# Cardiff Economics Working Papers


Working Paper No. E2013/1

# What we can learn about the behavior of firms from the average monthly frequency of price-changes: an application to the UK CPI data

*Huw Dixon and Kun Tian*

January 2013

Cardiff Business School
Aberconway Building
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

# What we can learn about the behavior of firms from the average monthly frequency of price-changes: an application to the UK CPI data.*

Huw David Dixon† Kun Tian.‡

December 6, 2012

**Abstract**

The monthly frequency of price-changes is a prominent feature of many studies of the CPI micro-data. In this paper, we see how much this ties down the behavior of price-setters ("firms") in steady-state in terms of the average length of price-spells across firms. We are able to divide an upper and lower bound for the mean duration of price-spells averaged across firms. We use the UK CPI data at the aggregate and sectoral level and find that the actual mean is about twice the theoretical minimum consistent with the observed frequency. We estimate the distribution using the hazard function and find that although the estimated hazard differs significantly from the Calvo distribution, the means and medians are similar. However, despite the micro differences, we find that the artificial Calvo distributions generated using the sectoral frequencies result in very similar impulse responses to the estimated hazards when used in the Smets-Wouters (2003) model.

JEL: E50.

Keywords: Price-spell, steady state, duration.

*For helpful comments and suggestions we would like to thank Erwan Gautier, James Costain, seminar participants at the Cardiff University, the University of Manchester, and the T2M conference at Nantes in 2012. Huw Dixon would like to thank the Foundation Banque de France for financial support and Kun Tian the Julian Hodge foundation. Faults remain our own.

†Cardiff Business School, Colum Drive, Cardiff. CF10 3EU, dixonh@cardiff.ac.uk.

‡Cardiff Business School, Tiank1@cardiff.ac.uk.

# 1 Introduction

In recent years, there have been many studies using comprehensive microdata on pricing. In the Euro area, there has been the inflation persistence network (*IPN*) consisting of national studies of the CPI and PPI microdata[1], which are summarized in Alvarez et al (2006). In the US there have been similar studies: Bils and Klenow (2004), Klenow and Kryvtsov (2008), Nakamura and Steinsson (2008)[2]. One common focus of these studies has been the statistic of the proportion of prices changing per month (this can either be an average over several months, or a monthly statistic). This statistic can be presented in several ways, depending on the level of disaggregation and the treatment of temporary sales and so on. In this paper, we seek to analyze what this statistic implies for the behavior of firms (or more accurately price-setters) in the economy. Each period firms set prices: they may either choose to leave the price unchanged or to change it. The proportion of firms resetting price corresponds to the proportion of prices changing (for simplicity we take a 1-1 correspondence between firms and prices). The prices of some product types change frequently (e.g. gasoline, tomatoes) and some very infrequently.

We can think of the CPI dataset as a panel of observations, each cross-section corresponding to the prices set by "firms" at that date. The cross-sectional mean completed price spell can be seen as capturing the mean behavior of the price-setters, which represents the "structure" of the economy in this respect (i.e. the average behavior of the firms in the economy). For a given frequency of price change, what can we infer about the behavior of the firms? In this paper we are able to derive a lower bound (and an upper bound) for the mean length of price-spells across firms, interpreted as the cross-sectional mean completed price-spell. The cross-sectional distribution is needed if we are to model price-setting as a Taylor process. We then use the UK CPI data for the period 1996-2007 and consider frequency data at three different levels of disaggregation: the 11 sector COICOP, the 67 sector COICOP and the highest possible level of disaggregation at 570 items, to see how the actual data on price-spell durations compares to the theoretical minimum. We find that the actual mean estimated from the CPI is 10.9 months, which is 62-90% higher than the theoretical minima generated from the frequency data. This is not surprising, in Proposition 1 we find that the theoretical minimum consistent with a given frequency is only attained if all price-spells have the same or almost the same duration, whilst the

[1] See Beaudry et al (2007) and Alvarez and Hernando (2006) for France and Spain *inter alia*.

[2] See Bunn and Ellis (2012) for the UK and Baharad and Eden (2004) for Israel.

actual distribution contains a lot of heterogeneity in durations which implies a longer cross-sectional mean.

We also interpret the frequency data under the hypothesis that within the sector the frequency is generated by a Calvo distribution, as has been assumed in applied work by Dixon and Kara (2010, 2011). We look at this in two ways. First, we aggregate over all sectors to derive the aggregate distribution under this assumption: thus we have the distribution of durations in each sector and for each duration we aggregate over sectors using the sectoral CPI weights. We can then compare this to the "true" distribution derived from the estimated hazard function.

- The aggregate distribution derived under the Calvo hypothesis at the sectoral level has a similar mean and median to the true distribution, with the mean increasing with the level of disaggregation. For example, the 570-item model yields a mean of 10.8 and median of 7.8 months, whilst the true values are 10.9 and 7.8 respectively.

- However, the implied Calvo distributions differ to the true distribution in significant ways: (i) there is a 12 month spike in the true distribution absent from the Calvo distributions, (ii) the proportion of short price-spells (1-3 months) is less in the Calvo than in the true distribution.

We also examine whether the Calvo is a good fit in each of the 11 COICOP sectors. Since the data set is large, even small deviations of the actual distribution from the hypothetical Calvo distribution cause the Calvo null to be rejected under the Kolmogorov-Smirnov test, which is indeed the case. However, whilst in some sectors the hypothesized Calvo distribution looks completely different (for example in Health which has a very large 12 month spike), in others the Calvo looks more similar (Transportation). There is clearly a variety of patterns across sectors[3] when we compare the true distributions within the 11 COICOP sectors.

Whilst the Calvo distributional hypothesis might not provide a good statistical fit in terms of the aggregate or sectoral distributions, does this matter in terms of how the economy behaves? Since the Calvo hypothesis yields a mean and median close to the true distribution, perhaps the differences will not result in different behavior of the economy in terms of impulse response functions. We explore this in the context of two macro models: a simple Quantity Theory model and the Smets and Wouters (2003) Euro area model. The pricing models used are the Generalized Taylor and Calvo as in

[3]This heterogeneity across sectors was also found in the French data by Fougere et al (2007).

Dixon and Le Bihan (2012), which are both consistent with any micro distributions of durations and can be calibrated to the true data and the data under the Calvo distributional hypothesis. What we find is that the impulse response functions for output and inflation are very similar when we use the true distribution and hypothetical Calvo distribution at the different levels of disaggregation for both the Generalized Taylor and Generalized Calvo. In the Smets-Wouters model, the IRFs are almost identical. This indicates that if we are interested in modelling macroeconomic properties of an economy, using the sectoral frequency data under the Calvo distributional hypothesis might be a useful shortcut and alternative to estimating the distribution using the hazard function. Indeed, where the actual hazard is not available or not estimated reliably, we can be confident that the use of sectoral frequencies with the Calvo distributional hypothesis can be a good working approximation.

The structure of the paper is follows. In Section 2, we give a theoretical description of the steady state distributions of durations. In Section 3, we derive the propositions in which the average duration across firms consistent with the mean frequency of price changes. We show an application to the UK CPI data in Section 4. In Section 5, we simulate a quantitative theory model and a Smets-Wouters model, both with Generalized Taylor and Generalized Calvo price-setting respectively. We conclude in Section 6.

# 2 Steady state distributions of durations.

The statistical framework for understanding the CPI microdata is outlined in detail by Dixon (2012), so in this paper we just summarize in a less technical manner the key properties needed for this paper.

There is a continuum of price-setting firms $f \in [0,1]$, time is discrete[4] and infinite $t \in Z_+ = \{0, 1, 2...\infty\}$. The price set by firm $f$ at time $t$ is $p_{ft}$. A price spell is a duration, a sequence of consecutive periods that have the same price. For every $\{t, f\} \subseteq [0,1] \times Z_+$ we can assign an integer $d(t,f)$ which is the duration of the price-spell to which $p_{ft}$ belongs[5]. The distribution of price-spell durations is simply the proportions of all durations having length $i = 1...F$: $\boldsymbol{\alpha}^d = \left\{\alpha_i^d\right\}_{i=1}^F \in \Delta^{F-1}$. We assume a steady-state, so that the distribution of durations of new price-spells is the same for each new cohort

[4] Typically, CPI data are collected on a monthly basis, the price observations being obtained in the first two weeks of the calender month.

[5] Note that in assigning an integer to a duration, we start with 1 by convention: it would be equally valid to start with 0. With our convention, a new price-spell is 1 month old, rather than becoming 1 on completion of the first month.

of price-spells. This means that the distribution of all price-spells is exactly the same as the distribution of new price spells at any period.

Whilst the distribution of durations $\boldsymbol{\alpha}^d \in \Delta^{F-1}$ is one way of looking at the microdata, it ignores the *panel* structure of the data. Each row of the panel is a trajectory of prices corresponding to a particular firm (or more accurately product sold at an outlet). Each column is a cross-section of all of the prices set by firms at a point in time. The cross-sectional distribution of *completed* price-spell durations is $\boldsymbol{\alpha} \in \Delta^{F-1}$. In effect, we take a representative $t$, and for each firm we see the completed price-spell duration at that time $d(f,t)$. $\alpha_i$ is then the proportion of firms that set prices for $i$ periods. Equivalently, it is the proportion of price-spells at any time t which last for $i$ periods.

The proportion of firms re-setting price each month is denoted as $\bar{h}$ : in the UK this is equal to 21%. We define the mean duration[6] of price-spells across the panel as a whole as

$$\bar{d}\left(\boldsymbol{\alpha}^d\right) = \sum_{i=1}^{F} i\alpha_i^d$$

and cross-sectional mean (across firms) as

$$\bar{T}\left(\boldsymbol{\alpha}\right) = \sum_{i=1}^{F} i\alpha_i \tag{1}$$

Note that the cross sectional mean in general be larger than the mean duration $\bar{T} \geq \bar{d}$: this is because in cross-section you have length-biased sampling, since the probability of a price-spell being observed in cross-section is proportional to its duration. Indeed, the two can be equal ($\bar{T} = \bar{d}$) *if and only if* $\alpha_F = \alpha_F^d = 1$, so that all price-spells are $F$ months long and there is no heterogeneity to generate a length bias.

From Dixon (2012), we know that[7]:

$$\bar{h} = \bar{d}^{-1} \tag{2}$$

$$= \sum_{i=1}^{F} \frac{\alpha_i}{i} \tag{3}$$

[6] Again, this way of defining the mean is consistent with our convention of assigning the integer 1 to the first time period. Had we instead assigned a 0 to this value, then we would have the expression $\sum_{i=1}^{F}(i-1)\alpha_i$ (as we do with human ages). An equally acceptable measure is to take the midpoint and have $\sum_{i=1}^{F}(i-0.5)\alpha_i$ We can move between these definitions simply by adding or subtracting a constant.

[7] In continuous time, we have $\bar{d} = -\frac{1}{\log(1-\bar{h})}$, which allows for the price to change more than once per period. Again, we are using a discrete time setting in which durations are integer valued.

That is, the proportion of firms resetting price is equal to the reciprocal of the mean duration $\bar{d}$. Furthermore, the proportion of firms resetting price is related to the cross-sectional distribution by equation (3). In steady-state, a proportion $i^{-1}$ of the $\alpha_i$ $i-$duration firms reset their price. The aggregate proportion is simply the sum over the durations $i = 1...F$.

# 3 The average duration across firms consistent with $\bar{\mathbf{h}}$.

Now, for a given frequency $\bar{h}$ there are many possible distributions across firms (DAF) $\boldsymbol{\alpha} \in \Delta^{F-1}$ consistent with identity (3) and each distribution results in a corresponding mean across firms $\bar{T}(\boldsymbol{\alpha})$. We can define the mapping $H\left(\bar{h}\right) : [0,1] \to \Delta^{F-1}$

$$H\left(\bar{h}\right) = \left\{ \boldsymbol{\alpha} \in \Delta^{F-1} : \sum_{i=1}^{F} \frac{\alpha_i}{i} = \bar{h} \right\}$$

$H\left(\bar{h}\right)$ is the set of all $DAFs$ which are consistent with a given mean duration of price-spells $\bar{d}$ expressed in terms of the corresponding proportion of firms resetting prices $\bar{h}$. Clearly, since the maximum duration is $F$, we have $\bar{h} \geq F^{-1}$ so that $H$ is non-empty. Since $H\left(\bar{h}\right)$ is defined by a linear restriction on the sector shares $\boldsymbol{\alpha}$, $H\left(\bar{h}\right) \subset \Delta^{F-2}$ and is closed and bounded. We can then ask what is the minimum (maximum) $\bar{T}$ consistent with a given $\bar{h}$. Since $H\left(\bar{h}\right)$ is non-empty, closed and bounded, with $\bar{T}\left(\boldsymbol{\alpha}\right)$ continuous, both a maximum and a minimum will exist. Turning to the minimization problem first, we have:

$$\min \ \bar{T}\left(\boldsymbol{\alpha}\right) \quad s.t. \ \boldsymbol{\alpha} \in H\left(\bar{h}\right) \tag{4}$$

**Proposition 1** Let $\boldsymbol{\alpha}^{\min} \in \Delta^{F-1}$ solve (4) to give the shortest average contract length $\bar{T}^{\min}$.

(a) No more than two sectors $i$ have values greater than zero

(b) If there are two sectors $\alpha_i > 0$, $\alpha_j > 0$ then will be consecutive integers ($|i - j| = 1$).

(c) There is one solution iff $\bar{h}^{-1} = k \in Z_+$. In this case, $\alpha_k = 1$.

(d) The minimum is $\bar{T}^{\min} = \bar{h}^{-1} = \bar{d}$.

We can also ask what is the *maximum* average contract length consistent with a proportion of re-setters $\bar{h}$:

$$\max \ \bar{T}\left(\boldsymbol{\alpha}\right) \quad s.t. \ \boldsymbol{\alpha} \in H\left(\bar{h}\right) \tag{5}$$

**Proposition 2** Let $\boldsymbol{\alpha}^{\max} \in \Delta^{F-1}$ solve (5). Given the longest contract duration $F$, the distribution of contracts that maximizes the average length of contract subject to a given proportion $\bar{h}$ of firms resetting price

$$\begin{aligned}
\alpha_F^{\max} &= \frac{F}{F-1}\left(1-\bar{h}\right) \\
\alpha_1^{\max} &= \frac{F}{F-1}\bar{h} - \frac{1}{F-1}
\end{aligned}$$

with $\alpha_i^{\max} = 0$ for $i = 2...F-1$. The maximum average contract length is

$$\bar{T}^{\max} = F\left(1-\bar{h}\right) + 1$$

To understand Propositions 1 and 2, we just need to think of what is generating the mean duration $\bar{d}$ and the proportion of firms changing price each period $\bar{h}$. There is the unit interval of firms, divided into proportions with different price-spell durations $i = 1...F$. Firms with price-spell lengths $i$ will set prices once every $i^{-1}$ periods: the longer the price-spell, the more infrequently the firm will reset price. Hence, we can have the same proportion of firms re-setting price (and hence same mean duration) and increase the mean duration across firms by more longer price-spells. The maximum $T^{\max}$ is reached when we have as many $F$ period contracts as possible, consistent with $\bar{h}$. In effect, this means we have a mix of 1 period and $F$ period price-spells. The existence of a maximum relies on us assuming an upper bound $F$: clearly, as $F \to \infty$, $T^{\max} \to \infty$. The minimum occurs when all firms have similar price-spells: if $\bar{d}$ happens to be an integer, then all price-spells have that length and the two distributions are the same: $\boldsymbol{\alpha}^d = \boldsymbol{\alpha}$.

# 4 An application to the UK CPI data[8].

In this section, we take the frequency data from the UK and apply the two propositions to derive the implied upper and lower bounds for the cross-sectional mean duration. We then estimate the actual distribution and the corresponding mean and how it compares to the theoretical distributions that yield the maximum and minimum mean durations across-firms. The dataset

[8] This work contains statistical data from ONS which is Crown copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates

we use for is the locally collected CPI price microdata covering the period January 1996 to December 2007 which is described in detail in the Appendix. The period covered corresponds to the Great Moderation period when the frequencies would have been stationary. We also want to see how the level of disaggregation affects the results. For the UK, we have the following levels of disaggregation available from the ONS:

- 11 COICOP categories
- 67 disaggregated COICOP categories.
- 570 items.

Each of these disaggregations represents exactly the same data. To get an idea of the level of aggregation, we can depict the broad 11 COICOP categories (excluding education which is not included in the VML dataset) in Table 1. For example, there is the category "food and non-alcoholic beverages" which represents 17.6% of the CPI weight in the subsample available in the dataset. The second level of disaggregation subdivides these into a total of 67 COICOP subcategories. For example, within ""food and non-alcoholic beverages" there are 11 subcategories: 2 for drinks (tea, coffee and cocoa; mineral water, soft drinks and juices) and 9 for food (such as meat, fish, fruit). The lowest level of disaggregation is the item level. An item is a particular product or service on which the price observation is made. For example: canned sweet corn (198g-340g); coffee - take-away; fresh lettuce (iceberg). The 570 items we include are all of the items which were included throughout the sample period - it excludes old items which were either discontinued or new items introduced in this period. These items represent over 66.4% of the total CPI.

Firstly, we present in Table 1 and for each category, we have the frequency data, which gives the proportion of items changing price in a given month[9]. The items are weighted by the appropriate CPI weight. The data are represented in Table 1.

In the first column of the Table 1 is the COICOP sector, in the second is the CPI weight for the sector, normalized so that they add up to 100 (since Education is excluded) and the third the frequency (to three d.p.). In the fourth we have the minimum average duration (MAD) in that sector from

[9] In a given month, the percentage of current prices (for items at a location) that are different from the price set in the previous month for the same product at the same location. The figure excludes items for which there was no observation the month before (e.g. it is the first price observation of the item at the outlet).

| COICOP Category. | CPI adj | freq | Min | Max |
|---|---|---|---|---|
| Transport | 10.4 | 36.0 | 2.8 | 29.2 |
| Alcoholic Beverages and Tobacco | 7.1 | 27.6 | 3.6 | 32.9 |
| Clothing and Footwear | 9.3 | 27.2 | 3.7 | 33 |
| Food and Non-Alcoholic Beverages | 17.6 | 26.0 | 3.8 | 33.6 |
| Furniture and Home Maintenance | 11.3 | 22.7 | 4.4 | 35 |
| Communications | 0.2 | 22.5 | 4.4 | 35.1 |
| Recreation and Culture | 9.9 | 20.0 | 5 | 36.2 |
| Housing and Utilities | 8.3 | 13.7 | 7.3 | 39 |
| Miscellaneous Goods and Services | 6.5 | 12.7 | 7.9 | 39.4 |
| Restaurants and Hotels | 17.5 | 10.5 | 9.5 | 40.4 |
| Health | 1.9 | 10.4 | 9.6 | 40.4 |
| Total | 100 | 21.4 | 4.7 | 35.6 |

Table 1: COICOP 11 sectoral frequencies

*"Freq" denotes the frequencies of prices changes, which are reported in percent per month. "CPI adj" denotes the adjusted CPI expenditure weight of the CPI sectors after excluding the Education sector. "Min" denotes the minimum average duration. "Max" denotes the maximum average duration*

Proposition 1, and in the fifth the maximum from Proposition 2 based on the assumption that the longest price-spell is 44 months.

We next generate the cross-sectional distribution in the whole economy corresponding to the minimum average duration consistent with the observed frequencies. From Proposition 1, in each sector we will have one or two durations with a non-zero share. In Recreation & Culture, since 20% of prices change per month, there are just 5 month price-spells. In Food & Non-Alcoholic Beverages there will be a mixture of 20% 3 month and 80% 4 month price-spells. The shortest durations are 2 months (in Transport) and the longest 10 months (in Health). For each duration, we can then add up across the 11 sectors to get the weighted cross-sectional distribution. This is depicted in Figure 1:

We estimate the actual cross-sectional distribution as in Dixon (2012) using the hazard function. We can see that the distributions are completely different, although derived from exactly the same data. The "minimum duration" distribution has no 1 month, 6 month, 11 month or 12 month durations; the most common durations are 4 months (32% ), then followed by the 3 months (16%) and 5 months (14%). Among the rest, we find that the share of distribution coincidently according with the length of duration, such as 10 months (10%) and 9 months (9%), 8 months (8%), 7 months (7%), 2

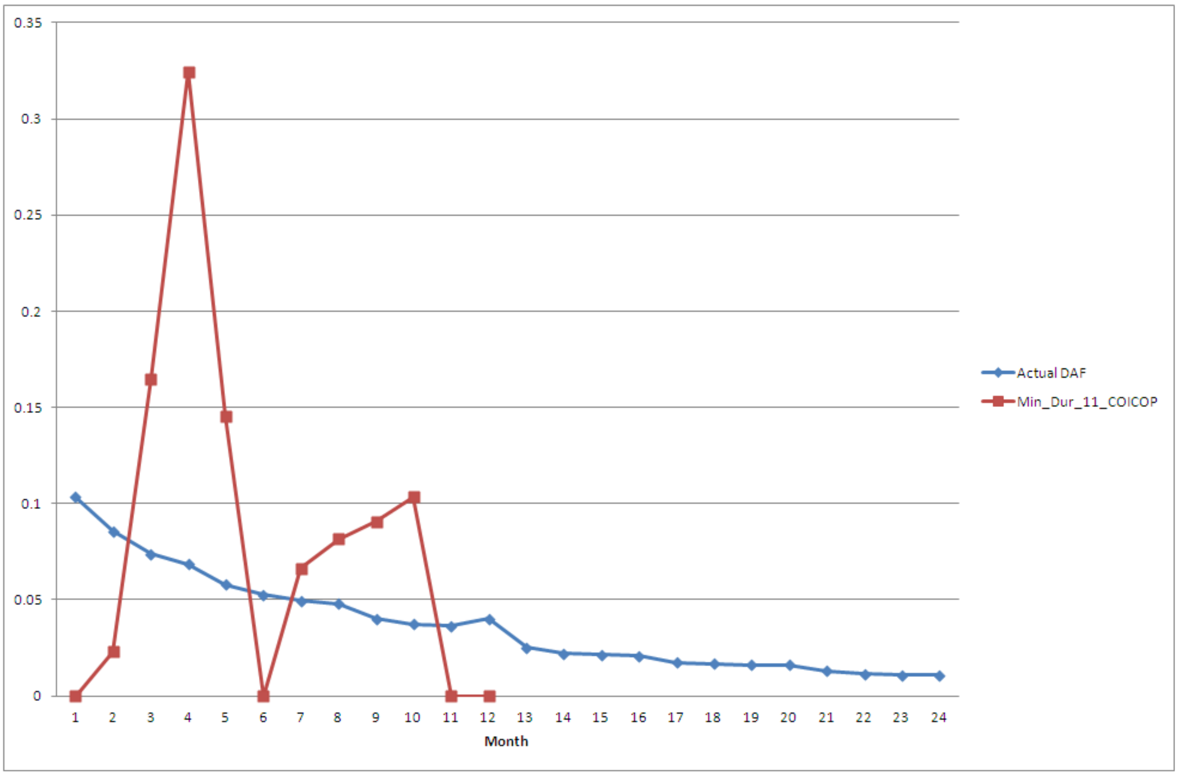



Figure 1: Actual DAF vs. *minumum duration* distribution

| | Mean DAF | Median DAF |
|---|---|---|
| True | 10.9 | 7.8 |
| 11_min | 5.5 | 4 |
| 67_min | 6.1 | 4.5 |
| 570_min | 6.7 | 5.8 |

Table 2: The minimum mean and median duration accross firms comparison *"True" denotes the actual cross-sectional distribution implied by hazard function, "11_min", "67_min", and "570_min" denote the cross-sectional distributions derived from the "minimum method" at different disaggregation level, corresponding to* 11 COICOP categories,67 disaggregated COICOP categories,and 570 items *respectively. "Mean DAF" and "Median DAF" denote the mean and median length of duration across firms, and both of them are in unit of month.*

months (2%). In contrast, the "true" distribution is much flatter with a long tale (which we have truncated at 24 months). The most common duration is 1 month (10.3%) closely followed by 2 months (8.5%). There is an annual spike at 12 months (4%). Whilst the longer durations tend it have lower shares the cross-sectional distribution is non-monotonic. The maximum duration distribution is to have a mix of one month and the maximum duration (44 months). As a first approximation, the share of the one month durations in each sector is a little less than the frequency. This is clearly very different from the actual distribution.

We can now see what the effect of further disaggregation is on the "minimum duration" distributions: we perform the same procedure for the COICOP 67 and the 570 item level. These are all depicted in Figure 2, along with the COICOP 11 and the true distribution. They share some common features when compared to the true distribution: they all put too little weight on month 1, month 12, and after. They all put too much weight on months 3-5 and months 9 and 10. However, they are also quite different. The level of disaggregation clearly matters when constructing a possible cross-sectional distribution. We can see this from the mean and median durations in Table 2.

These are far too short, reflecting the fact that the minimum duration distributions put a large weight on the shorter distributions and do not have a long fat tail as in the data. In fact, the minimum durations are just over half the actual mean duration.

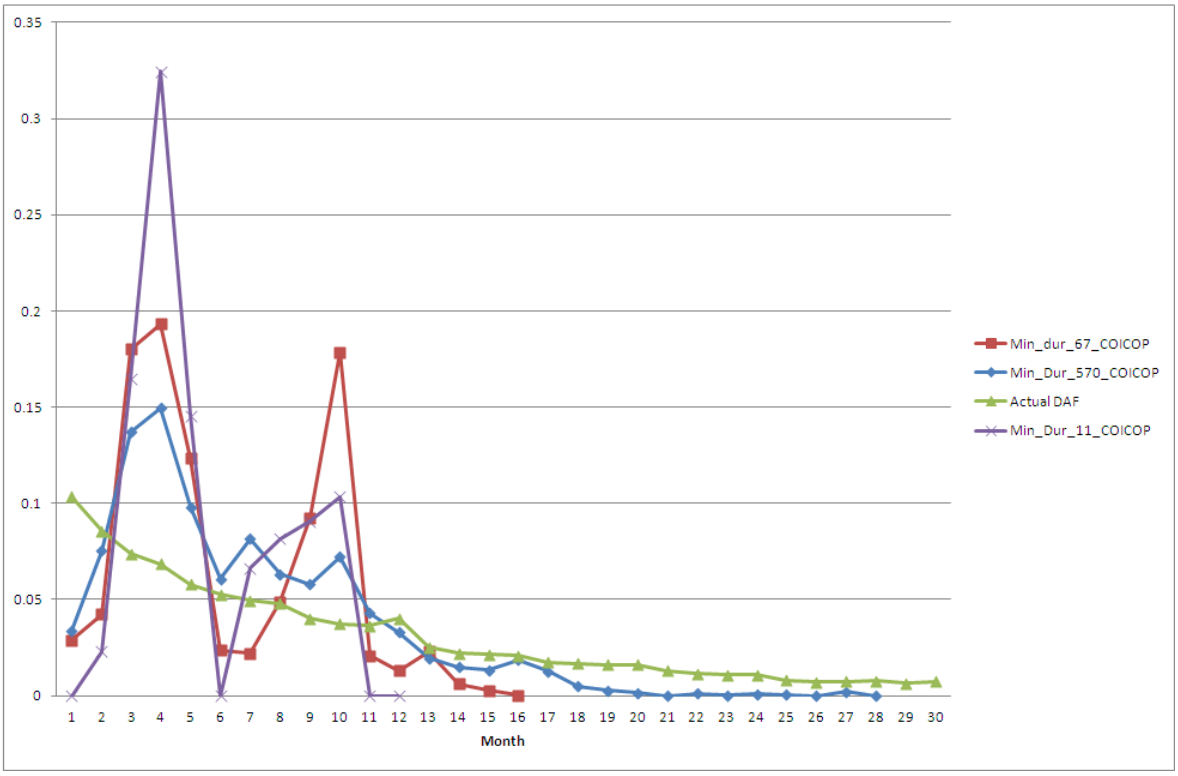



Figure 2: Actual DAF vs. *minimum duration* distribution.

### 4.1 Calvo distributions.

Clearly, the "minimum duration" distributions corresponding to Proposition 1 do not look at all like the true distribution. In this section we look at the distribution generated by the hypothesis that the sectoral frequencies are generated by a Calvo distribution. Again, as in the previous section, we are looking at exactly the same data, just at different levels of aggregation. Within each sector $k = 1...N$, we observe a sectoral frequency of $h_k$. As shown in Dixon and Kara (2006), this corresponds to the cross-sectional distribution for that sector $\boldsymbol{\alpha}_k = \{\alpha_{ik}\}_{i=1}^{\infty}$ where:

$$\alpha_{ki} = i.h_k^2(1 - h_k)^{i-1} \tag{6}$$

Each sector has a CPI weight $c_k$. We can then aggregate across the $N$ sectors using the CPI weights to get the share of each duration across all sectors $\boldsymbol{\alpha} = \{\alpha_i\}_{i=1}^{\infty}$ where:

$$\alpha_i = \sum_{k=1}^{N} c_k \alpha_{ki} \tag{7}$$

The mean duration of the Calvo distribution at the sectoral level is[10]:

$$\bar{T}_k^C = 2h_k^{-1} - 1. \tag{8}$$

The mean of the aggregate distribution is thus:

$$\bar{T} = \sum_{k=1}^{N} c_k \bar{T}_k^C$$

This is the method used in Dixon and Kara (2010, 2011) for generating the Bils-Klenow distribution based on the Bils Klenow (2004) appendix dataset of 350 sectoral frequencies for the US.

It is important to note that by assuming a Calvo *distribution*, we are not assuming a Calvo *pricing model* within each sector. We are simply describing the distribution of price-spell durations in each sector generated by a constant hazard rate that is equal to the sectoral frequency. This is purely descriptive of the distribution. It is perfectly compatible with a Taylor model, where within each sector the length of the price-spells is known *ex ante*. What we are doing in effect is constructing $\boldsymbol{\alpha} = \{\alpha_i\}_{i=1}^{\infty}$ using (7): that means we take out all of the $i$ duration spells from each sector $k$ and put them together

[10] Dixon and Kara (2006), Theorem 1.

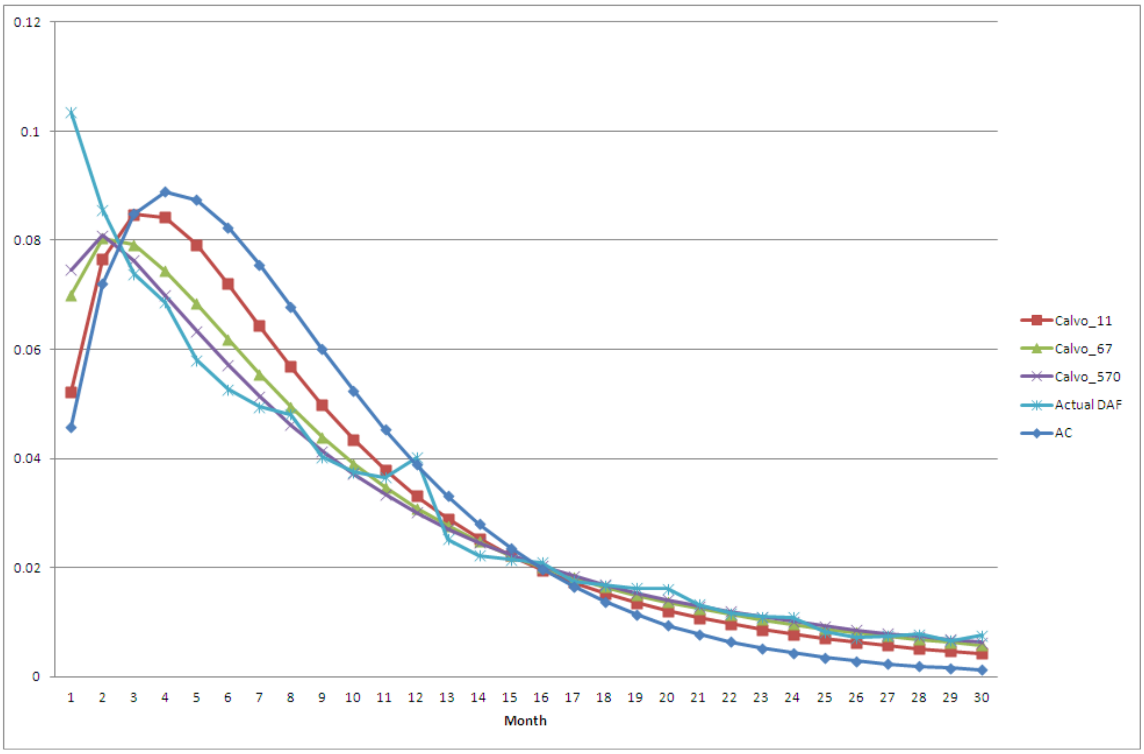



Figure 3: Calvo distributions at different levels of aggregation vs. Actual DAF.

into a "duration sector" $\alpha_i$, which includes all of the price-spells of length $i$ in the economy. The key difference between the Calvo and Taylor pricing frameworks is that under Taylor the firms know the length of the price-spell when they set the price, whereas in the Calvo they do not.

In Figure 3, we represent the Calvo distributions at different levels of aggregation: the one sector "aggregate Calvo" (AC) distribution based on the mean UK frequency of 0.2140; the 11 sector COICOP, the 67 sector CIOCOP and the 570 item level. We have truncated the theoretical Calvo distributions at 44 months.

The first observation is that the level of aggregation influences the shape of the aggregate distribution. The distributions all have a "hump" shape, which peaks at 2 months (COICOP 67 and Item 570) , 3 months (COICOP 11), and 4 months(AC). They all have far too few one-month shares and of course miss the 12 month spike, as pointed out by Alvarez and Burriel (2010). However, from month 8, CIOCOP 67 and Item 570 both track the true distribution fairly well (except for month 12). AC and COICOP 11 both

| | Mean DAF | Median DAF |
|---|---|---|
| True | 10.9 | 7.8 |
| 11c | 9.1 | 6.8 |
| 67c | 10.5 | 7.2 |
| 570c | 10.8 | 7.8 |
| ac | 8.3 | 6.5 |

Table 3: Mean and median durations of Calvo distributions
*"True" denotes the actual cross-sectional distribution implied by hazard function, "ac","11c", "67c", and "570c" denote the cross-sectional distributions derived from the "Calvo distribution" at different aggregate level, corresponding to one aggregate sector, 11 COICOP categories,67 disaggregated COICOP categories,and 570 items respectively. "Mean DAF" and "Median DAF" denote the mean and median length of duration across firms, and both of them are in unit of month.*

overestimate the share of durations between 3 months and 16 months, and they underestimate the share of durations longer than 16 months. However, COICOP 11 is relatively closer to the true distribution than AC. If we look at the Item 570 Calvo distribution, the most disaggregated one we have, this peaks at month 2 and is the only Calvo distribution to be roughly close to (a little less than) the true proportion of 1 months. Furthermore, the Calvo distribution generated by Item 570 is quite similar from the months 2 onwards, only missing the 12 months spike. The UK distribution has a fatter tail, but the Calvo tails are certainly quite substantial. Essentially, the Calvo distributions put too much weight on the shorter months (2-7) and hence puts less weight on the remaining durations.[11]

The means and medians of the different Calvo distributions are listed in Table 3. Here we can see that as the category becomes more disaggregated, the mean and median of Calvo distributions become closer to the true distribution. Indeed, the Calvo distribution generated by Item 570 almost has the same mean and median value as what we get from the true distribution. If we compare the means of the Calvo distributions, these are linked to the distributions in Table 1, since the mean of the minimum duration distribution is

$$\bar{T}^{\min} = \sum_{k=1}^{N} c_k h_k^{-1}$$

[11]We then look at the sectoral data to see what the sectoral distributions actually look like. Most of these do not fit the Calvo implied distribution. We use a Komogorov-Smirnov test to confirm this finding. See appendix 2 for the details.

which yields the theoretical relation $\bar{T}^C = 2.\bar{T}^{\min} - 1$. Hence the Calvo means are similar to the actual mean, whilst the theoretical minimum is just over half. No such exact relation holds for the medians. If we look at Tables 2 and 3, we can see that the Calvo means are less than the theoretical means. This is because we have truncated the Calvo distributions at 44 months: if we extend this then the mean will approach its theoretical value. Truncation reduces the mean quite significantly, since the long tail of the Calvo distribution will be allocated to the 44th month: whilst the shares of these longer durations are small, they are long and so affect the mean.

Whilst the Calvo assumption gives us a mean that is about right, it differs considerably from the true distribution. There are not enough one-period price-spells: in the data, there are a lot of products that have perfectly flexible prices that change almost every month (petrol, vegetables etc.) Hence the first period hazard rate needs to be higher than in the standard Calvo model. Second there are too many 3-8 month spells. This implies that the hazard needs to be lower for these months. Then of course there is a 12 month peak.

In fact we can get a very good approximation to the monthly DAF if we do the following "simplified" hazard function:

- allow for a high and declining hazard for months 1-4.($h_1 = 0.41, h_2 = 0.29$, $h_3 = h_4 = 0.22$)
- When the length of duration less than a year , the hazard is constant at $h_i = 0.16$ for $5 \leq i \leq 11$.
- There is a 12 month peak: $h_{12} = 0.2$.
- Thereafter, we the hazard is constant at $h_i = 0.125$ or $i > 12$.

We can depict the "true" hazard and the "simplified" hazard in Figure 4. And accordingly, we can have the true cross-sectional distribution (DAF) and the one implied from the "simplified" hazard in Figure 5.

We can also look at the distributions within each of the 11 COICOP sectors. This we do in Appendix 3. There is considerable heterogeneity in the sectoral distributions. Whilst most have a 12 month peak, there are several sectors which have little if any 12 month peak in the DAF or hazard: these latter include Food and non-alcoholic beverages, alcoholic beverages, clothing and footwear, communication. Also, there are sectors which have peaks other than 12 months which are important: housing and utilities has peaks every 4 months, communications at 5-6, 11 and 18 months. We compare each sectoral distribution with the corresponding Calvo distribution. Since

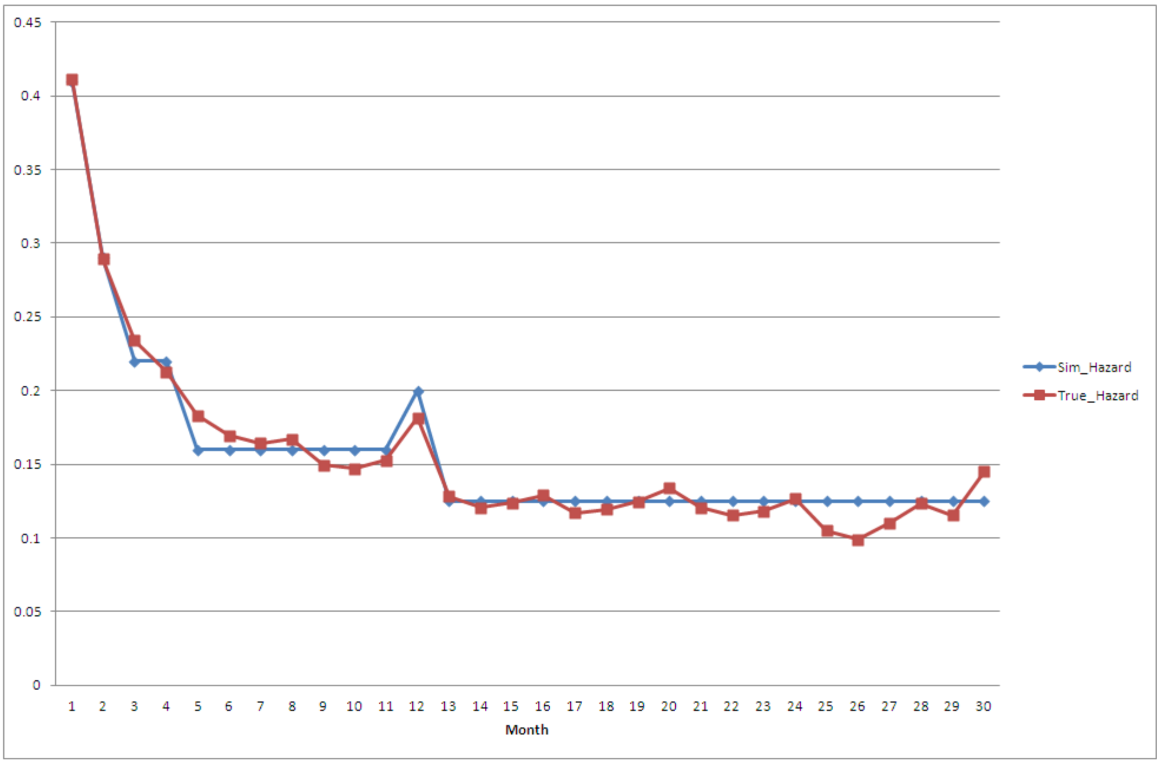



Figure 4: True and simplified hazard

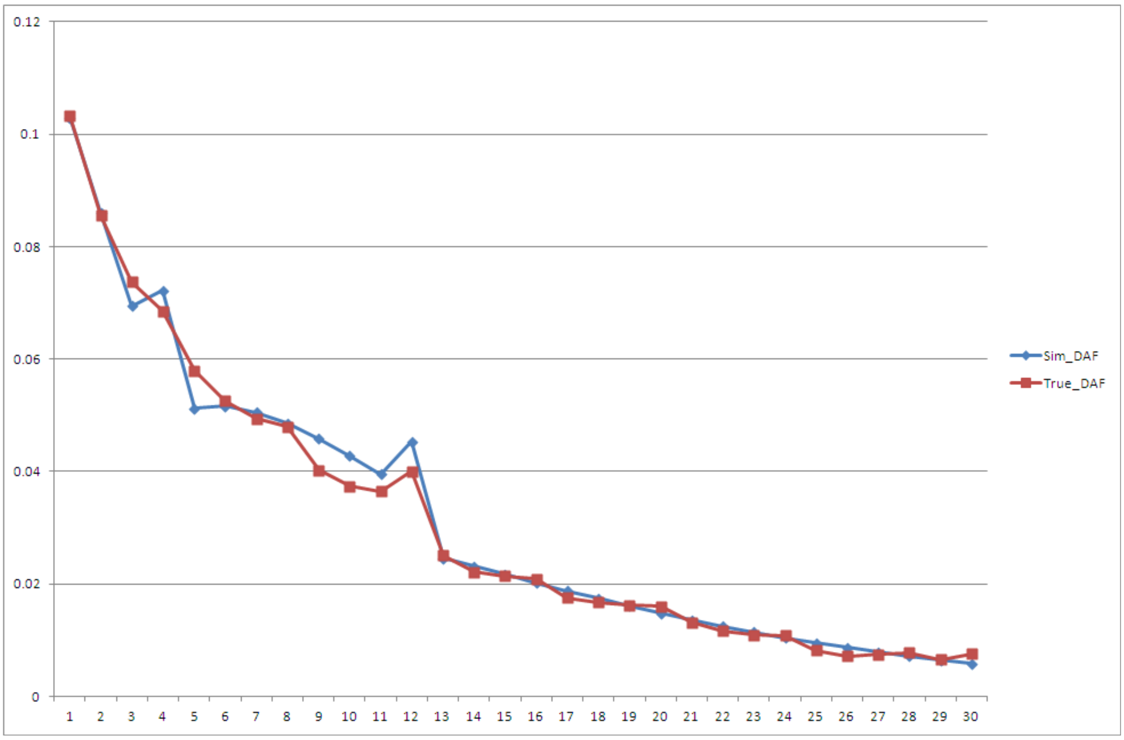



Figure 5: True and simplified DAF.

we have such a large data set, the formal Kolmogorov-Smirnov test rejects the null hypothesis that the two distributions are the same in all of the 11 sectors. We also measure the degree of "overlap" of the two distributions for each sector. That is, the extent to which the two distributions allocate the same mass to the same values: it is the sum of the absolute deviations for each value (in this case number of months) relative to the total mass. A value of 1 means that there is no overlap at all: 0 that they are identical. In the case of the distributions, the difference is as low as 0.18 for Alcoholic beverages, 0.20 for restaurants & hotels, and 0.23 for transportation. For the rest it is over 0.30 peaking at 0.58 for Health. Given that there must be some overlap here (all of the values for both distributions are strictly positive for months 1-44), these figures indicate a wide divergence in most sectors. In conclusion, we can say that the Calvo distribution is not a good description of the data either at the aggregate level or the COICOP 11 level.

# 5 The Simulation of different pricing models.

We have found that the we can use the sectoral frequencies to generate the corresponding hypothetical Calvo distributions. For the UK data at least, we find that at high levels of disaggregation, the resultant hypothetical aggregate distribution matches the true distribution quite well in terms of both the mean and the median. There are significant differences, most notably the hypothetical distribution has no 12 month spike and too few flexible prices. Since the mean and median are close, do the differences matter at the aggregate level? If we simulate a DSGE macro model using the hypothetical Calvo distribution, will it yield a good approximation to the simulations using the true distribution found in the UK data? If the answer is "yes", then it implies that the absence of the 12 month spike and too few flexible prices does not matter from the perspective of the macroeconomic properties of the DSGE model. This would validate the approach taken in Dixon and Kara (2010,2011) and Kara (2011) which used the hypothetical Calvo distribution derived from the Bils-Klenow table of sectoral frequencies in order to calibrate their US pricing models.

We will perform our simulations using two DSGE models: a simple Quantity Theory model (QT) and the Smets and Wouters (2003) (SW) model. We will look at two pricing models in both of these cases: the Generalized Taylor (GT) and Generalized Calvo (GC) model as in Dixon and Le Bihan (2012).

## 5.1 Price setting.

There are two general time-dependent models which are capable of reflecting the underlying distribution found in the micro-data: the Generalized Taylor (GT)[12] and Generalized Calvo (GC) models[13]. The key difference between the models is that in the GT the firms know how long the price spell will last when they set the price, and so each duration of price-spell will have a different reset price. In the GC, in contrast, the firms do not know how long the price spell will last and have a distribution over possible price-spells durations. All firms have the same distribution and hence there is only one reset price every period as in the simple Calvo model. In the *Generalized Taylor Economy* (*GT*) there are $N$ sectors, $i = 1, ..., N$. In sector $i$ there are $i-$period contracts: each period a cohort of $i^{-1}$ of the firms in the sector sets a new price (or wage). If we think of the economy as a continuum of firms, we can describe the $GT$ as a vector of sector shares: $\alpha_i$ is the proportion of firms that have price-spells of length $i$. If the longest observed price-spell is $F$, then we have $\sum_{i=1}^{F} \alpha_i = 1$ and $\alpha \in \Delta^{F-1}$ is the $F$-vector of shares $\boldsymbol{\alpha} = \{\alpha_i\}_{i=1}^{F}$. We can think of the "sectors" as "duration sectors": we can classify the economy by the length of price-spells. The essence of the Taylor model is that when they set the price, firms know exactly how long its price is going to last. The simple Taylor economy is a special case where there is only one length of price-spell (e.g. $\alpha_2 = 1$ is a simple Taylor "2 quarters" economy).

The log-linearised equation for the aggregate price $p_t$ is a weighted average of the sectoral prices $p_{it}$, where the weights are $\alpha_i$ :

$$p_t = \sum_{i=1}^{F} \alpha_i p_{it} \tag{9}$$

In each sector $i$, a proportion $i^{-1}$ of the $\alpha_i$ firms reset their price at each period. Assuming imperfect competition and standard demand curve, the optimal reset price in sector $i$, $x_{it}$ is given by the first-order condition of an intertemporal profit-maximisation program under the constraint implied by price rigidity. The log-linearised equation for the reset price, as in the standard Taylor set-up, is then given by :

$$x_{it} = \left( \frac{1}{\sum_{k=0}^{i-1} \beta^k} \right) \sum_{k=0}^{i-1} \beta^k E_t p^*_{t+k} \tag{10}$$

[12] See Taylor 1993, Coenen et al. 2007, Dixon and Kara 2010, 2011.

[13] See Wolman 1999, Mash 2003 and 2004, Guerrieri 2006, Sheedy 2010, Paustian and von Hagen 2008.

where $\beta$ is a discount factor, $E_t$ is the expectation operator conditional on information available at date $t$, and $p^*_{t+k}$ is the optimal flex price at time $t+k$. The reset price is thus an average over the optimal flex prices for the duration of the contract (or price-spell). The formula for the optimal flex price will depend on the model: clearly, it is a markup on marginal cost. We will specify the exact log-linearised equation for the optimal flex-price when we specify the exact macroeconomic model we use.

The sectoral price is simply the average over the $i$ cohorts in the sector:

$$p_{it} = \frac{1}{i}\sum_{k=0}^{i-1} x_{it-k} \tag{11}$$

In each period, a proportion $\bar{h}$ of firms reset their prices in this economy: proportion $\iota^{-1}$ of sector $i$ which is of size $\alpha_i$ resulting in equality (3).

In the *GC*, firms have a common set of duration-dependent reset probabilities: the probability of resetting price $i$ periods after you last reset the price is given by $h_i$. This is a time-dependent model, and the profile of reset probabilities is $\mathbf{h} = \{h_i\}_{i=1}^{F}$. Clearly, if $F$ is the longest price-spell we have $h_F = 1$ and $h_i \in [0, 1)$ for $i = 1...F-1$. Again, the duration data can be represented by the hazard function. Estimated hazard function can then be used to calibrate $\mathbf{h}$. Since any distribution of durations can be represented by the appropriate hazard function, we can choose the *GCE* to exactly fit micro-data.

In economic terms, the difference between the Calvo approach and the Taylor approach is that when the firm sets its price, it does not know how long its price is going to last. Rather, it has a *survivor function* $S(i)$ which gives the probability that its price will last at up to $i$ periods. The survivor function in discrete time is[14]:

$$\begin{aligned} S(1) &= 1 \\ S(i) &= \prod_{j=1}^{i-1}(1-h_j) \quad i = 2, ..., F \end{aligned} \tag{12}$$

Thus, when they set the price in period $t$, the firms know that they will last one period with certainty, at least 2 periods with probability $S(2)$ and so on. The Calvo model is a special case where the hazard is constant $h_i = \bar{h}$, $S(i) = (1-\bar{h})^{i-1}$ and $F = \infty$. Of course, in any actual data set, $F$ is finite.

[14]Note that the discrete time survivor function effectively assumes that all "failures" occur at the end of the period (or the start of the next period): this corresponds to the pricing models where the price is set for a whole period and can only change at the transition from one period to the next.

In the $GC$ model the reset price is common across all firms that reset their price. The optimal reset price, in the same monopolistic competition set-up as mentioned above, is given in log-linearised form by:

$$x_t = \frac{1}{\sum_{i=1}^{F} S(i)\beta^{i-1}} \sum_{i=1}^{F} S(i)\beta^{i-1} E_t p^*_{t+i-1} \tag{13}$$

The evolution of the aggregate price-level is given by:

$$p_t = \sum_{i=1}^{F} S(i) x_{t-i+1} \tag{14}$$

That is, the current price level is constituted by the surviving reset prices of the present and previous $F-1$ periods.

## 5.2 A simple quantity theory model with price-setting.

We will first examine the $GC$ and $GT$ models of prices in a quantity theory model with labour as the only input of production. This model has the great advantage being very simple, because almost all its dynamic properties are generated by the pricing models alone. DSGE models like the $SW$ model in contrast are quite complicated with dynamic properties emerging from the interaction of pricing with many other features of the model. The model we present is in its log-linearised version (see Ascari 2003, Dixon and Kara 2010 for the derivation from microeconomic foundation).

To model the demand side, we use the constant-veloticity Quantity Theory:

$$y_t = m_t - p_t$$

where $(p_t, y_t)$ are aggregate price and output and $m_t$ the money supply. We model the monetary growth process as an autoregressive process of order one $AR(1)$ :

$$\begin{aligned} m_t &= m_{t-1} + \varepsilon_t \\ \varepsilon_t &= \nu \varepsilon_{t-1} + \xi_t \end{aligned}$$

where $\xi_t$ is a white noise error term (effectively a monetary growth shock). Following Chari et al( 2000) we set $\nu = 0.5$.

The optimal flexible price $p^*_t$ at period $t$ in all sectors is given by:

$$p^*_t = p_t + \gamma y_t \tag{15}$$

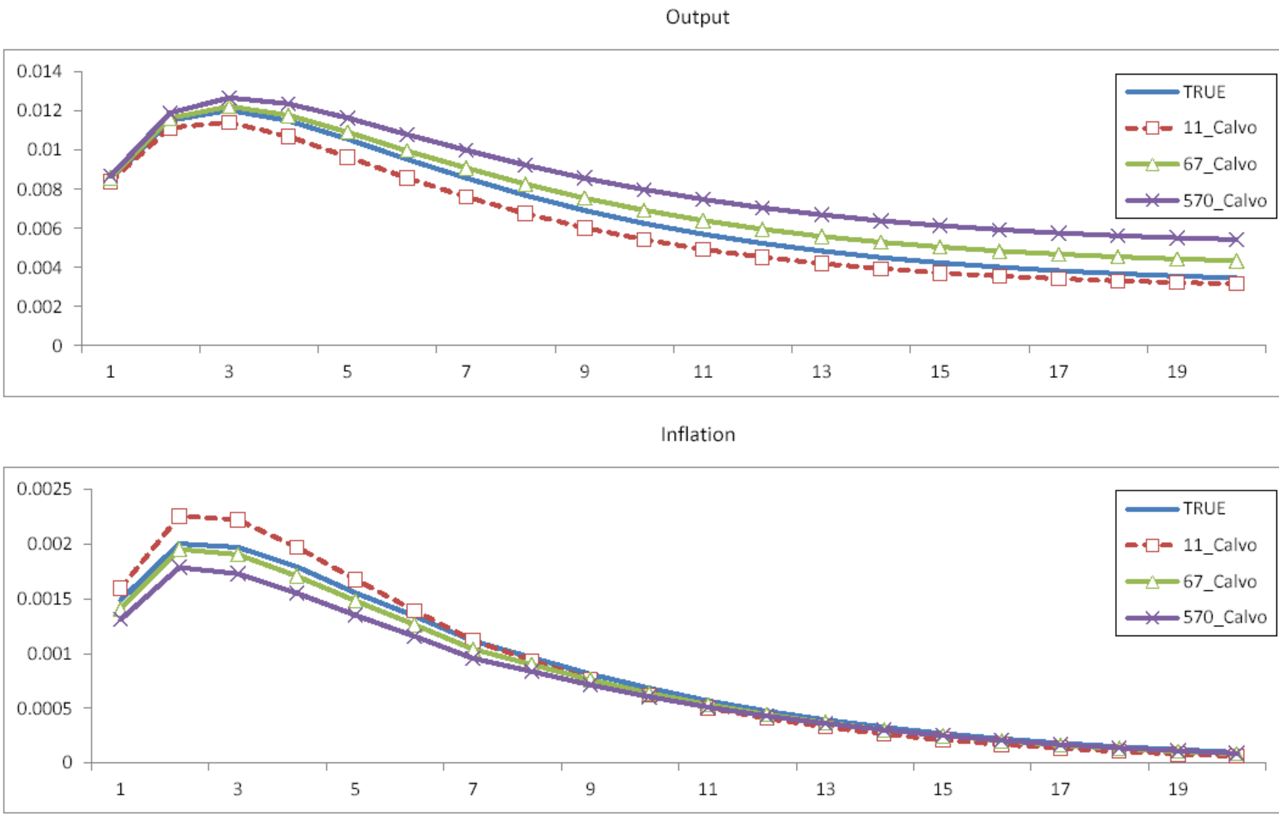



Figure 6: IRF of money growth shock in Quantity Theory model for GTE price-setting.

The key parameter $\gamma$ captures the sensitivity of the flexible price to output[15]. As discussed in Dixon and Kara (2010), there are a range of calibrated and estimated values for $\gamma$: for illustrative purposes, we use the "moderate" case of $\gamma = 0.1$ as in Mankiw and Reis (2002). As shown by Ascari (2003) and Edge (2002), the value of $\gamma$ can be interpreted as resulting from either wage or price-setting.

In Figures 6 and 7, we see the IRFs for the QT model responding to a monetary 1% monetary growth shock. As we can see, the IRFs look similar in shape for both the GC and GT: for both output and inflation. There are four IRFs reflecting the distributions generated from the estimated ("true")

[15] This can be due to increasing marginal cost and/or an upward sloping supply curve for labour. See for example Walsh (2003, chapter 5) and Woodford (2003, chapter 3).

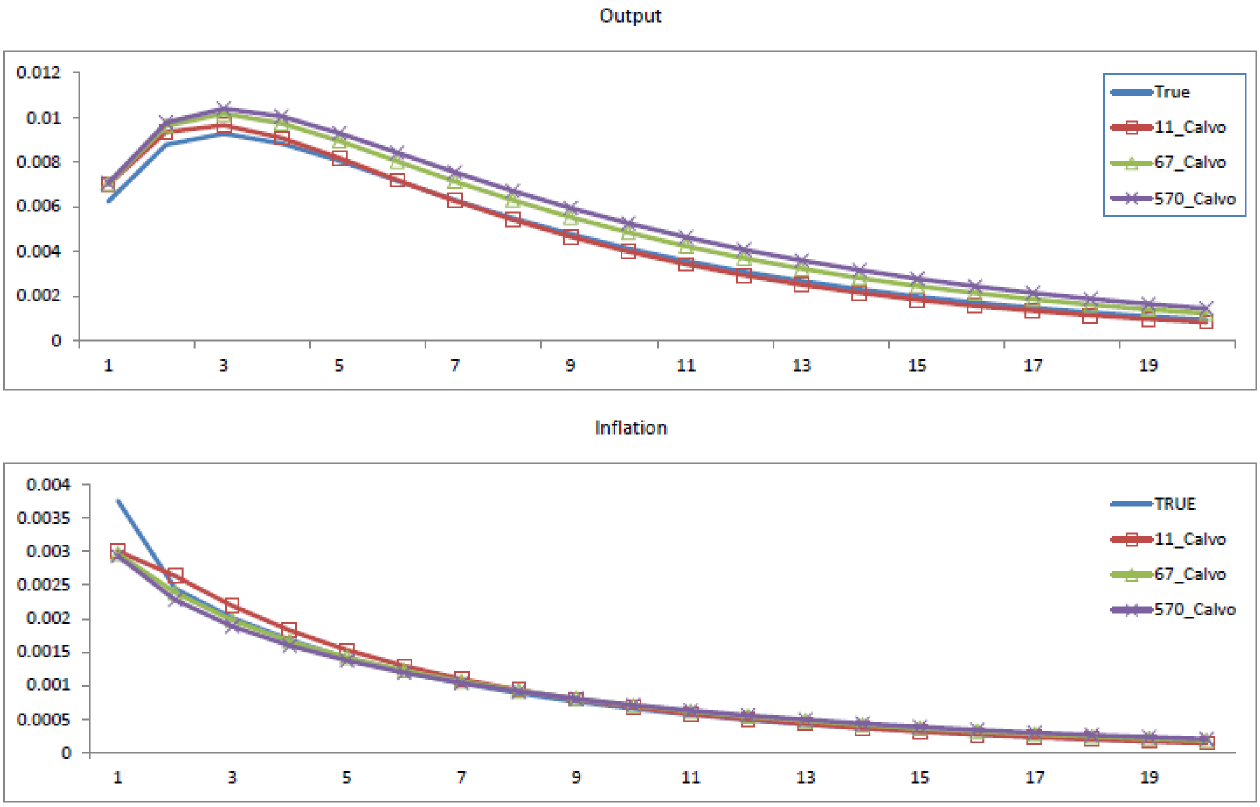



Figure 7: IRF of money growth shock in Quantity Theory model for GCE price-setting.

|  | GTE |  |  | GCE |  |  |
|---|---|---|---|---|---|---|
|  | True -11c | True -67c | True -570c | True -11c | True -67c | True -570c |
| AD in output | 9.12% | 8.73% | 21.82% | 4.34% | 14.41% | 21.42% |
| AD in inflation | 9.93% | 5.36% | 11.87% | 9.01% | 7.99% | 10.49% |

Table 4: The average relative difference AD in IRF from QT model
*"True" denotes the actual cross-sectional distribution implied by hazard function, "11c", "67c", and "570c" denote the cross-sectional distributions derived from the "Calvo distribution" at different aggregate level, corresponding to 11 COICOP categories, 67 disaggregated COICOP categories, and 570 items respectively.*

UK distribution, and the three Calvo distributions derived from the sectoral frequencies at different levels of aggregation. The results are striking. We can summarize them in a few points:

(a) the IRFs from the four distributions are similar, with the "true" IRF lying in the middle.

(b) the 570-item Calvo has the largest and most persistent effect on output, followed by the 67 COICOP Calvo, then the "true" and the 11 COICOP Calvo showing the least effect.

(c) for inflation, the 11COICOP has the biggest immediate effect, while it dies out relatively faster than the other cases. The 570-item Calvo has the smallest immediate effect, but it has the most persistent effect. The "true" and 67 COICOP Calvo lie between those two.

(d) the GT has a hump shaped reaction function for inflation, the GC does not.

In order to quantify differences between the IRFs, we define the point-by-point absolute difference as a percentage of the mean "true" IRF $\delta_i = \frac{|IRF_{true} - IRF_{calvo}|}{mean(IRF_{true})} \times 100\%$. Summing these differences over the first 20Q $\delta_i$ $(i = 1, \cdots, 20)$ and dividing by 20, we can get the average relative difference $(AD)$, which is shown in the Table 4.

Here we can see that with the GT model, the IRF generated by the COICOP 67 frequencies is the one most close to those generated from "true" distribution. However, in the GC model, the results are kind of mix. For the IRF in output, the COICOP 11 is the one has the smallest average difference. While for the IRF in inflation, the COICOP 67's performance is the best.

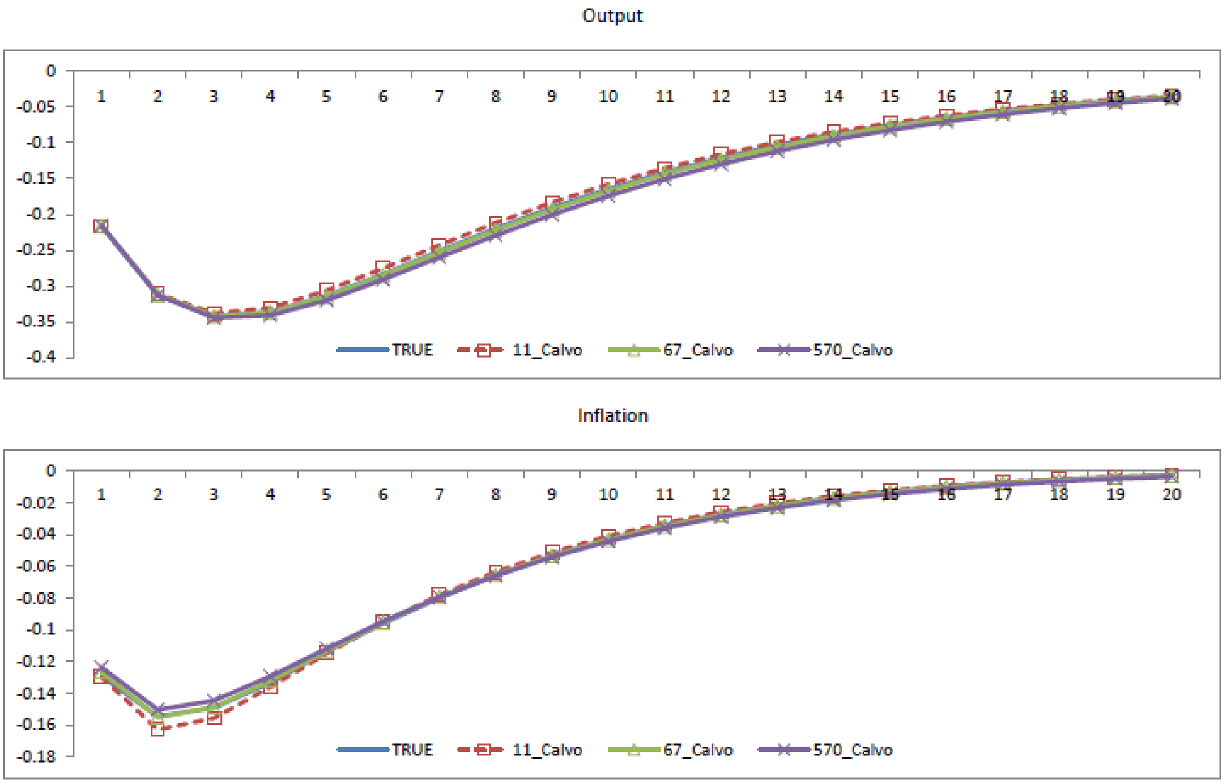



Figure 8: IRF for monetary shock in SW model with GTE price-setting.

## 5.3 A DSGE model: Smets and Wouters (2003)

In this section, we use the Smets and Wouters (2003) model of the euro area commonly employed for monetary policy analysis. The *SW* model is much more complicated than the simple QT model we have just used: there are many sources of dynamics other than prices and wages, including capital adjustment, capital utilization, consumer dynamics with habit formation, and a monetary policy reaction function. The behavior of the model is the outcome of the interaction of all of these processes together as it should be in a DSGE model. Hence the effect of pricing dynamics is not isolated as in the simple QT framework of the previous section. The details of the model and calibration are outlined in Dixon and Le Bihan (2012) using a notation consistent with this paper, and the Dynare program can be downloaded from huwdixon.org.

We depict the IRFs for an interest rate shock, which causes output and inflation to fall initially (as shown in Figures 8 and 9. Here we see that

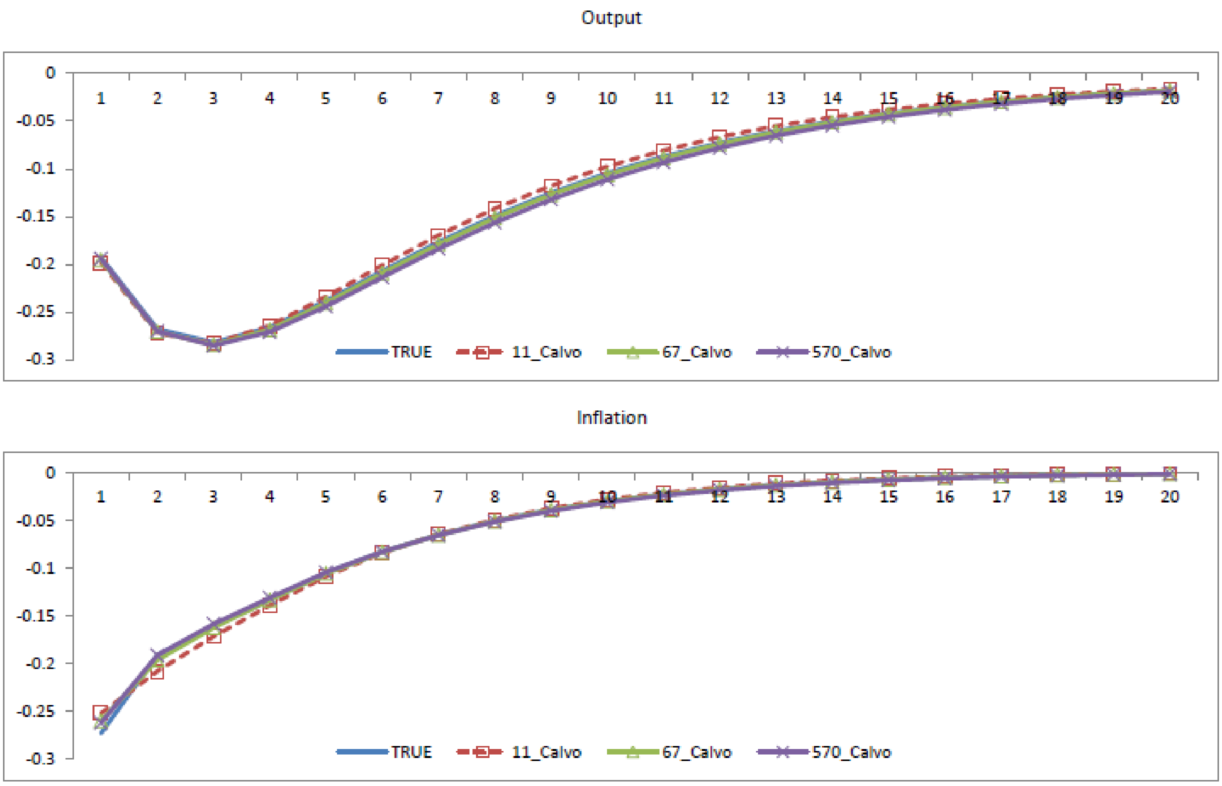



Figure 9: IRF for monetary shock in SW model with GCE price-setting.

| | GTE | | | GCE | | |
|---|---|---|---|---|---|---|
| | True -11c | True -67c | True -570c | True -11c | True -67c | True -570c |
| AD in output | 2.58% | 0.88% | 3.58% | 3.71% | 1.28% | 3.58% |
| AD in inflation | 3.08% | 0.89% | 3.27% | 6.12% | 2.06% | 2.34% |

Table 5: The average relative difference AD in IRF from SW model
*"True" denotes the actual cross-sectional distribution implied by hazard function, "11c", "67c", and "570c" denote the cross-sectional distributions derived from the "Calvo distribution" at different aggregate level, corresponding to 11 COICOP categories, 67 disaggregated COICOP categories, and 570 items respectively.*

the differences between the 4 IRFs are much smaller and less visible when compared to the QT model. This is probably because the structure on the dynamics is also determined by the rest of the model's complex dynamics, which leaves less room for the precise distribution of price-spell durations to matter. The AD are shown in the Table 5 and are quite small compared to the AD in the QT model.

If we take the results from the simulations of both the QT model and the SW model, we can see that the microeconomic differences in durations do not matter that much. We are comparing hypothetical distributions derived from the sectoral frequencies under the assumption that within each sector there is a Calvo distribution corresponding to the frequency. As we have seen, whilst the aggregate distributions implied by this may have a similar mean and median to the true distribution, the shape differs significantly and in particular there is no 12 month spike. The results of the simulations implies that these micro differences do not matter in practice. In the SW model, the differences in distribution seem to have almost no observable effect on the IRFs.

Why is it that the hypothetical Calvo distributions seem to work well despite their poor fit at the microeconomic level. We believe that there is one prime reason for this: the macroeconomic models use a *quarterly calibration*, which in effect smooths out some of the differences we observed in the monthly data. From the perspective of GTE, for example, the 12 month spike gets smoothed out. This is shown in the Figure 10, in which the"true" quarterly distribution is compared to the corresponding quarterly distributions for the "11c", "67c", "570c" distributions and "ac" the distribution

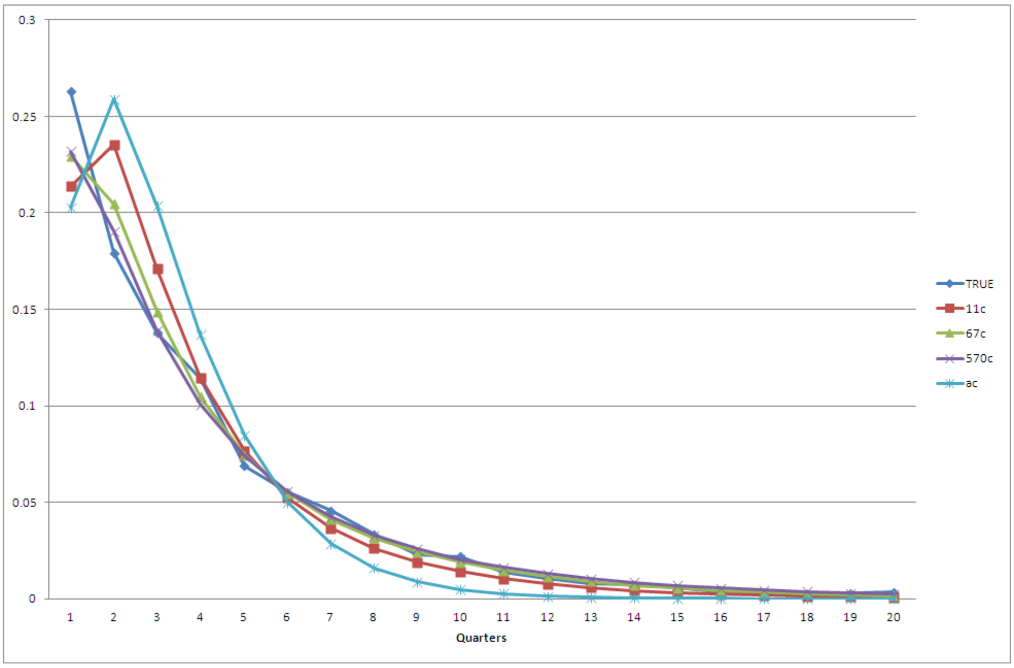



Figure 10: Quarterly distribution comparison.

implied by the single aggregate Calvo frequency. The quarterly "True" distribution is monotonic and has no 4 quarter spike. And the Calvo implied distributions at different aggregate level are all quite similar to the "True" one, except for the "ac" which has a big hump in quarter 2.

If we look at the quarterly model from the perspective of the GCE, we need to compare the quarterly hazards for the Calvo pricing at different aggregate level and the "True" one which is from an estimated hazard function. Following Dixon (2012), given a distribution across firms $\alpha \in \Delta^{F-1}$, the corresponding hazard profile that will generate this distribution in steady state is given by $\mathbf{h} \in [0,1]^{F-1}$ where:

$$h_i = \frac{\alpha_i}{i} \left( \sum_{j=i}^{F} \frac{\alpha_j}{j} \right)^{-1} \tag{16}$$

Therefore, we can calculate the monthly hazard for the Calvo pricing at different aggregate level by using the equation 16, and the relevant monthly survival rate will be obtained accordingly. This can then be converted into a quarterly hazard rate.

We plot the quarterly hazard functions in the Figure 11. The hazard functions from Calvo pricing at different aggregate levels are smoothly downward

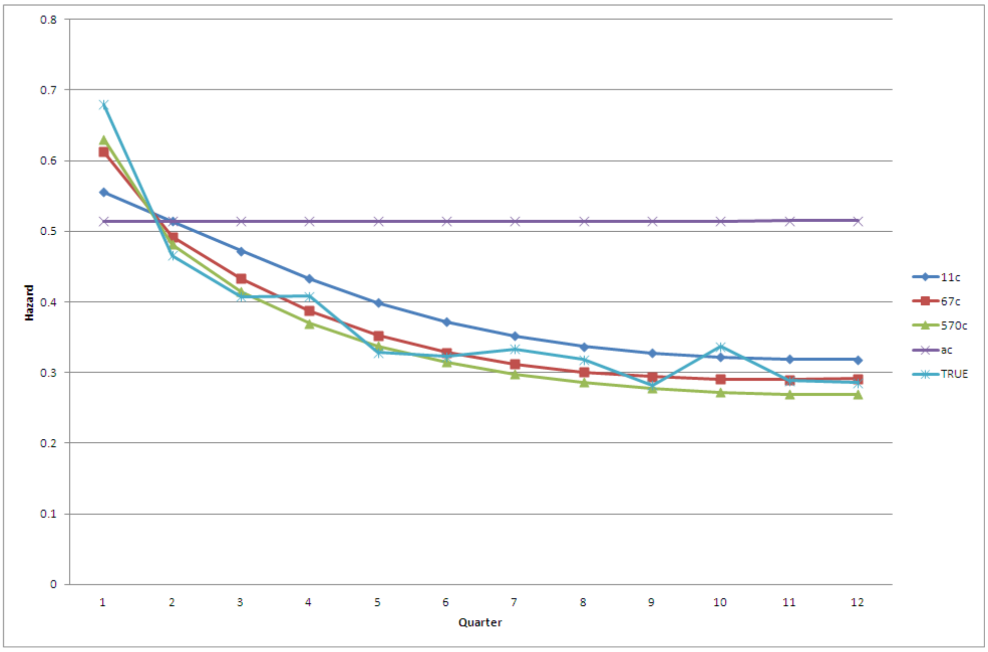



Figure 11: Quarterly hazards comparison.

sloping [16] resulting from aggregation of heterogenous price setters (see Alvarez 2008). The "true" hazard function is also generally downward sloping, with several small spikes. In general, the hazards from different disaggregate level of Calvo pricing are quite similar to the "true" hazard. The aggregate Calvo on the other hand has a constant hazard and looks quite unlike the "true" hazard.

# 6 Conclusion

In this paper we asked the question what can the sectoral data on the frequency of price-change tell us? On the theoretical level, sectoral frequencies tell us what expected duration of a price-spell is. This is of some interest, but from a macroeconomic perspective we are more interested in the behavior of the economic agents setting prices - the *cross-sectional* distribution is

[16] The Calvo implied hazards are generally downward sloping. However, after 12 quarters, these hazards become upward sloping. This is due to the truncation. For Calvo pricing, truncation means that the $\sum_{j=i}^{F} \frac{\alpha_j}{j}$ is smaller than it is under the infinite sum. The reciprocal of the smaller sum is the main reason that the hazard is biased upwards and results in an upward sloping hazard after 12 quarters.

of much more interest. Unfortunately the frequency itself says little about the cross-sectional distribution: to uncover this we need to make additional assumptions. However, we are able to say what the theoretical minimum cross-sectional mean duration is consistent with an observed frequency: it is the mean duration of a price-spell which occurs when all price-spells in the sector have the same or almost the same length. However, the cross-sectional mean can be much longer: intuitively, firms that have prices that last for a long time do not reset their prices very often. These firms contribute little to the monthly frequency but add a lot to the cross-sectional mean.

When we look at the UK data using an estimated hazard function, we find that the UK data is a long way from the distribution implied by the theoretical minimum. Looking at different levels of disaggregation, we find that whilst the minimum theoretical mean duration is around 5.5-6.7 months, the actual data reveals a mean of almost 11 months. We also find that the more disaggregated the data, the closer are both the median and mean to the true values. We also look at the aggregate data using the hypothesis that the sectoral frequencies are generated by a Calvo distribution. Under this assumption, the cross-sectional mean is much larger than the minimum, and gets closer to the actual mean as you become more disaggregated, with the most disaggregated having almost exactly the same mean and median as the data. Whilst the mean and median of the hypothetical Calvo distribution can be close to the actual values, the shape differs in two distinct ways: firstly, there is no 12 month spike, secondly there are not enough flexible prices. When we look at the COICOP 11 sectors, we find that the Calvo distribution hypothesis does not work very well: there is considerable heterogeneity across sectors in terms of the cross-sectional distribution and hazard function.

However, we do find that whilst the Calvo distribution hypothesis differs from the estimated distribution both at the aggregate and sectoral levels, it nonetheless works at the aggregate level. This is because of a combination of two factors. First, along with getting the mean and median correct, the Calvo distribution also generates a nice long fat tail. Second, DSGE models are calibrated using quarterly periods rather than monthly: when we move to quarterly data, the differences which look significant at the monthly level get averaged out to a large extent. This means that when we look at the behavior of DSGE models under the Calvo distribution hypothesis, they behave in a very similar way to models calibrated with the microdata. This suggests that we can use the disaggregated frequency data (the more disaggregated the better) to calibrate DSGE models when we do not have reliable hazard function estimates or access to the price microdata as was done by Dixon and Kara (2010, 2011).

# 7 Bibliography.


1. Alvarez, L. (2008). What do Micro Price Data Tell Us on the Validity of the New Keynesian Phillips Curve? *Economics E-Journal,* 2, 2008-19

2. Alvarez, L., and Burriel, P., (2010), " Is a Calvo Price Setting Model Consistent with Individual Price Data?," *The B.E. Journal of Macroeconomics*, Berkeley Electronic Press, 10(1), 13.

3. Alvarez, L., Dhyne, E., Hoeberichts, M., Kwapil, C., Le Bihan, H., Lunnemann, P., Martins, F., Sabbatini, R., Stahl, H., Vermeulen, P., and Vilmunen, J. (2006). "Sticky prices in the euro area: a summary of new micro-evidence", *Journal of the European Economic Association*, 4, 575-584.

4. Alvarez, L and Hernando, I (2006). Price setting behaviour in Spain: Stylised facts using consumer price micro data, *Economic Modelling,* 23, 699-716.

5. Ascari, G (2003). "Price/Wage Staggering and Persistence: A Unifying Framework," *Journal of Economic Surveys*, Wiley Blackwell, vol. 17(4), pages 511-540, 09.

6. Baharad, E., and Eden, B. (2004) "Price Rigidity and Price Dispersion: Evidence from Micro Data," *Review of Economic Dynamics*, Vol7, 613-641.

7. Baudry L, Le Bihan H, Sevestre P and Tarrieu S (2007). What do thirteen million price records have to say about consumer price rigidity? *Oxford Bulletin of Economic Statistics,* 69, 139-183.

8. Bils, M and Klenow, P (2004). Some evidence on the importance of Sticky prices, *Journal of Political Economy*, 112, 947-985.

9. Bunn P, Ellis C. (2009). Price setting behavior in the UK, *Bank of England Quarterly Bulletin*, 49, 28-36.

10. Bunn P, Ellis C. (2012). Examining The Behaviour Of Individual UK Consumer Prices. *Economic Journal*, 122, F35-F55.

11. Calvo, G. (1983) "Staggered Prices in a Utility-Maximizing Framework", *Journal of Monetary Economics,*12(3),pp.383-398

12. Carvalho, C (2006) "Heterogeneity in Price Stickiness and the Real Effects of Monetary Shocks," *Frontiers of Macroeconomics*: Vol. 2 : Iss. 1, Article 1.

13. Chari, V. V., Kehoe, P. J. and McGrattan, E. R.: 2000, Sticky price models of the business cycle: Can the contract multiplier solve the persistence problem?, *Econometrica* 68(5), 1151–79.

14. Coenen, G., Levin, A.T., and Christoffel, K., (2007), "Identifying the influences of nominal and real rigidities in aggregate price-setting behavior", *Journal of Monetary Economics*, 54, 2439-2466.

15. Dixon, H and Kara, E (2006): "How to Compare Taylor and Calvo Contracts: A Comment on Michael Kiley", *Journal of Money, Credit and Banking.,* 38, 1119-1126.

16. Dixon H, Kara E (2010): Can we explain inflation persistence in a way that is consistent with the micro-evidence on nominal rigidity,. *Journal of Money, Credit and Banking*, 42, 151-170.

17. Dixon H, Kara E (2011): Contract Length Heterogeneity and the Persistence of Monetary Shocks in a Dynamic Generalized Taylor Economy, *European Economic Review*, 55, 280-292.

18. Dixon H. (2012), A unified framework for using micro-data to compare dynamic time-dependent price-setting models, *BE Journal of Macroeconomics* (Contributions), volume 12.

19. Dixon H and Le Bihan H (2012), Generalized Taylor and Generalized Calvo price and wage-setting: micro evidence with macro implications, *Economic Journal* (forthcoming, May).

20. Edge, R.: 2002, The equivalence of wage and price staggering in monetary business cycle models, *Review of Economic Dynamics* 5, 559–585.

21. Fougere, D., Le Bihan, H., and Sevestre, P. (2007), "Heterogeneity in Price Stickiness: A Microeconometric Investigation," *Journal of Business & Economic Statistics, 25(3), 247-264.*

22. Guerrieri, L., (2006), "The Inflation Persistence of Staggered Contracts," *Journal of Money, Credit and Banking*, 38(2), 484-494.

23. Kara E (2010): Optimal Monetary Policy in the Generalized Taylor Economy, *Journal of Economic Dynamics and Control,* 34, 2023-2037.

24. Klenow P and Kryvtsov O (2008). State-Dependent vs. Time-Dependent Pricing: Does it Matter for Recent U.S. Inflation? *The Quarterly Journal of Economics,* 123, August 2008, 863-904.

25. Kolmogorov, A. (1933) "Sulla determinazione empirica di una legge di distribuzione" G. Inst. Ital. Attuari, 4, 83

26. Mankiw, G. and Reis, R (2002), "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117, 1295-1328.

27. Mash, R. (2003) "New Keynesian Microfoundations Revisited: A Calvo-Taylor-Rule-of-Thumb Model and Optimal Monetary Policy Delegation", *Oxford University Discussion Paper 174.*

28. Mash, R. (2004), "Optimising Microfoundations for Inflation Persistence", *Oxford University Discussion Paper 183.*

29. Matusoka T (2009) "Goodness-of-fit test for price duration distributions", Mimeo

30. Nakamura E and Steinesson J, 2008. "Five Facts about Prices: A Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics,* MIT Press, vol. 123(4), pages 1415-1464,November.

31. Paustian, M., and von Hagen, J., (2008). "The Role of Contracting Schemes for Assessing the Welfare Costs of Nominal rigidities," CFPR Discussion Papers 6726.

32. Sheedy, K., (2010). "Intrinsic inflation persistence", *Journal of Monetary Economics*, 57(8), 1049-1061.

33. Smets, F. and Wouters, R.: 2003, An estimated dynamic stochastic general equilibrium model of the euro area, *Journal of the European Economic Association* 1(5), 1123–1175.

34. Smirnov, N.V. (1948) "Tables for estimating the goodness of fit of empirical distributions", *Annals of Mathematical Statistic*, 19, 279.

35. Taylor, J. B., (1993), *Macroeconomic Policy in a World Economy*, Norton.

36. Walsh, C.:2003, Monetary Theory and Policy, 2nd Edition, *MIT Press.*

37. Wolman, A., (1999), "Sticky prices, marginal cost and the behaviour of inflation", *Federal reserve bank of Richmond quarterly,* 85, 29-47.

38. Woodford, M.: 2003, Interest and prices: Foundations of a theory of monetary policy, *Princeton University Press*, Princeton, NJ .

# 8 Appendix:

## 8.1 Proofs.

**Proof of Proposition 1..** Firstly we will prove (a) and (b). We do this by contradiction. Let us suppose that the solution such that $\alpha_k > 0$ and $\alpha_j > 0$ and $k - j \geq 2$ We will then show that there is another feasible $\boldsymbol{\alpha}' \in H(\bar{h})$ with $\alpha_j > 0$ and $\alpha_{j+1} > 0$ which generates a shorter average contract length.

For the proposed solution $\boldsymbol{\alpha}$, the two sectors $k$ and $j$ have sector shares satisfying:

$$\begin{aligned} \alpha_k + \alpha_j &= \rho = 1 - \sum_{i=1, i \neq j,k}^{F} \alpha_i \\ \frac{\alpha_k}{k} + \frac{\alpha_j}{j} &= \eta = \bar{h} - \sum_{i=1, i \neq j,k}^{F} \frac{\alpha_i}{i} \end{aligned} \tag{17}$$

$\rho$ is the total share of the two sectors: $\eta$ is the contribution of these two sectors to $\bar{h}$. Since $k > j$, $\rho > \eta j$, hence we can rewrite (17) as

$$\begin{aligned} \alpha_j &= \frac{kj}{k-j}\eta - \frac{j}{k-j}\rho \\ \alpha_k &= -\frac{kj}{k-j}\eta + \frac{k}{k-j}\rho \end{aligned} \tag{18}$$

Hence we can choose $\left(\alpha'_j, \alpha'_{j+1}\right)$ which satisfies satisfies (18), but yields a lower average contract length:

$$\begin{aligned} \alpha'_j &= j(j+1)\eta - j\rho \\ \alpha'_{j+1} - \alpha_{j+1} &= (j+1)\rho - j(j+1)\eta \end{aligned} \tag{19}$$

Define $\Delta\alpha_{j+1} = \alpha'_{j+1} - \alpha_{j+1}$. What we are doing is redistributing the total proportion $\rho$ over durations $j$ and $j+1$ so that the aggregate proportion of firms resetting the price is the same: $\boldsymbol{\alpha}'\in H\left(\bar{h}\right)$ ,since (19) is equivalent to (18) implies

$$\begin{aligned} \Delta\alpha_{j+1} + \alpha'_j &= \rho \\ \frac{\Delta\alpha_{j+1}}{k} + \frac{\alpha'_j}{j} &= \eta \end{aligned} \tag{20}$$

$\boldsymbol{\alpha}'$ has a lower average contract length. Since we leave the proportions of other durations constant, their contribution to the average contract length is unchanged. From (18) the contribution of durations $k$ and $j$ is given by

$$T_k = ka_k + j\alpha_j = \rho\,(k+j) - kj\eta$$

Likewise the contribution with $\alpha'$ is given by

$$T_j = (j+1)\,\Delta\alpha_{j+1} + j\alpha'_j = \rho\,(2j+1) - (j+1)\,j\eta$$

Hence (noting that $\rho > \eta j$):

$$\begin{aligned} T_k - T_{j+1} &= \rho\,(k+j-2j-1) - \eta\,(kj - (j+1)\,j) \\ &> \eta\,[j\,(k-j-1) - kj + (j+1)\,j] = 0 \end{aligned}$$

That is $\bar{T}\,(\boldsymbol{\alpha}) - \bar{T}\,(\boldsymbol{\alpha}') = T_k - T_{j+1} > 0$ the desired contradiction. To prove (c) for sufficiency, if $\bar{h}^{-1} = k \in Z_+$, then $\alpha_k = 1 \in H\left(\bar{h}\right)$. If $\alpha_k < 1$ any other element of $H\left(\bar{h}\right)$ must involve strictly positive $\alpha_j$ and $\alpha_i$ with $j - i \geq 2$, which contradicts the parts $(a)$ and $(b)$ of the proposition already established. For necessity, note that if $\bar{h}^{-1} \notin Z_+$, then no solution with only one contract length can yield the observed proportion of firms resetting prices. ■

**Proof or Proposition 2.** Assume the contrary, that there is a distribution $\boldsymbol{\alpha}$ with $\alpha_i > 0$ where $1 < i < F$ which solves (5). Redistribute the weight on sector $i$ between $\{1, F\}$ in order to ensure that we remain in $H\left(\bar{h}\right)$ so that:

$$\Delta\alpha_1 + \frac{\Delta\alpha_F}{F} = \frac{\alpha_i}{i}; \Delta\alpha_1 + \Delta\alpha_F = \alpha_i$$

which implies:

$$\Delta\alpha_F = \alpha_i \frac{F\,(i-1)}{i\,(F-1)}; \Delta\alpha_1 = \alpha_i \frac{F-i}{i\,(F-1)}$$

Hence:

$$\begin{aligned} \Delta\bar{T} &= \alpha_i \left[\frac{F\,(i-1)}{i\,(F-1)}\,(F-i) - \frac{F-i}{i\,(F-1)}\,(i-1)\right] \\ &= \frac{\alpha_i\,(i-1)\,(F-1)}{i\,(F-1)}\,[F-1] > 0 \end{aligned}$$

The desired contradiction. Given that all contracts must be either 1 or $F$ periods long, the rest of the proposition follows by simple algebra. ■

## 8.2 Data description

The data is described in some detail by Bunn and Ellis (2009, 2012) so our description will be brief. The ONS collect a longitudinal micro data set of monthly price quotes from over ten thousands of outlets to compute the national index of consumer prices. There are two basic price collection methods: local and central. Local collection is used for most items. The UK was divided into its standard regions (eg Wales, East Midlands etc) and a number of locations are random selected in each region according to the total expenditure for the region. There are about 150 locations around the UK, and around 120,000 quotations are obtained each month by local collection. For some items, collection in individual shops across the 150 locations is not required- for example, for larger chain stores who have a national pricing policy or where the price is the same for all UK residents or the regional variation in prices can be collected centrally. The data that we were able to access for this study via the VML at Newport (Wales) consists of the locally collected data covering about two thirds of total CPI (centrally collected data covers about 33% of CPI). The sample spans over the time period from January 1996 to December 2007 and contains between 112,676 (1996) and 99,524 (2007) elementary price quotations per month, with a resulting dataset of around 14 million price observations. The coverage and classification of the CPI indices are based on the international classification system for household consumption expenditures known as COICOP (classification of individual consumption by purpose). This is a hierarchical classification system comprising: *divisions* e.g. 01 Food and non-alcoholic beverages, *groups* e.g. 01.1 Food, and *classes* ( the lowest published level) e.g. 01.1.1 Bread and cereals. As table 6 shows, the division Food and non-alcoholic beverages accounts for about 17% of the CPI weight in the subsample available in the dataset. Education is not contained in the VML dataset, as these prices are all collected centrally: but all other CPI divisions have locally collected observations and are included in the dataset.

In our CPI research data set, each individual price quote consists of information on the item code, the outlet, the region, the date etc. The product category at the elementary level is defined as an item - for example large loaf, white, unsliced (800g). However, the data has been anonymized with respect to the variety and brand of the product. With the information on the item $i$, the shop $j$, the location $k$, and the date $t$, we can construct a price trajectory $P_{ijk,t}$, which is sequence of price quotes for a specific item belonging to a product category in a specific shop over time. Specifically, we take two sequential price quotes belong to the same price trajectory if the they have the same product identity, location and shop code. There are

| COICOP division | Freq. | Percent | Cum. |
|---|---|---|---|
| Food and Non-Alcoholic Beverages | 25,191.51 | 17.62 | 17.62 |
| Alcoholic Beverages and Tobacco | 10,083.28 | 7.05 | 24.67 |
| Clothing and Footwear | 13,323.33 | 9.32 | 33.98 |
| Housing and Utilities | 9,350.23 | 6.54 | 40.52 |
| Furniture and Home Maintenance | 16,211.75 | 11.34 | 51.86 |
| Health | 2,705.55 | 1.89 | 53.75 |
| Transport | 14,800.15 | 10.35 | 64.1 |
| Communications | 237.2797 | 0.17 | 64.27 |
| Recreation and Culture | 14,085.51 | 9.85 | 74.12 |
| Education | 6.340364 | 0 | 74.12 |
| Restaurants and Hotels | 25,087.06 | 17.54 | 91.67 |
| Miscellaneous Goods and Services | 11,918.02 | 8.33 | 100 |
| Total | 143,000 | 100 | |

Table 6: CPI share in COICOP sectors

about $614,000$ price trajectories. And the average length of each price trajectory is about 24 months. Each trajectory will consist of a sequence of one or more price-spells: there are 3,174,692 price-spells in the data (i.e. on average about 5 price-spells per trajectory).

## 8.3 Comparing sectoral distribution of DAF with distribution of Calvo

This section examines the distribution of the duration of price spells at the COICOP 11 sectoral level. As in the aggregate data, we estimate the sectoral hazard functions using the KM non-parametric method and the corresponding cross-sectional DAF. The exact method is the same as Dixon and Le Bihan (2012) and is described in detail in Dixon (2012): in brief, we estimate the non-parametric KM estimator excluding all left-censored data and treating right-censoring as the end of a price-spell. We then compare these with the sectoral Calvo distributions in three ways. First, and most straightforwardly, we comparing the distribution of DAF with distribution of Calvo in each of 11 COICOP sectors with the "eye ball test". Even though this is not a strict statistical test, it can give us an impression how close or how far the two kind of distributions are different. Second, we examine the distributional assumption using Kolmogorov-Smirnov test. The Kolmogorov-Smirnov statistic quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the *reference distribution* (in our case, the reference distribution is Calvo distribution). Since we have such a large sample of price spells, this is a very strict test: standard errors are very small so that even small deviations of the reference distribution will lead to its rejection. Third, we propose a method to calculate the relative absolute difference. Instead of calculating the difference with the cumulative distribution function like Kolmogrov-Smirnov test, we compute the sum of the absolute differences between the two probability density distributions relative to their combined mass of 2. This method can provide us with a measure of how much the two distributions differ: this is a descriptive statistic rather than a test, but captures the extent of "overlap".

### 8.3.1 Distribution graphs: DAF vs. Calvo

As the distribution graphs show, there are significant differences between the distribution of DAF and the distribution of Calvo in most COICOP sectors. First, the Calvo distribution has less one-period price-spells than DAF, indicating that Calvo price setting mechanism fails to replicate the evidence of large volumn of flexible price setters. Second, there is no 12-month spike in the Calvo distribution. But this 12-month spike does appear in all COICOP sectors, it is absent in health, housing, culture, etc. Third, in most sectors the DAF has a fatter tail compared to the Calvo distribution. The comparisons between the actual cross-sectional distribution with the Calvo distribution in each COICOP sector can be shown as Figure 12-Figure 14.

### 8.3.2 Kolmogorov-Smirnov test

The Kolmogorov-Smirnov statistic for a given cumulative distribution function $F(x)$ is

$$D_n = \sup_x |F_n(x) - F(x)|$$

where $\sup x$ is the supremum of the set of distances. By the Glivenko-Cantelli theorem, if the sample comes from distribution $F(x)$, then $D_n$ converges to 0 almost surely. Kolmogorov strengthened this result, by effectively providing the rate of this convergence. In practice, the statistic requires relatively large number of data to properly reject the null hypothesis.

Under null hypothesis that the sample comes from the hypothesized distribution $F(x)$,

$$\sqrt{n} D_n \overset{n\to\infty}{\to} \sup_t |B(F(t))|$$

in distribution, where $B(t)$ is the Brownian bridge. The Kolmogorov-Smirnov test is constructed by using the critical values of the Kolmogorov distribution.

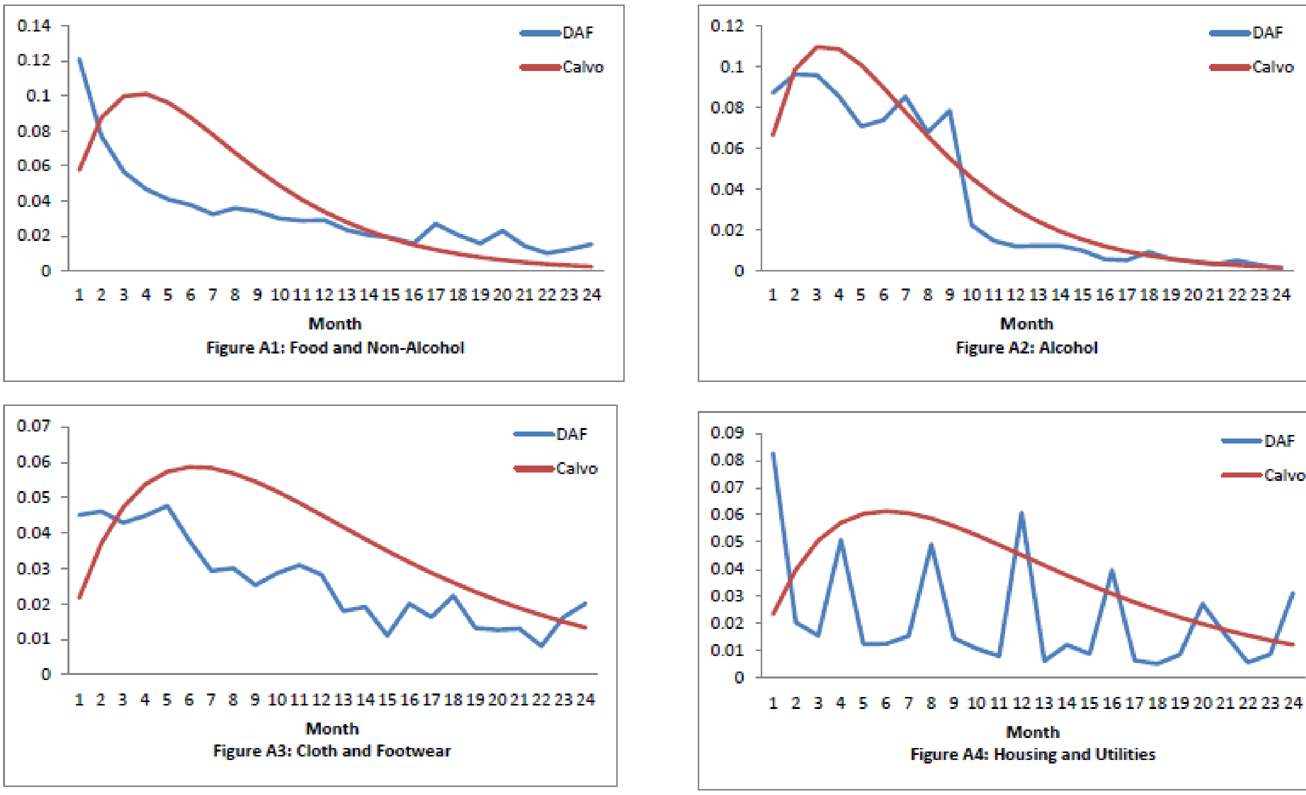



Figure 12: DAF vs. Calvo in each COICOP sector

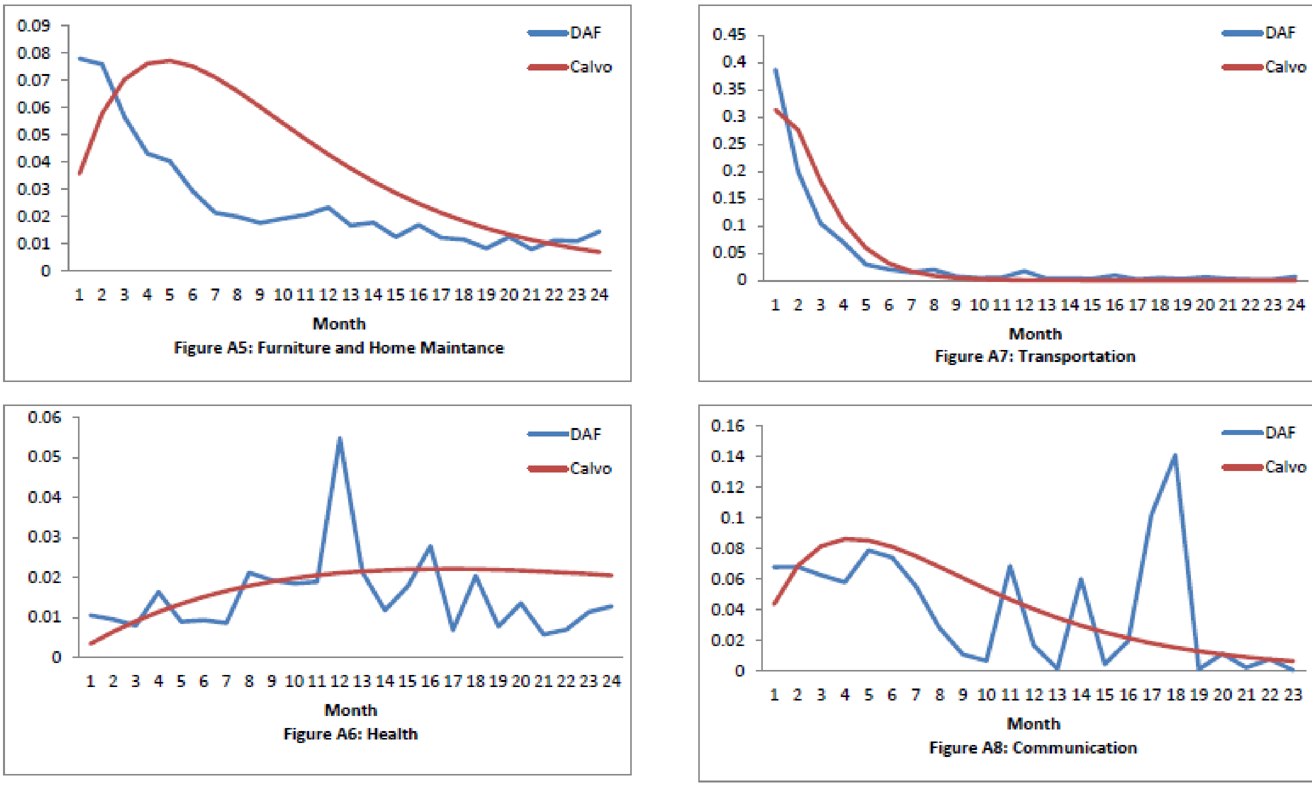
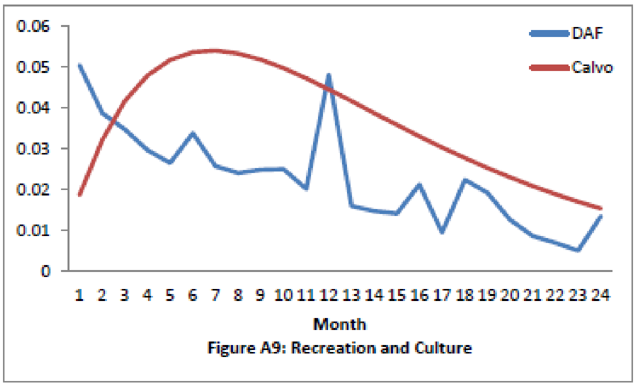
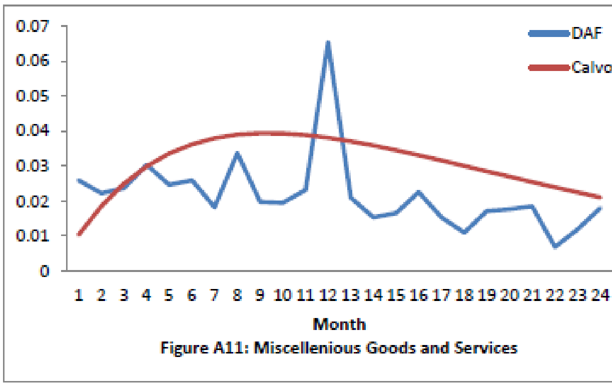



Figure 13: DAF vs. Calvo in each COICOP sector (cont).



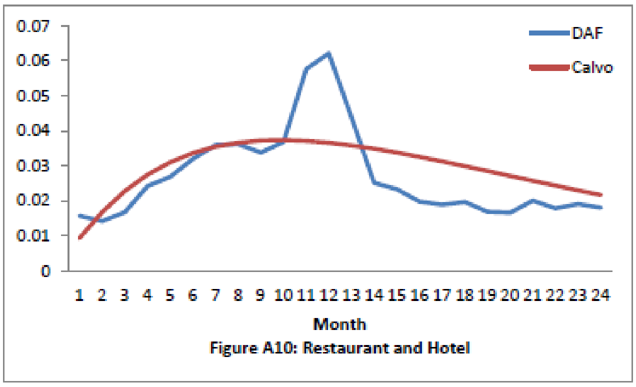




Figure 14: DAF vs. Calvo in each COICOP sector. (cont).

| COICOP Sectors | Distance | K-alpha/sqrt(n) | Test result |
|---|---|---|---|
| Food & Beverages | 0.293 | 0.009 | Reject null |
| Alcoholic & Tobacco | 0.130 | 0.060 | Reject null |
| Clothing & Footwear | 0.281 | 0.060 | Reject null |
| Housing & Utilities | 0.401 | 0.049 | Reject null |
| Furniture & Home | 0.377 | 0.013 | Reject null |
| Health | 0.399 | 0.078 | Reject null |
| Transport | 0.161 | 0.111 | Reject null |
| Communications | 0.227 | 0.136 | Reject null |
| Recreation & Culture | 0.337 | 0.090 | Reject null |
| Restaurants & Hotels | 0.105 | 0.102 | Reject null |
| Miscellaneous | 0.295 | 0.082 | Reject null |

Table 7: Kolmogorov-Smirnov test results.

The null hypothesis is rejected at level $\alpha$ if

$$\sqrt{n}D_n > K_\alpha,$$

where $K_\alpha$ is found from

$$\Pr\left(K \leq K_\alpha\right) = 1 - \alpha.$$

As described in Table 7, the test results reject that sectoral distribution of DAF is the same as sectoral Calvo distribution. This result is consistent with the finding in Matsuoka (2009), who found that over 90 percent of the 429 tested items in the Japanese retail price data for 2000-2005 reject the hypothesis that the underlying distribution is exponential, which corresponds to the time-dependent pricing model of Calvo.

#### 8.3.3 Relative difference: DAF vs. Calvo

We propose a new method to calculate the relative absolute difference between the DAF and the Calvo distribution. We calculate and absolute difference point-by-point for the probability density function between two distributions:

$$d_i = |f_{DAF} - f_{Calvo}|$$

and then we add up all the $d_i$ and let it be divided by 2. We denote the result as Relative Difference $RD$ :

$$RD = \frac{\sum_{i=1}^{T} d_i}{2}$$

| COICOP | Relative Difference (DAF vs. Calvo) |
|---|---|
| Food and No-Alcoholic | 0.36 |
| Alcoholic and Tobacco | 0.18 |
| Clothing and footwear | 0.33 |
| Housing and utilities | 0.49 |
| Furniture and home maintenance | 0.44 |
| Health | 0.58 |
| Transportation | 0.23 |
| Communication | 0.35 |
| Recreation and culture | 0.41 |
| Restaurant and hotel | 0.20 |
| Miscellaneous goods and services | 0.37 |

Table 8: Relative difference between True and Calvo in each COICOP sector

If the two distributions are identical, the $RD = 0$. If the two distributions do not overlap at all (these is no value for which both pdfs are strictly positive), then $RD = 1$. The results of *relative difference* are shown in Table 8. As can be seen, the $RD$ between the two distributions varies: Whilst some sectors show a low level of $RD$ (Alcohol and Tobacco) others have a much higher level (Housing and Utilities, Health).

We also calculate the relative difference for the aggregate distribution of durations (as shown in Table 9) for the different levels of disaggregation. Indeed, the aggregate differences are much less than those at the sectoral levels. The 67 and 570 Calvo distributions are quite similar to the True distribution. The relative differences become even smaller when we look at the quarterly data, suggesting that the differences are averaged out to some extent.

| | Relative Differences (True vs. Calvo) | |
|---|---|---|
| | Monthly data | Quarterly data |
| ac | 0.20 | 0.16 |
| 11c | 0.12 | 0.09 |
| 67c | 0.07 | 0.05 |
| 570c | 0.06 | 0.04 |

Table 9: Relative differences between the True and Calvo distributions
*"True" denotes the actual cross-sectional distribution implied by hazard function, "ac", "11c", "67c", and "570c" denote the cross-sectional distributions derived from the "Calvo distribution" at different aggregate level, corresponding to one aggregate sector, 11 COICOP categories, 67 disaggregated COICOP categories, and 570 items respectively.*