. However, there is an aggregation issue. This was accompanied by a welter of small value price falls, leading to the underlying movements in basket prices being obscured.

In the present paper, we are not concerned directly with whether prices are “sticky” or not; in some senses we view this focus on individual prices as misplaced. Nevertheless, our results relate centrally to mechanisms for transmission of inflationary pressures. Within our time frame, there is one major inflationary incident amid general mild inflation, namely the hike in costs and prices in early 2008, and we are able to observe pricing behaviours before, during and after this incident. However we cut the data, what shines out is the core response, which we argue is consistent with attempts to obfuscate (and hence to facilitate) the fact that basket prices are rising.

In a recent report to the Treasury Select Committee, Bean (2011) (Deputy Governor of the Bank of England) noted that one of the reasons CPI inflation experience in the UK has been markedly higher than the Bank had expected is that “… we appear to have significantly under-estimated the degree of pass-through from Sterling’s 2007-8 depreciation …”. Our work relates to this quite closely. The big rise in commodity prices in early 2008 largely came from overseas. Our paper exposes the behaviour of key players in the inflationary process in their desire to raise prices in a competitive environment, hence providing some explanation for this changed outcome. By obscuring the true nature of underlying upwards price movements through significant small price cuts, it appears that the major supermarket chains facilitated the increased degree of pass-through in 2008.
Data Appendix

Data sources

As stated in the text, our main sources of data are the sets of pricing data downloaded on a weekly basis from two sources, Tesco Pricecheck (covering the period Nov 2003 to Nov 2008 for the retailers Tesco, Sainsbury’s, Safeway/Morrison’s, ASDA) and MySupermarket (from Feb 2008 to Nov 2010, for Tesco, Sainsbury’s, Ocado, ASDA). The datasets were obtained by matching item descriptions across weeks and across supermarkets. Where branded products are concerned, the match was exact. Where own label products are involved, they were matched across own label products at the same quality level. The approach was to collect data only for matched products, but within this to collect as near a full sample as possible. However, we discovered a significant degree of missing price data. Therefore, we engaged in various procedures in order to minimise the problems this might cause. These are described more fully below.

To this were appended various elements of additional data. We collected data and information on methodology from official statistics of pricing indices for the UK using the Office of National Statistics website,


We obtained data on index prices of various world-traded commodities from IMF from

http://www.indexmundi.com/commodities/?commodity=commodity-price-index&months=300

We created purchase frequency as PF=(Total number of trips/Total Number of shoppers) which were obtained from TNS (Taylor Nelson Sofres) group.

Using the Nielsen Media Research source, we aggregated across media and categories the media spend (advertising) information they provide.

Missing values and balanced Sample

In order to trace item price changes over time we derived a balanced panel of item prices for the ‘Big 3’ retailers, Tesco, Sainsbury’s and ASDA, (and in the case of our first dataset, Morrison in addition) common to the datasets mentioned above. Our main problem was to make comparison of items across retailers whilst also maintaining item coherence (the latter being a problem for branded items as well) as names and some quantities altered through the sample period. Tesco PriceCheck maintained comparisons of items across the three retailers and by comparing week on week for name components, size and price, 849 candidates for the balanced sample were found. By using similar search methods on the Mysupermarket data cross referencing from the Tesco PriceCheck
on name, price, weight and product type the initial sample was extended to cover in total the seven year period 10th November 2003 to 10th November 2010.

These items were subjected to rigorous testing for missing data through start, body and end of sample, and also by the similarity of Tesco PriceCheck and Mysupermarket prices during the common period where both datasets were available (Feb 2008 to Nov 2008). The ‘best’ 370 items were chosen predominantly by their minimal missing values. In balancing objectives, we negatively weighted longer continuous missing weeks’ data, say of length n, in an item series more than n weeks broken up into smaller groups of missing observations. The reasoning was that much more price variation can be traced (and imputed) from the latter item’s series. By reducing the sample from 849 to 370 the missing values as a percentage of the total observations dropped from 16.8% to 11.7% where about 80% of the gaps are accounted for by the 34 missing weeks in the Tesco PriceCheck sample during the Christmas/New year breaks, the remaining 20% caused essentially by missing stock. By selecting in this way, missing values were minimized and more heterogeneously dispersed. A similar process was used in generating the shorter 600 items sample of the big 4 supermarkets from the Tesco PriceCheck data but the number of missing observations were affected relatively more by losses early in the sample and the Christmas/New year breaks.

The handling of remaining missing values, in both samples, was an important issue in this work. The key consideration is not to generate apparent results on price movements where none are warranted. For cases where no apparent price change occurred, for example a series ‘25p 25p Missing 25p 25p’, then 25p would be inserted into the ‘Missing’ observation week. Likewise missing values at the very start or very end of the sample, which were limited to no more than 4 weeks, were filled by the nearest neighbour's value. For identifiable price changes, for example the series ‘25p Missing Missing Missing 26p’ a probability of price change variable was created for each week. Probabilities were imputed from the ratio of changes from available data. Where data were not available, because of missing weeks, they were imputed from knowledge that over the period of missing weeks a certain number of prices were altered, these were compared to the average ratio of the number of three week period of price change to the number of individual of price changes in each of the three weeks. Hence a good approximation of the probability of price change per week over the missing weeks could be found. These probabilities were then used to randomly associate a known price change to a week in the missing period. A key underlying assumption in this analysis is that where data are missing, this tells us nothing as such about the nature of the price or product that is idiosyncratic to the observation (in other words, in the language of imputation, the observations are “missing at random”). Clearly, we would not wish to analyse behaviour of supermarkets specifically over the Christmas period using these data, but since we are considering inflationary periods generally, the missing at random hypothesis seems innocuous.
References:


Competition Commission (2003), *Safeway plc and Asda Group Limited (owned by Wal-Mart Stores Inc); Wm Morrison Supermarkets PLC; J Sainsbury plc; and Tesco plc: A report on the mergers in contemplation*, London.


