EUROPEAN CENTRAL BANK

WORKING PAPER SERIES

NO. 511 / AUGUST 2005

EUROSYSTEM INFLATION PERSISTENCE NETWORK

TIME OR STATE DEPENDENT PRICE SETTING RULES?

EVIDENCE FROM PORTUGUESE MICRO DATA

by Daniel A. Dias
Carlos Robalo Marques
and João M. C. Santos Silva

EUROPEAN CENTRAL BANK

WORKING PAPER SERIES

NO. 511 / AUGUST 2005

EUROSYSTEM INFLATION PERSISTENCE NETWORK

TIME OR STATE DEPENDENT PRICE SETTING RULES?

EVIDENCE FROM PORTUGUESE MICRO DATA [1]

by Daniel A. Dias, [2]
Carlos Robalo Marques [2]
and João M. C. Santos Silva [3]

In 2005 all ECB publications will feature a motif taken from the €50 banknote.

1 We are indebted to Stephen Cecchetti, Jordi Gali, Andrew Levin and participants in the Eurosystem "Inflation Persistence Network" for inspiring discussions and many useful suggestions. We also thank Pedro Duarte Neves and Pedro Portugal for helpful discussions and comments. The usual disclaimer applies. João Santos Silva is thankful for the hospitality, working conditions and financial support provided by Banco de Portugal and gratefully acknowledges the partial financial support from Fundação para a Ciência e Tecnologia, program POCTI, partially funded by FEDER.

2 Banco de Portugal, Rua Francisco Ribeiro, 2 1150-165 Lisboa, Portugal; e-mail: dbdias@bportugal.pt
e-mail: cmrmarques@bportugal.pt

3 ISEG/Universidade Técnica de Lisboa, Rua do Quelhas, 6 1200 Lisboa, Portugal; e-mail: jmcss@iseg.utl.pt

The Eurosystem Inflation Persistence Network

This paper reflects research conducted within the Inflation Persistence Network (IPN), a team of Eurosystem economists undertaking joint research on inflation persistence in the euro area and in its member countries. The research of the IPN combines theoretical and empirical analyses using three data sources: individual consumer and producer prices; surveys on firms' price-setting practices; aggregated sectoral, national and area-wide price indices. Patterns, causes and policy implications of inflation persistence are addressed.

Since June 2005 the IPN is chaired by Frank Smets; Stephen Cecchetti (Brandeis University), Jordi Galí (CREI, Universitat Pompeu Fabra) and Andrew Levin (Board of Governors of the Federal Reserve System) act as external consultants and Gonzalo Camba-Méndez as Secretary.

The refereeing process is co-ordinated by a team composed of Günter Coenen (Chairman), Stephen Cecchetti, Silvia Fabiani, Jordi Galí, Andrew Levin, and Gonzalo Camba-Méndez. The paper is released in order to make the results of IPN research generally available, in preliminary form, to encourage comments and suggestions prior to final publication. The views expressed in the paper are the author's own and do not necessarily reflect those of the Eurosystem.

Address
Kaiserstrasse 29
60311 Frankfurt am Main, Germany

Postal address
Postfach 16 03 19
60066 Frankfurt am Main, Germany

Telephone
+49 69 1344 0

Internet
http://www.ecb.int

Fax
+49 69 1344 6000

Telex
411 144 ecb d

ISSN 1561-0810 (print)
ISSN 1725-2806 (online)

CONTENTS

Abstract

In this paper we analyse the ability of time and state dependent price setting rules to explain durations of price spells or the probability of changing prices. Our results suggest that simple time dependent models cannot be seen as providing a reasonable approximation to the data and that state dependent models are required to fully characterise the price setting behaviour of Portuguese firms. Inflation, the level of economic activity and the magnitude of the last price change emerge as relevant variables affecting the probability of changing prices. Moreover, it is seen that the impact differs for negative and positive values of these covariates.

JEL classification codes: C41, D40, E31.

Key Words: CPI data, Hazard functions, Inflation.

NON-TECHNICAL SUMMARY

Models of price rigidity suggested in the literature may broadly be classified into time dependent and state dependent. These two types of price setting mechanisms have different implications for the effects of monetary policy. While time dependent models imply that the effect of monetary policy shocks does not depend on the state of the economy, state dependent models predict that the probability of a price change varies according to the state of the economy and that the effects of nominal shocks on real activity and inflation are also state dependent. Thus, identifying the type of price rigidities which characterise the economy emerges as an issue of paramount importance.

In this paper we use two micro-datasets to analyse the ability of both time and state dependent models of price setting behaviour to describe the durations of price spells, or the probability of changing prices, observed in the Portuguese economy. The approach is based on the estimation of a hazard function for the duration of price spells, which allows us to test whether time varying regressors contribute to explain the probability of changing prices.

It is shown that the duration of price spells, and consequently the frequency of price changes, depends on variables such as accumulated sectoral inflation, the magnitude of the last price change and the level of demand (measured by the accumulated growth rate of the industrial production index).

In particular, it is found that the larger the accumulated sectoral inflation or the accumulated demand pressure, the larger the probability of price changes and thus, the smaller the expected duration of price spells. In contrast, the magnitude of the last price change has a negative effect on the expected duration of price spells. It is also found that the effects on the probability of changing prices differ according to whether we consider positive or negative values of the included time varying regressors.

The main message of this paper is, therefore, that time dependent models are unable to adequately describe the features of the investigated datasets, and that state dependent models are required to fully characterise the price setting behaviour of Portuguese firms.

1. INTRODUCTION

Price rigidities are usually seen as a major determinant of the response of production and inflation to nominal shocks hitting the economy. Therefore, identifying the type of price rigidities which characterise the economy emerges as an issue of paramount importance.

Models of price rigidity suggested in the literature may broadly be classified into two categories: time dependent and state dependent models. Time dependent models are characterized by the fact that the decision of changing prices is exogenous, i.e., prices change independently of the sate of the economy according to a certain statistical or deterministic rule. The best known examples of this kind of models are the ones proposed by Taylor (1980), where prices are kept unchanged during a fixed period of time, and by Calvo (1983), where prices have random durations. A major consequence of this type of models is that the effect of monetary policy shocks does not depend on the state of the economy.

In the case of state dependent models, the decision of changing prices is endogenous, i.e., at every moment economic agents decide whether or not to change their prices based on the evaluation of the costs and benefits of a price change. Thus, such models predict that the probability of a price change varies according to the state of the economy and that the effects of nominal shocks on real activity and inflation are also state dependent. Examples of this type of models are those proposed by Sheshinski and Weiss (1977, 1983), and by Dotsey, King and Wolman (1999).

From an economic point of view, state dependent models are clearly more attractive as they assume that economic agents base their decisions on a cost-benefit analysis. Time dependent models use simple ad hoc hypotheses to justify price changes and therefore may not be realistic descriptions of the behaviour of rational agents. Nevertheless, the possibility of such models being able to provide an accurate picture of the overall economy should not be excluded beforehand.

In this paper we use two micro-datasets, which were originally collected to compute the Consumer Price Index (CPI), to analyse the ability of both time and state dependent models of price setting to describe the durations of price spells, or the probability of changing prices, observed in the Portuguese economy. The approach adopted here is based on the estimation of a hazard function for the duration of price spells, which allows us to test whether time varying regressors contribute to explain the probability of changing prices. With this approach it is possible to fully explore the cross-section dimension of our datasets.

The paper is organized as follows. Section 2 describes the data used in this study and provides some preliminary results on the adequacy of time dependent models of price setting behaviour. Section 3 discusses and estimates a hazard function which includes state dependent features, and section 4 concludes.

2. THE DATA

2.1. Brief data description

This paper uses two micro-datasets on consumer prices collected by the Portuguese *Instituto Nacional de Estatística* (INE) in order to produce the aggregate Consumer Price Index for Portugal. These two datasets cover two different periods, January 1992 to December 1997 and January 1997 to December 2002. Hereafter, the two datasets will be referred to as CPI1 and CPI2, respectively. Although the two datasets have different characterisitcs (see Appendix 1 for further details), both include information on prices at the outlet and product level, covering outlets nationwide. The basic observation is that of a price of an item in a particular outlet at a given point in time. This item is followed over time within the same store. It is worth mentioning, however, that forced substitutions may occur in the case of the Consumer Price Index but the exact identification of such situation is not possible.

There are over $3,000,000$ observations in CPI1 and around $2,000,000$ in CPI2 up to January 2001.[1] This corresponds to data collected on over $10,000$ outlets and 460 items

[1] The sample periods considered in the analysis were shortened in order to exclude potential cash changeover effects (2001 and 2002 for CPI2) and to make sure that an homogeneous period will be analysed as far as seasonal pattern is concerned (for CPI1 and CPI2).

for CPI1 and $13,000$ outlets and 780 items for CPI2. Information on prices is collected on a number of outlets for each item, although brands and packages are not necessarily the same across stores. Consequently, prices for the same items across stores are not comparable.

During this period, the prices of some goods (notably energy) were controlled by the government. This, of course, has important consequences for the observed frequency of price changes and the duration of price spells of such products. Consequently, the prices of energy and other goods with prices set administratively are not considered in our analysis.

Apart from prices, product code and outlet code, the CPI datasets also include information on date, geographical location of the outlet (grouped in seven possible regions using the NUTS II classification), type of outlet (allowing for a distinction between big and small stores in CPI1 and hypermarkets, supermarkets, classical stores, discount stores, market and other in CPI2), a dummy for perishable food products and some information on the weights of the items in the typical consumer bundle. As the CPI records are under statistical secrecy, it is impossible to know the specific goods and services for which the prices are collected in the survey.

2.2. Preliminary data analysis

Figures 1 and 2 display the share of prices that change in each quarter (hereafter referred to as the "frequency of price changes") for CPI and CPI2. Two important characteristics emerge from these Figures. First, the average frequency of price changes is larger for CPI1 (0.39 in CPI1 against 0.35 in CPI2). Second, and perhaps more significantly, the frequency of price changes exhibits a downward trend over the sample period. This is especially clear for CPI1, whose sample period (1993-1997) is characterized by a higher average (aggregate) inflation, with a significant decreasing trend (average inflation was 4.1% during the sample period corresponding to CPI1 and 2.6% during the period corresponding to CPI2). These characteristics of the data are at odds with time dependent price setting mechanisms, which would imply that the fraction of prices that change in every period is (approximately) constant over time.

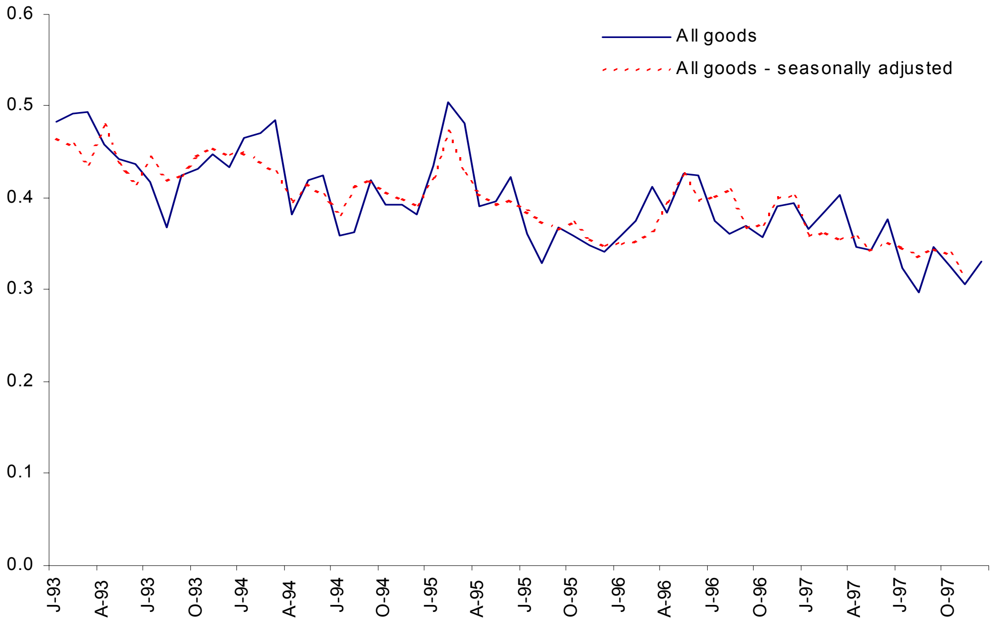

FIG. 1. CPI1 - Frequency of price changes

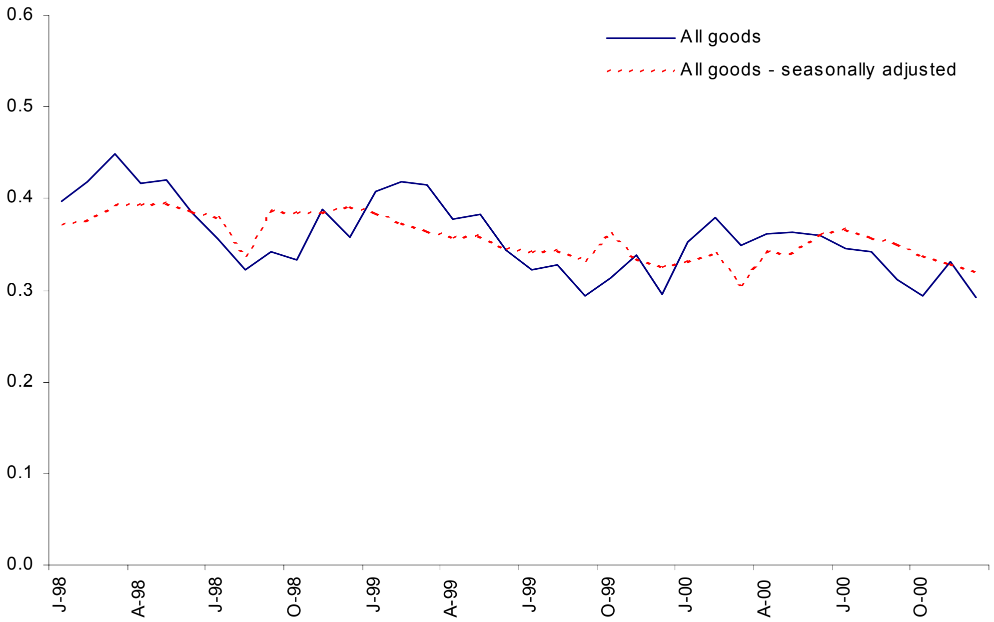

FIG. 2. CPI2 - Frequency of price changes

Figures 1 and 2 also display the seasonally adjusted frequency of price changes. It is seen that seasonality is more significant in case of CPI2.[2]

Seasonality in the frequency of price changes is usually seen as evidence of time dependent behaviour, because a strong seasonal factor, in a given quarter of the year, say, would imply that a large proportion of firms is likely to change their prices on a yearly basis, on that particular quarter. Consequently, the fact that seasonality is stronger in CPI2 than in CPI1 is in accordance with the expected result of an increased proportion of time dependent behaviour for low and stable levels of inflation, which characterizes CPI2. However, the fact that the importance of seasonality has changed over time as the state of the economy changed (from CPI1 to CPI2) must also be seen as evidence of state dependent behaviour in price setting practices by firms.[3]

In summary, the analysis of the evolution over time of the frequency of price changes suggests that we may expect significant state dependent characteristics to be present in the data. In the next section we try to identify which factors explain the probability of observing a price change at the micro level.

3. WHICH FACTORS DETERMINE THE PROBABILITY OF OBSERVING A PRICE CHANGE?

The analysis in this section proceeds in two different steps. We start by discussing some technical issues that must be taken into account when estimating and interpreting the aggregate hazard function. Then, we investigate which factors affect the probability of observing a price change at the micro level by estimating hazard functions with time varying regressors.

---

[2] The variance of the average quarterly seasonal factors is 0.008 in case of CPI2 and 0.004 in case of CPI1, which confirms, as a simple glance at Figures 1 and 2 would suggest, that seasonality is stronger in case of CPI2.

[3] For instance, the (average) estimated seasonal factor for the first quarter of the year equals 1.08 in case of CPI1 and 1.11 in case of CPI2. Notice, however, that part of the differences in the seasonal pattern between CPI1 and CPI2 may also be due to methodological changes in the data collection procedures underlying the two datasets (see Appendix 1 for details).

3.1. The aggregate hazard function: some estimation issues

Estimating and interpreting the aggregate hazard function for the economy as a whole would not raise any special difficulties if the population of firms/stores were homogeneous with respect to the distribution of price spells. However it is well known (see for instance Dias, Dias and Neves, 2004, for Portugal) that the frequency of price changes differs significantly not only across sectors but also across firms/stores. In particular, the frequency of price changes in the food sector is significantly higher than the frequency of price changes in the services sector (0.37 against 0.11). Similarly, the frequency of price changes is higher in big stores (supermarkets and hypermarkets) than in small stores. This heterogeneity, if not properly taken into account, may have severe distorting effects on the resulting estimates of the aggregate hazard function.

Two major implications of heterogeneity are worth discussing. The first one is well known from the literature on duration models, and concerns the bias of estimated hazards towards *negative duration dependence,* brought about by the fact that products (in the population and in the sample) are heterogeneous with respect to the average price durations (see, for instance, Heckman and Singer, 1984, or Lancaster, 1990). This characteristic of the data implies that the aggregate estimated hazard, as obtained for instance using the nonparametric Kaplan-Meier estimator, tends to display a decreasing pattern irrespective of the specific hazard functions underlying each individual homogeneous group of firms or products. Intuitively, this is so because the share of shorter price spells (spells corresponding to products with more flexible pricing rules) decreases as the larger time horizons are considered and consequently the hazard rate decreases, creating the illusion of a stronger negative duration dependence than what actually exists.[4] An obvious consequence of this heterogeneity is that the estimates for the aggregate hazard function cannot be taken at face value. In particular, an aggregate downward sloping hazard function may be compatible with the existence (at the homogeneous group level)

[4] Note that according to the Kaplan-Meier estimator, the hazard rate, $h(t)$, defined as the rate at which spells are completed after duration t, given that they last at least until t, is estimated as $\hat{h}(t) = d_i/n_i$ where d_i is the number of spells that end at time t and n_i is the number of spells that last at least until time $t-1$.

of both constant hazard functions as implied, for instance, by the Calvo model, and upward sloping hazard functions, as predicted by some state dependent models.

A second consequence of neglected heterogeneity turns up when the dataset, as in our case, is a panel of firms (or prices) observed over a fixed period of time such that for each product there are several available price spells. As it is shown in Appendix 2, this situation creates a particular form of length biased sampling, and appropriate estimators have to be used to deal with this issue. Failing to account for this particular sampling mechanism will exacerbate the bias of the estimated aggregate hazard function towards negative duration dependence. Intuitively, the idea is that heterogeneity in a dataset of price spells results in an over-representation of shorter spells (and a corresponding under-representation of the longer ones) and this implies inconsistency of the standard estimator. As explained in Appendix 2, this situation can be overcome using an estimator in which the number of spells of duration t for unit i, say, are weighed by the inverse of the number of spells for unit i (see, equation A.2 in Appendix 2). Alternatively, an estimator with a fixed and equal number of spells for each product can be used (equation A.5 in Appendix 2). In the next sub-section we will work with a single spell for each product randomly drawn from the full set of spells, after excluding the left-censored ones.

3.2. Hazard functions for the duration of price spells

3.2.1. Unconditional hazard functions

Figures 3 and 4 display the Kaplan-Meier estimates of the unconditional aggregate hazard functions for CPI1 and CPI2. Both functions display a downward sloping pattern. Given the discussion above, this does not come as a surprise and should not be seen as having significant implications for the shape of the hazard functions at the firm or homogenous group level.

The estimated unconditional hazard function may help to unveil the type of heterogeneity that is flagged by the presence of non-monotonicity. For instance, the estimated hazard functions exhibit significant spikes at durations of 1 and 4 quarters for both CPI1

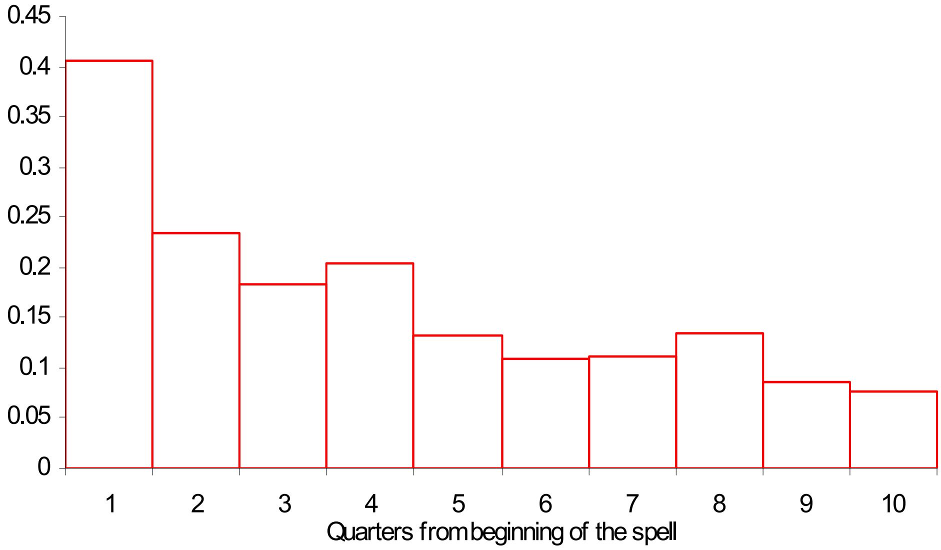

Fig. 3. CPI1 - Unconditional hazard function

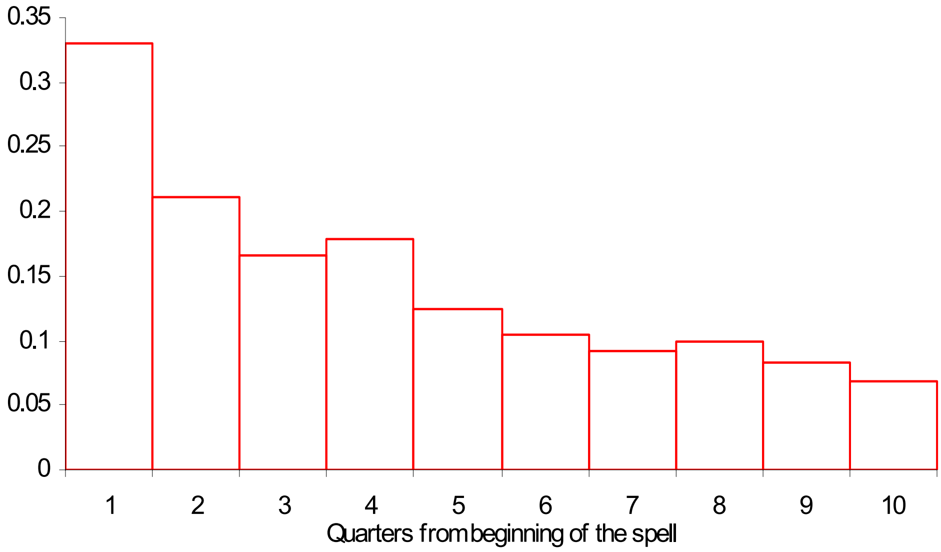

Fig. 4. CPI2 - Unconditional hazard function

and CPI2. This means that price spells of 1 and 4 quarters are likely to constitute a large share of the whole sample of spells. This evidence may also suggest that there are groups of firms with different average probability of changing prices. In particular, the spikes at durations of 1 and 4 suggest that there may be two different groups of firms/products, one that changes prices very frequently (in the limit, in each and every quarter), and another one that changes prices on a yearly basis (every four quarters).

Notice that the existence of different groups of firms cannot be seen as evidence in favour of time dependent behaviour because what matters for the distinction between time or state dependent behaviour is not the existence of groups of firms with different average price spells but whether the expected duration of such price spells changes over time as the state of the economy changes.

It is also important to stress that the unconditional hazard functions, as depicted in Figures 3 and 4, shed no light whatsoever on whether firms follow time dependent or state dependent price setting rules. Indeed, the unconditional hazard function is estimated under the assumption that the probability of a price change is constant over time. However, the constancy over time of the probability of changing prices is exactly what we need to investigate to be able to distinguish between time dependent and state dependent price setting schemes. Therefore, in order to distinguish between time and state dependent behaviour using the unconditional hazard function, one would need to estimate several hazard functions corresponding to different periods of time and compare them with each other. Stability of the hazard functions over time would constitute evidence of time dependent behaviour, while a time varying profile of the different hazard functions would be seen as evidence favouring state dependent behaviour.[5] Due to the limited time span of our data sets, this approach is not pursued here.

3.2.2. A discrete time duration model with time varying covariates

The analysis of the unconditional hazard function has highlighted the problems caused by uncontroled heterogeneity. However, if the purpose is to test which type of price set-

[5] The implicit identifying assumption is that the state of the economy varies over time. In a completely stable economy with constant inflation it would not be possible to distinguish between time and state dependent models since both would imply hazard functions that are stable over time.

ting practices – time or state dependent – better characterizes the behaviour of economic agents, we need to look for a different instrument of analysis. To investigate whether or not there is significant evidence of state dependent behaviour by Portuguese firms, we use a discrete time duration model.[6]

A discrete time duration model may be regarded as a sequence of binary choice problems defined on the surviving population at each duration. The parametric model we consider is characterized by the fact that $h_i(t)$, the hazard function for product i in period t, is specified as a complementary log-log model (see, for instance, Jenkins, 1995) of the form

$$h_i(t) = P[T_i^* = t | T_i^* \geq t, W_i(t)] = 1 - \exp\{-\exp[W_i(t)]\}. \tag{1}$$

Because of the specification of $h_i(t)$ that is chosen, this model is the discrete-time counterpart of an underlying continuous-time proportional hazards model (see Prentice and Gloeckler, 1978)

For reasons that will soon be clear, the following specification for $W_i(t)$ in (1) is adopted

$$W_i(t) = \theta(t) + \alpha' Z_i + \beta' X_{it}^a + \delta' X_{it}^- + \gamma' T_t. \tag{2}$$

In equation (2), $\theta(t)$ is the function of t whose specification is discussed later, Z_i is the vector of time constant covariates, X_{it} is the vector of time varying conditioning variables whose elements enter $W_i(t)$ in two different ways, X_{it}^a represents a vector containing the absolute value of the elements of X_{it}, X_{it}^- is a vector containing the product of the elements of X_{it} by dummies which equal 1 when the element of X_{it} is negative, being 0 otherwise, and, finally, T_t is the vector of time dummies (seasonal and yearly dummies). In the estimated models below, Z_i includes several dummy variables related to the type of store, the type of product, the geographical location and, in case of CPI1, also changes in the VAT rates.[7]

[6] For similar approaches see Aucremmane and Dynne (2005) and Fougère, Le Bihan and Sevestre (2005).

[7] As the two CPI datasets have different levels of information, the number of dummy variables included in Z_i is slightly different in the two models. In the model for CPI1, Z_i includes one dummy variable to signal whether a store is big or small, three dummy variables indicating the type of product

The vector X_{it} includes time varying regressors that economic theory suggests may be relevant factors in explaining the probability of changing prices over time. In particular, for the two models estimated below, X_{it} includes: i) the magnitude of the last price change (w_{it}); ii) the sectoral inflation rate accumulated since the beginning of the spell (π_{it}) lagged one quarter;[8] and iii) the accumulated growth rate of industrial production index (g_t). The economic rationale for the use of such variables can be seen, for instance, in Cecchetti (1986) and Dotsey *et al.* (1999). In particular, Cecchetti (1986) develops a model of firm price changes in the presence of adjustment costs, in which the probability of changing prices varies positively with the cumulative inflation since the last price change (because it measures the disequilibrium between the actual price level and the optimum price level), negatively with the magnitude of the previous period price change and positively with the industry sales growth (used as a proxy to measure demand pressure).

In order to capture duration dependence, instead of imposing a given functional form on $\theta(t)$, we use a more flexible approach which consists in introducing an additive dummy variable for each duration. Thus we introduce a variable dur_t ($t = 1, 2, ..., m$) which equals 1 if the hazard corresponds to a duration of a price spell of t quarters, being 0 otherwise. In our application we have decided to saturate the hazard function, i.e., we consider a different parameter for each duration. One major advantage of this approach is

(Unprocessed Food, Processed Food, Non-Energetic Industrial Goods, and Services) and six dummy variables relative to the geographical location of the stores (North, Centre, Lisbon, Alentejo, Algarve, Azores and Madeira). In the model for CPI2, Z_i includes the same six dummy variables relative to the geographical location of the stores, but it also includes five dummy variables related to the type of store (Hypermarket, Supermarket, Discount Store, Classical Store, Market and Private/Public Services) and five dummy variables relative to the type of product (Unprocessed Food, Processed Food, Non-Durable goods, Semi-Durable Goods, Durable Goods and Non Administered Price Services). We note from the outset that the poorer level of information for CPI1 (regarding the type of store and the sectoral disaggregation) is expected to cause additional difficulties at the empirical level, because it may imply some residual neglected heterogeneity, with consequences on the properties of the estimators for the parameters of the included regressors.

[8] Sectoral inflation is measured at the five digits aggregation level (COICOP level).

that it allows us to capture the effect of the covariates without the need for any additional parametric assumptions on the distribution of neglected individual heterogeneity.[9]

In our sample there are negative as well as positive changes not only at the product level (price changes) but also at the sectoral or national level (negative and positive sectoral inflation as well as industrial production growth). This situation raises the issue of how to distinguish the effects of negative and positive changes of the covariates on the probability of a price change since, at least from some of these variables, asymmetric effects are expected. In fact, one of the costs usually incurred by firms when they change prices is the negative reaction from customers (customer anger) following price increases, but such a cost is not expected to be relevant for price decreases. In (2) we account for the possibility of asymmetric effects by introducing both the absolute value of X_{it} and X_{it}^{-}. This specification has the advantage of allowing simple tests of symmetry to be performed.

In the models estimated below, the sectoral inflation rate accumulated since the beginning of the spell (π_{it}) enters lagged one quarter. This is done to circumvent the potential simultaneity problem stemming from the fact that, at this level of disaggregation, inflation as a regressor may be expected to reflect changes in prices occurring at the product/firm level. Finally, notice that π_{it} and g_t are defined at each point in time during the spell and not only at the end of the spell. This allows for the possibility of the effect of such regressors to vary over the spell, as inflation or production growth accumulate, thus increasing the disequilibrium between the actual price and the optimum price that would prevail in the absence of price rigidities.

[9] It is well known (see, Ridder, 1987) that neglected heterogeneity is immaterial if the baseline hazard is reasonably flexible to begin with. Portugal and Addison (2000) investigated this claim by estimating hazard functions with and without the additional heterogeneity controls and concluded that the shape of the hazard function is largely unaffected by the incorporation of such additional controls. For this reason, we decided to make the baseline hazard as flexible as possible by saturating it with dummies. Some authors use an intermediate approach, which consists in using a flexible baseline hazard and explicitly integrating out the neglected heterogeneity (see, for instance, Bover *et al.*, 2002). However, when the hazard function is saturated, it is not possible to identify the effects of neglected heterogeneity. For this reason, in the models estimated below, no additional unobserved individual heterogeneity is accounted for.

With the hazard function specified as in (1) and (2), estimation of the parameters of interest can be performed using standard likelihood methods. Let us assume that we have N products (identified at the outlet or store level) indexed by $i = 1, 2, ..., N$. Let m_i denote the duration of the spell corresponding to product i, which by assumption starts in period $t = t_i$ and ends at time $T_i = t_i + m_i$. It can be shown (see, Jenkins, 1995, or Bover, Arellano and Bentolila, 2002) that the log-likelihood function for this problem may be written as

$$\ln(L) = \sum_{i=1}^{N} \left\{ (1-\delta_i) \left[\sum_{t=t_i}^{T_i-1} \ln\left[1-h_i(t)\right] + \ln h_i(T_i) \right] + \delta_i \sum_{t=t_i}^{T_i} \ln\left[1-h_i(t)\right] \right\}, \quad (3)$$

where δ_i is a dummy variable which equals 1 if spell i is (right) censored and is zero otherwise.

Table 1 reports the main estimation results for CPI1 and CPI2.[10] The two models were estimated using a random sample that includes a single price spell for each product. The model for CPI1 uses data for the period 1992-1997, while model for CPI2 was estimated with data for the period 1997-2000. Data from 2001 and 2002 were not used in the estimation to avoid the problems brought about by the euro cash changeover.

The first point to be stressed is the fact that the coefficients associated to the time varying regressors, which measure the state of the economy, are in general individually significant and, using the likelihood ratio test, the null hypothesis that the included time varying regressors are not jointly significant is strongly rejected. Moreover, in view of Cecchetti's (1986) model, the regressors display the expected sign. Cumulative inflation, π_{it}, and the level of activity, as measured by the cumulative growth rate of industrial production, g_t, appear as significant variables that increase the probability of changing prices and thus reduce the expected duration of price spells. In contrast, but also as expected, the magnitude of the last price change, w_{it}, appears with a negative sign.

[10] Results for all the controls used in the estimation are displayed in table A1 in Appendix 3. A brief look at those results shows that the dummy variables for the type of outlet, the type of sector and the region in which the good is sold are usually significant, which constitutes evidence that it is important to account for such heterogeneity in the estimated models. A similar conclusion applies for the quarterly seasonal dummies as well as for the yearly dummies, which are meant to capture changes over time of the overall economic conditions not explained by the included time varying regressors.

Table 1 - Main estimation results for CPI1 and CPI2

	CPI1		CPI2	
	Coef.	t-stat	Coef.	t-stat
Sectoral accumulated inflation (π)				
$\|\pi\|$	0.780	3.971	2.404	6.387
π^-	-1.603	-2.151	-6.645	-9.301
Magnitude of the last price change (ω)				
$\|\omega\|$	-0.783	-14.741	-0.292	-5.471
ω^-	-1.251	-21.173	-0.681	-11.709
Accumulated industrial production growth rate (g)				
$\|g\|$	0.593	3.061	4.140	8.573
g^-	0.308	1.505	-3.560	-5.693
VAT				
$1995Q1$	0.584	20.202	–	–
$1996Q3$	-0.230	-7.135	–	–
$1996Q4$	-0.200	-6.490	–	–
Sample size	183923		186284	
Number of Spells	58957		51218	
Log-likelihood	-84676		-77545	
Likelihood ratio test* ($p-value$)	1002 (0.00)		2651 (0.00)	

* The null hypothesis is that the all coefficients in the table are zero.

The fact that, even controlling for different sources of heterogeneity, coefficients associated to the time varying regressors are statistically very significant, suggests that simple time dependent models are not likely to provide a reasonable approximation to the data and that state dependent models may be required to fully characterize the price setting behaviour of Portuguese firms.

In the case of CPI1, the estimated model also includes three dummy variables to account for changes in the VAT rates occurred in the first quarter of 1995 and the third and fourth quarters of 1996, respectively. Investigating how the probability of changing prices reacts to changes in the VAT rates is an interesting exercise because the reaction by firms to changes in the VAT rates, if significant, may be seen as evidence in favour of state dependent behaviour. From the estimated parameters in Table 1, we see that the coefficient of the VAT dummy variable for the first quarter of 1995 is positive and highly significant, as expected, but the coefficients of the other two dummy variables are negative. Such an outcome may, however, be easily understood. Changes in the VAT rates in the first quarter of 1995 were broad in scope and quantitatively significant, while the changes that occurred in 1996 were quantitatively smaller and affected only some specific sectors. This explains why in Figure 1 a significant peak in the frequency of price changes emerges during the first quarter of 1995, but no such behaviour is discernible in the second half of 1996. On the contrary, for these two observations, we see that, if anything, the frequency of price changes exhibits a small trough, which means that the effect (if any) of a change in the VAT rates may have been more than compensated by shocks with the opposite effect. The fact that the two dummy variables remain significant in the estimated model, strictly speaking, means that the reasons for the trough in the second half of 1996 are not fully captured by the regressors in the model.

The estimated models also show that the assumption of symmetric effects is rejected by the data (the coefficients associated with the regressors in X_{it}^{-} are generally statistically significant), allowing us to conclude that positive and negative values of accumulated inflation, the magnitude of the last price change and accumulated activity growth have different effects on the probability of changing prices. For instance, for both CPI1 and CPI2, negative inflation exerts a larger effect on the probability of changing prices than positive inflation.

In order to get a quantitative assessment of the importance of the conditioning variables for the hazard function, we carried out two different exercises. One aimed at investigating how alternative trajectories for the time varying regressors impact on the probability of changing prices and the other aimed at assessing the impact on the expected duration of price spells from changes on those covariates.

In the first exercise we simulate the estimated hazard functions for alternative trajectories of the time varying covariates. For inflation and production index we consider the two alternative hypotheses of a constant annual inflation and a constant annual growth rate of the production index of $+4\%$ and -4%. Notice that these hypotheses imply that the corresponding covariates (accumulated inflation and accumulated production growth) vary over the spell (they are strictly increasing or decreasing over time). For the magnitude of the last price change we consider the two alternative figures of $+8\%$

and -8%, which, in this case, imply that the corresponding covariate is constant over the spell.[11] Figures 5 and 6 depict the different hazard functions for the reference group implicit in the estimated models.[12] Table 2 presents the differences between the baseline hazard and the hazard corresponding to a given trajectory of the conditioning variable, for the durations of 1, 4, 6, 8 and 10 quarters.[13]

In interpreting these results it is useful to recall that in this context the hazard function is the probability of observing a price spell of duration t, conditional on the fact that the price was kept unchanged for the previous t periods. Therefore, the differences between the baseline hazard and the hazard corresponding to a given trajectory of the conditioning variable may be seen as the variation in this conditional probability resulting from the assumed trajectory of the covariate. For instance, in case of CPI2, an annual inflation of 4% increases the probability of a price change by 3.1 percentage points after one year, while, under the same circumstances, production growth increases the probability of changing prices by 5.9 percentage points (conditional on the spells lasting for at least one year). It can also be seen from Table 2 and Figure 6 that, in case of CPI2, variations in the probability of changing prices are larger for negative inflation and negative production growth. Finally, both for CPI1 and CPI2, the effect of the magnitude of the previous price change appears to have little economic relevance.

In the previous exercise the trajectories defined for accumulated inflation and accumulated production growth are time varying. We now consider an alternative exercise

[11] The magnitude of 8% for the previous period price change is close to the sample median of the magnitudes of price changes in case of CPI2.

[12] For CPI1 the reference group corresponds to a spell occurring in the first quarter of 1993 for the price of a good of the "unprocessed food sector", sold in a "big store" in "Lisbon". For CPI2 the reference group corresponds to a spell occurring in the first quarter of 1998 for the price of a good of the "unprocessed food sector" sold in a "hypermarket", in "Lisbon". Although the two reference groups are as close as possible, the fact that they are not defined in the same way suggests some caution is needed in the comparison of the two sets of results. Naturally, the need for caution is even stronger when comparing results for different countries, unless one can ensure that they refer to the same reference groups and are based on comparable datasets.

[13] The baseline hazard for the reference group is obtained by setting the trajectories for the time varying regressors equal to zero, over the spell.

in which the trajectories for those covariates are constant over the spell.[14] Such an exercise is close to a *ceteris paribus* analysis that allows us to investigate the impact on the expected duration of price spells of changes in the values of those covariates, and, therefore, get a better idea of how economically relevant are the state-dependent features of price setting practices by Portuguese firms.

Table 2 - Effects of time varying regressors on the probability of changing prices*

Inflation				
	CPI1		CPI2	
Quarters	+4% *year*	−4% *year*	+4% *year*	−4% *year*
1	0.003	0.008	0.008	0.032
4	0.011	0.034	0.031	0.128
6	0.013	0.041	0.038	0.168
8	0.020	0.062	0.049	0.222
10	0.018	0.059	0.050	0.247
Previous price change				
	CPI1		CPI2	
Quarters	+8%	−8%	+8%	−8%
1	−0.022	0.013	−0.009	0.011
4	−0.022	0.013	−0.008	0.011
6	−0.017	0.010	−0.006	0.009
8	−0.019	0.012	−0.006	0.008
10	−0.014	0.009	−0.005	0.007
Industrial Production Index				
	CPI1		CPI2	
Quarters	+4% *year*	−4% *year*	+4% *year*	−4% *year*
1	0.002	0.001	0.015	0.028
4	0.009	0.004	0.059	0.111
6	0.010	0.005	0.074	0.144
8	0.015	0.007	0.095	0.190
10	0.014	0.007	0.100	0.209

*The entries in the table are the difference between the hazard function and the baseline hazard.

[14] This corresponds to assuming that inflation or production growth is non zero in the first quarter and zero afterwards, so that the corresponding accumulated values are constant over the spell. In the case of the magnitude of the previous price change, the assumption is the same as in the first exercise because this variable does not change over the spell.

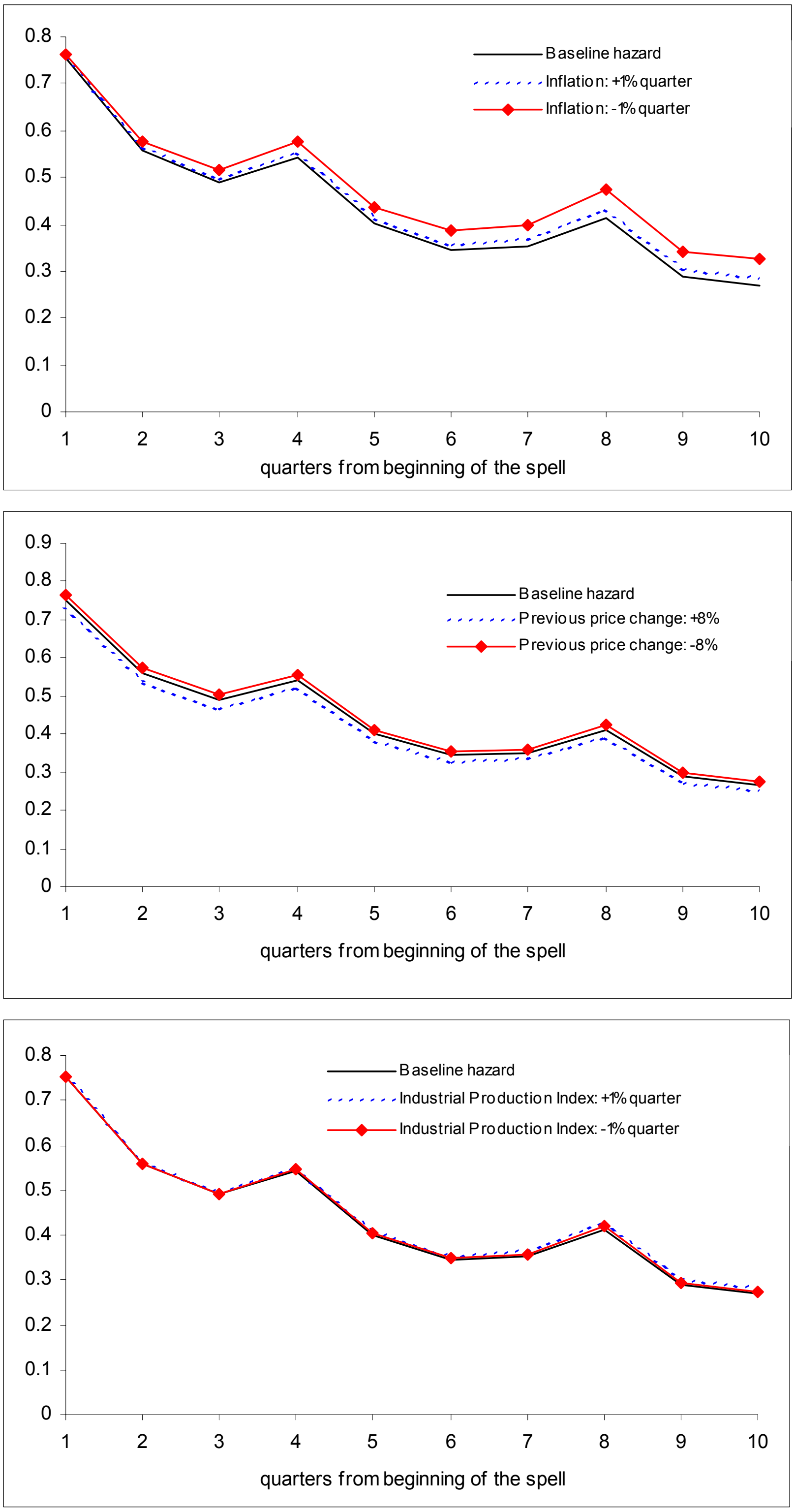

FIG. 5. CPI1 - Conditional hazard function

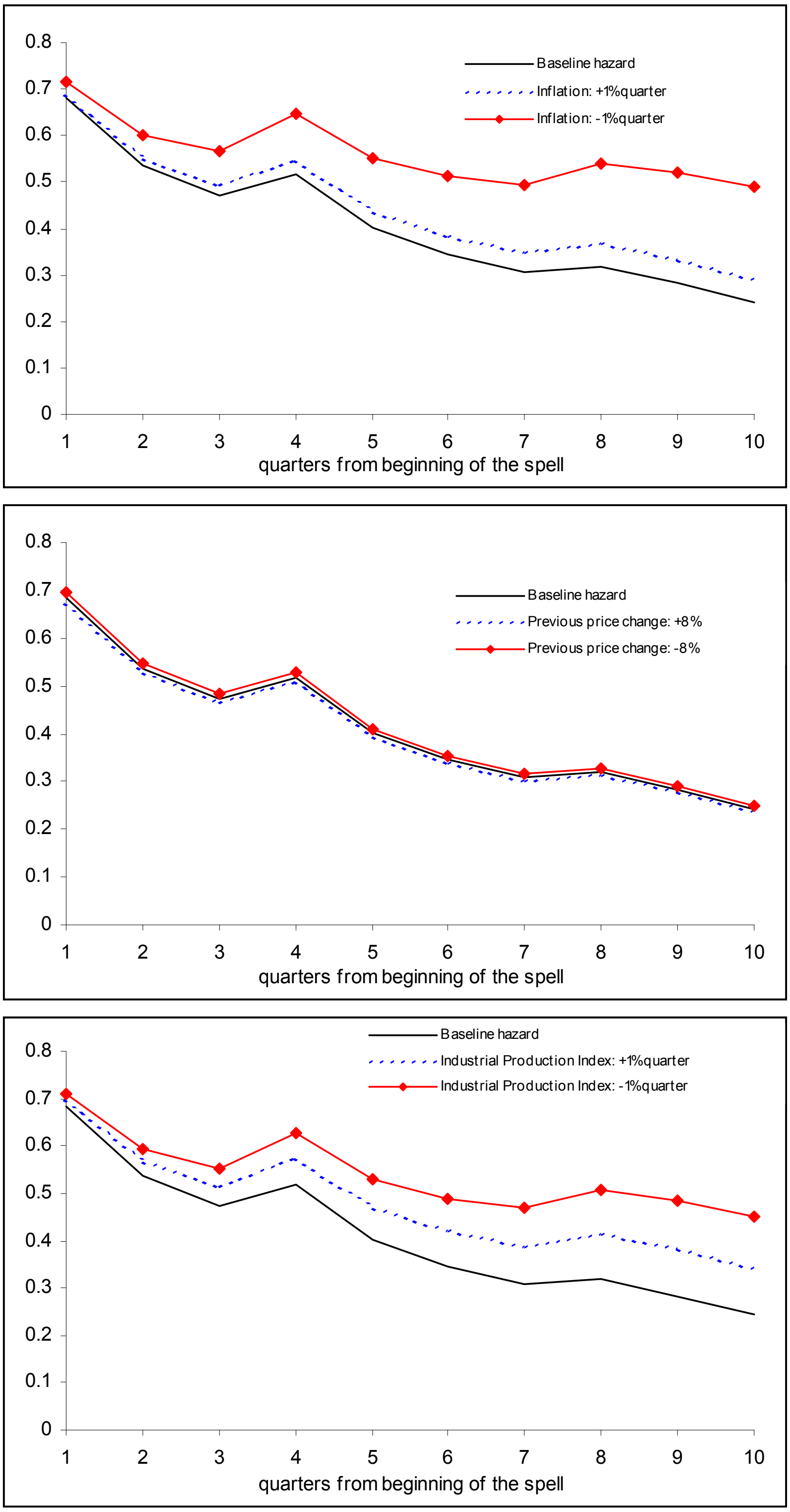

Fig. 6. CPI2 - Conditional hazard function

Table 3 and Figure 7 display the expected duration of price spells for different values of the covariates. For instance, the right-upper panel of Figure 7 displays the expected duration of price spells of CPI2 for different levels of sectoral accumulated inflation (the corresponding figures can be found in Table 3). When all the time varying regressors are equal to zero the expected duration of a price spell is 1.67 quarters (baseline case). When accumulated inflation is equal to 4% (everything else equal), the expected duration of a price spell is reduced to 1.55 quarters and further to 1.46 quarters if accumulated inflation is equal to 8%. In case of negative inflation the expected duration of a price spell is reduced to 1.31 quarters when accumulated inflation is equal to −4% and to 1.13 quarters when inflation is −8%. From Table 3 and the lower-right panel of Figure 7 we see that, in case of CPI2, production growth also has a strong impact (even somewhat larger than inflation) on the expected duration of a price spell. For instance, the expected duration of a price spell (which is 1.67 quarters in the baseline case) is reduced to 1.47 quarters when accumulated production growth is equal to 4%. As the findings of the previous exercise suggest, Table 3 and panels of Figure 7 show that the magnitude of the last price change has a minor impact on the expected duration of a price spell. For instance, for a previous period price change of 8% the expected duration of a price spell increases only to 1.71 quarters from a baseline of 1.67 quarters.

Table 3 - Expected durations of price spells for alternative trajectories of the covariates

Accumulated Inflation	−8%	−4%	−2%	0%	2%	4%	8%
CPI 1	1.28	1.35	1.39	1.43	1.41	1.40	1.37
CPI 2	1.13	1.31	1.46	1.67	1.61	1.55	1.46
Previous price change	−8%	−4%	−2%	0%	2%	4%	8%
CPI 1	1.40	1.41	1.42	1.43	1.44	1.46	1.49
CPI 2	1.63	1.65	1.66	1.67	1.68	1.69	1.71
Accumulated IPI growth	−8%	−4%	−2%	0%	2%	4%	8%
CPI 1	1.41	1.42	1.42	1.43	1.42	1.41	1.39
CPI 2	1.16	1.34	1.48	1.67	1.56	1.47	1.32

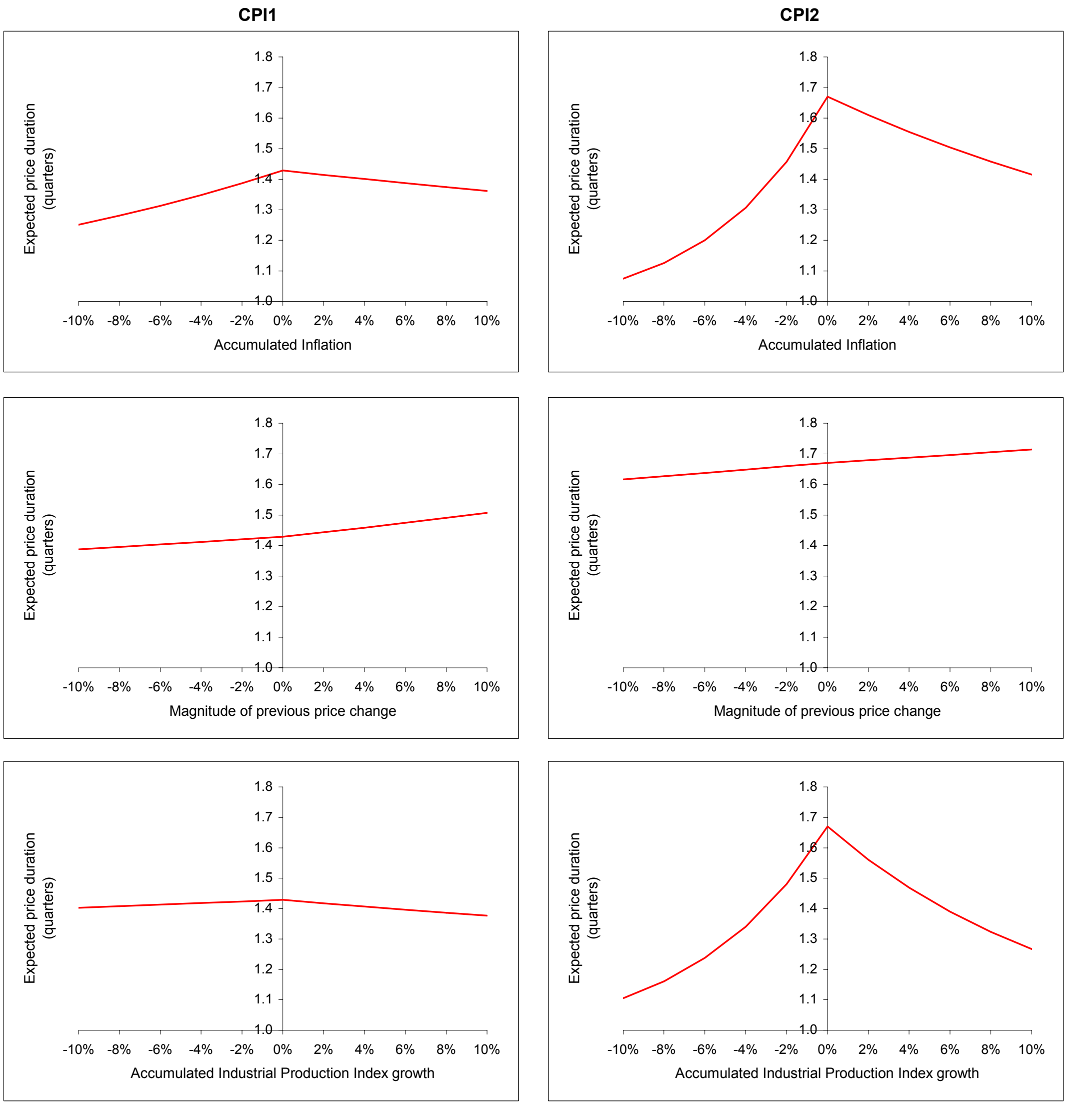

FIG. 7 - Expected duration of price spells for different values of the covariates

Overall, the two exercises allow us to conclude that inflation has a significant impact on the probability of changing prices or the expected duration of price spells and that negative inflation has a larger impact than positive inflation[15]. The fact that negative

[15] Two explanations for the lower estimated effect of positive inflation in case of CPI1 can be suggested. On the one hand, during the time period covered by CPI1, aggregate inflation displays a strong decreasing trend which is reflected in a large number of sectors (at COICOP level) displaying negative inflation, especially so for some items that are likely to be over represented in our sample (food items, for instance). This implies too little information for the precise estimation of the effect of positive

inflation both for CPI1 and CPI2 appears as having a more significant impact than positive inflation is intuitive. On the one hand, both positive and negative accumulated inflation are expected to impact positively on the probability of changing prices because they are proxying the increasing disequilibrium between the actual price and the optimum price. On the other hand, as noted above, price decreases are not expected to bring about "customer anger", thus reducing the true adjustment costs of decreasing prices when there is negative accumulated inflation.

Concerning the impact of business cycle conditions, as measured by the accumulated growth rate of industrial production since the last price change, we see that both positive and negative changes are associated with a larger probability of changing prices and thus a smaller durations of price spells. It is also seen that, in the case of CPI2, negative accumulated demand shocks have a larger effect than positive shocks and that the effect of demand shocks is estimated to be larger for CPI2 than CPI1[16].

Finally, in what concerns the effect of the magnitude of the previous price change (the single variable which is defined at the product level), it is seen that positive changes have a negative impact on the probability of changing prices, while a negative price change has a positive impact. Thus, prices appear to last longer after a positive price change. However, quantitatively the impact of such regressor is not as large as the impact of inflation or production growth.

3.3. Sensitivity analysis

Two limitations of the models estimated above are worth discussing. First, so far it has been assumed that the effect of the regressors is the same across the different sectors. Second, the specification used imposes that the effects of the time varying regressors on the probability of changing prices does not depend on the direction of the price change.

---

inflation. On the other hand, we have seen that the level of information for CPI1 on some potential non observed heterogeneity (poorer disaggregation level by type of store and by sector) makes it more difficult to identify the factors influencing the probability of occurring a price change in CPI1. This latter fact may also explain why the effect of economic activity also appears as having a smaller effect on the probability of changing prices for CPI1 than for CPI2.

[16] See previous footnote for an explanation.

We start by addressing the first issue. There is no theoretical reason to expect the magnitude of the effects of the regressors on the probability of changing prices to be the same across the different sectors. On the contrary, we can easily imagine that the importance of accumulated inflation on the probability of changing prices may differ, say, between an "unprocessed food" and a "services" item. More generally, it may be the case that some sectors exhibit a close to time dependent behaviour (not reacting to inflation or demand shocks) while others display a strong state dependent behaviour. Such potential parameter heterogeneity is not accounted for in the previously estimated models, which impose the same coefficients across all goods in the economy. To investigate whether there are significant differences in the responses of the probability of changing prices for the different sub-sectors, we estimated the above models for each of the sub-sectors of CPI2. These are: "unprocessed food", "processed food", "non-food" and "services", where the "non-food" sector is further disaggregated into "non-durable", "semi-durable" and "durable" items.[17] The major conclusions of such a disaggregate analysis are as follows. Positive inflation emerges as having a larger effect in case of "unprocessed" and "processed food", while negative inflation is more important in the "services", "processed food" and "durables". Positive inflation appears wrong signed in the case of "semi-durables" and statistically non-significant in the case of "durables". As for the magnitude of the last price change, the effects are relatively homogeneous for most of the sub-sectors and somewhat larger in case of "services". Finally, for the impact of business cycle conditions, it is found that a positive accumulated growth rate of industrial production has a especially large effect on "unprocessed" and "processed food" and is not significant in the case of "services", "non-durables" and "durables". With the exception of "semi-durables", where there is no evidence of asymmetric effects, negative growth rates of industrial production emerge as more significant than positive growth rates in all the sub-sectors considered.[18]

[17] We restrict the disaggregate analysis to CPI2 items because the information set of CPI1 is clearly poorer not only about the level of disaggregation by "sub-sector", but also about the classification by the type of store (see, sub-section 3.2). This way we expect potential bias stemming from neglected heterogeneity to be smaller in case of CPI2.

[18] Detailed output of the estimated models is available from the authors upon request.

Let us now address the second issue. The models estimated so far allow us to tell whether positive or negative changes in the regressors have a positive or a negative impact on the probability of changing prices, but not whether such regressors have a positive or negative impact on the probability of changing prices upwards or downwards. However, from a theoretical point of view, one would like to be sure that positive accumulated inflation or positive accumulated demand shocks are associated to positive price changes and similarly that negative inflation or negative activity growth are associated to negative price changes. A simple way of investigating this issue is to estimate two different regressions for two different sub-samples, which separate the spells according to whether they end with prices changing upwards or downwards.[19]

From the estimated models, both for CPI1 and CPI2, it is seen that, in fact, positive accumulated inflation increases the probability of changing prices upwards and decreases the probability of changing prices downwards. In turn, negative accumulated inflation increases the probability of changing prices downwards, as expected, but the effect on the probability of changing prices upwards is not very well defined. It does not emerge as significant in case of CPI1 and is wrong signed in the case of CPI2.

As for the industrial production index, it is found that positive accumulated production growth increases the probability of changing prices upwards and negative accumulated production increases the probability of negative price changes both for CPI1 and CPI2. However, the effect of positive (negative) accumulated production growth on the probability of changing prices downwards (upwards) is not well identified. The difficulty in identifying those effects is likely to stem from the fact that we are using an economy-wide defined variable to measure demand pressure, when in fact a sector specific measure (if available) would be more adequate.[20]

[19] This corresponds to a competing risks model in which the risks are independent.

[20] We note that in the estimated models both for CPI1 and CPI2 the censored spells are added to each of the sub-samples defined for positive and negative price changes. This way the proportion of censored spells in the models estimated for each sub-sample is substantially larger than in the models estimated for the whole sample. This may also contribute to explain the difficulty in identifying the separate effects of industrial production growth on positive and negative price changes. The results of the estimated models are available from the authors upon request.

As for the effects of the magnitude of the last price change, it is seen that, as could be expected, a positive price change decreases the probability of changing prices upwards and increases the probability of changing prices downwards (more so in case of CPI2), while a negative change is associated with a larger probability of changing prices upwards and a smaller probability of changing prices downwards.

Overall, the sensitivity analysis conducted in this section allows us to conclude that, despite the differences found, the conclusions obtained from the analysis of the results in Table 1 are mostly unaffected by changes in the model specification.

4. CONCLUSIONS

This paper investigates the ability of time and state dependent price setting models to capture the main characteristics of price setting practices in Portugal. Two micro-datasets on consumer prices designed to produce the aggregate Consumer Price index are used.

By estimating a discrete time duration model, it is shown that the duration of price spells, and consequently the frequency of price changes, depends on variables such as accumulated sectoral inflation, the magnitude of the last price change and the level of demand (measured by the accumulated growth rate of the industrial production index).

In particular, it is found that the larger the accumulated sectoral inflation or the accumulated demand pressure, the larger the probability of price changes and thus, the smaller the expected duration of price spells. In contrast, the magnitude of the last price change has a negative effect on the expected duration of price spells. It is also found that the effects on the probability of changing prices differ according to whether we consider positive or negative values of the included time varying regressors. The robustness checks performed suggest that the main conclusions are mostly unaffected by changes in the model specification.

The main message of this paper is, therefore, that time dependent models are unable to adequately describe the features of the investigated datasets, and that state dependent models are required to fully characterise the price setting behaviour of Portuguese firms.

APPENDIX 1 - DATA DESCRIPTION

The two CPI datasets (CPI1 and CPI2) share a similar longitudinal structure but are collected using different criteria.

First, the composition of the datasets at the product level is determined using information on family expenditure patterns from the Portuguese Family Income and Expenditure Surveys. Two different surveys underlie the two datasets, thus introducing differences in composition by product between CPI1 and CPI2, although within each CPI the product composition is kept unchanged for the whole duration of the survey. However, the classification itself changed from CPI1 to CPI2 so that it is not possible to obtain a correspondence between the CPI1 and CPI2 products at the item level.

Second, the periodicity of data collection is product-dependent, varying between monthly, quarterly and yearly information. This means that some outlets are visited every month while others are only visited once a year. In CPI1 yearly, quarterly and monthly observations represent 1%, 51% and 48% of the consumer bundle while in CPI2 these proportions are, respectively, 4%, 58% and 38%. In dealing with such diversity, we need to standardize the time unit for comparison purposes. We start by excluding items observed on a yearly basis, because this information is too poor for the purpose of studying the price setting behaviour. Since in both datasets there are some items whose prices are collected monthly, whereas for other items prices are observed quarterly, in order to use data on all items, we have opted for transforming monthly data into quarterly data. This transformation was done by randomly selecting one month (first, second or third) in the quarter for each item and discarding the other two records for the entire observation period.

Third, the definition of price differs between CPIs due to the way sales and promotions are dealt with. Both datasets report retail prices at the moment the purchase occurs, but CPI1 excludes any sales or special prices being applied at the time of data collection, while CPI2 reports the effective price, including sales and promotions. Since CPI2 contains information about the occurrence of promotions/sales, we have decided to transform the CPI2 to be as similar as possible to CPI1 in what the treatment of sales/promotions is concerned.

Finally, missing observations are also treated differently in CPI1 and CPI2. Missings can occur either because the product is out-of-stock or the outlet is temporarily (or permanently) closed. In such cases, a price is generally reported in both datasets, although the procedure used to estimate it differs. CPI1 uses the last observed price as an estimate of the non-observed price. CPI2 uses an estimate of what the non-observed price would be, had it changed at the average rate of change observed in the remaining outlets. This procedure is applied for up to 3 consecutive periods. After this, the store is replaced if it remains closed (giving rise to an incomplete price trajectory) or the item is replaced by the most popular alternative within category and store, if it remains out-of-stock. As in the case of promotions/sales, CPI2 contains information about when the missing occurs. Such information was used to estimate non-observed prices in the same way as in CPI1, i.e., replacing non-observed prices by the last observed price. The products for which price trajectories are incomplete were discarded from CPI1 or CPI2 for estimation purposes.

APPENDIX 2 - SAMPLING ISSUES IN ESTIMATING HAZARD FUNCTIONS

The micro-datasets on prices compiled by national statistical offices are panels which have information on the prices of different products in different outlets or firms, recorded at regular intervals, for a fixed period of time.

In this Appendix we examine the complications that arise when these data are used to study price durations. To start with, it is important to notice that the first price that is observed for each unit is part of a spell that begun before the observation period was initiated. This means that, not only is this spell left-censored, but more importantly, it is an observation of the population of price spells elapsing when the observation starts. In other words, these observations are collected through a form of length biased sampling (see Cox, 1969) that, like the well known stock sampling (see Lancaster, 1990), tends to over-represent long spells.

To see why this is so, consider an individual for which the price spells have some finite mean. Assuming that the duration of price spells does not have a degenerate distribution, it is clear that *most of the time* the price set by this individual will be part of spells with duration above the mean. Therefore, if the price is observed at some random date, it is more likely that it will be part of a long price spell. If the duration of price spells varied across individuals but not for the same individual, this problem would disappear.

Unless the researcher is prepared to make strong assumptions about the distribution of price durations, the combination of the left-censoring with the stock sampling scheme makes the first observed price spell for each unit unusable. In what follows, we will assume that these observations are discarded. This will not lead to sample selection problems as long as at least one price change is observed for every unit.

After the first spell for each individual is concluded, the sampling scheme changes, as we are now able to observe the flow of spells until the end of the sampling period.[21] For each individual, the observations from the flow of spells are random draws from the distribution of the duration of price spells. Because each unit is observed until a given date, the number of spells that is observed varies across units and is endogenous in the sense that it depends on the average duration of price spells.

The estimation of the probability that a random price spell drawn from the population has duration d, is an important building block in the estimation of hazard functions. Assuming that the population is heterogeneous, this probability can be defined as

$$P\left(d\right)=\sum_{i}P\left(d|i\right)\mathcal{P}\left(i\right), \tag{A.1}$$

where $P\left(d|i\right)$ denotes the probability that a random price spell for unit i has duration d and $\mathcal{P}\left(i\right)$ is the probability of observing unit i in the population.

Assuming that the units in the sample are representative of the population of interest, $P\left(d\right)$ can be estimated taking $\mathcal{P}\left(i\right)$ as a constant. In this case, for a sample of N units, the analog of $P\left(d\right)$ is given by

$$\hat{P}\left(d\right)=\sum_{i=1}^{N}\frac{d_i}{s_i}\frac{1}{N}, \tag{A.2}$$

[21] The last spell is necessarily right-censored, but, for the moment, we will neglect this complication.

where d_i denotes the number of spells of duration d for unit i and s_i denotes the total number of spells for unit i. Since it is assumed that the units in the sample are representative of the population, this estimator gives equal weight to all units.

However, if $P(d)$ is estimated ignoring that different units have different numbers of spells, every spell is given the same weight. Consequently, the proportion of observations coming from each unit is $s_i / \sum_{i=1}^{N} s_i$ and the estimate of $P(d)$ is computed as

$$\hat{P}^*(d) = \frac{\sum_{i=1}^{N} d_i}{\sum_{i=1}^{N} s_i} = \sum_{i=1}^{N} \frac{d_i}{s_i} \frac{s_i}{\sum_{j=1}^{N} s_j}, \tag{A.3}$$

which is the sample analog of $P^*(d) = \sum_i P(d|i)\, \mathcal{P}^*(i)$, where $\mathcal{P}^*(i)$ denotes the probability that an observed spell comes from unit i. In a heterogeneous population, $\mathcal{P}(i)$ will be different from $\mathcal{P}^*(i)$ and therefore $\hat{P}^*(d)$ will be inconsistent for $P(d)$.[22]

To see how heterogeneity distorts the relation between $\mathcal{P}(i)$ and $\mathcal{P}^*(i)$, assume that there are only two units in the population with $\mathcal{P}(1) = \mathcal{P}(2) = 0.5$. Denoting by S_1 and S_2 the expected number of spells from each unit during a given time period, the expected number of spells in a representative sample collected over that period is

$$S = 0.5S_1 + 0.5S_2.$$

It is straightforward to see that the expected proportion of spells from unit 1 in the sample is

$$\mathcal{P}^*(1) = \frac{0.5S_1}{0.5S_1 + 0.5S_2} = \frac{S_1}{S_1 + S_2},$$

which is different from $\mathcal{P}(1) = 0.5$, unless $S_1 = S_2$.[23]

When the expected number of spells is constant for all units, $s_i \big/ \sum_{j=1}^{N} s_j \to 1/N$, and therefore $\hat{P}^*(d)$ will be a consistent estimator of $P(d)$. Furthermore, if the population is truly homogeneous in the sense that, not only $E(s_i|i) = E(s_i)$, but also $P(d|i) = P(d)$,

[22] Notice that, because under $P^*(i)$ the probability of observing unit i depends on the average length of its price spells, this distribution can be interpreted characterizing an artificial population resulting from a length biased sample from a population of units described by $P(i)$ (see Cox, 1969).

[23] This difference of representation in the sample, relatively to the population, will generate biases towards smaller (average) durations because units with more spells are over-represented. This over-representation problem becomes more severe as the ratio S_2/S_1 departs from 1.

any estimator of the form

$$\hat{P}^{w}(d)=\sum_{i=1}^{N}\frac{d_i}{s_i}\frac{w_i}{\sum_{j=1}^{N}w_j},\qquad w_i>0, \tag{A.4}$$

will be consistent for $P(d)$ since in this case the assumed distribution for i is irrelevant.

Naturally, even the hypothesis that $E(s_i|i)=E(s_i)$ is generally untenable and therefore estimating hazard functions using (A.3) in place of (A.2) can lead to very erroneous conclusions.

In some cases, it may be impractical to work with the entire sample and the researchers may want to draw a sub-sample to perform their analysis. A simple way to obtain a representative sub-sample is to randomly select $n \leq N$ units, giving each unit in the sample probability $1/N$, and then using a fixed number of spells, say J, for each unit in the sub-sample. In this case, because J is fixed across units, $P(d)$ can be estimated as

$$\tilde{P}(d)=\sum_{i=1}^{n}\frac{d_i}{J}, \tag{A.5}$$

where now d_i denotes the number of spells of duration d for unit i, out of the J spells considered. This is the approach we follow in sub-section 3.2, setting $J=1$.

Naturally, because it does not use all available information, (A.5) will be less efficient than (A.2).[24] However, with uncensored samples, (A.5) has the advantage of being an unbiased estimator of $P(d)$, rather than just a consistent one.

[24] Notice that this estimator is a special case both of (A.2) and (A.3), which shows that if the sampling scheme was such that the number of spells per unit was exogenous, the sampling problems discussed here would not arise.

APPENDIX 3 - OTHER ESTIMATION RESULTS

Table A1 - Estimation results for the controls

	CPI1		CPI2	
	Coef.	t-stat	Coef.	t-stat
Big store	0.417	36.033	—	—
Supermarket	—	—	−0.333	−15.761
Discount Store	—	—	−0.572	−5.558
Classical Store	—	—	−0.880	−41.399
Market	—	—	−0.181	−6.904
Private/Public	—	—	−0.969	−26.327
Processed Food	-0.503	-37.343	−0.317	−18.297
Non-energetic industrial goods	-0.860	-60.959	—	—
Non-durable goods	—	—	−0.549	−25.869
Semi-durable goods	—	—	−0.537	−19.466
Durable goods	—	—	−0.569	−25.097
Non administered price services	-0.936	-45.032	−1.044	−30.465
North	-0.086	-6.548	−0.161	−11.060
Center	-0.171	-11.172	−0.050	−2.846
Alentejo	-0.268	-14.770	0.101	5.120
Algarve	-0.430	-20.071	−0.312	−13.788
Azores	-0.050	-1.824	0.020	0.827
Madeira	-0.097	-2.535	−0.341	−9.575
dur_1	-0.082	-3.958	0.138	5.099
dur_2	-0.619	-26.316	−0.262	−8.861
dur_3	-0.814	-30.080	−0.448	−13.938
dur_4	-0.662	-22.727	−0.319	−9.685
dur_5	-1.088	-30.254	−0.667	−17.472
dur_6	-1.280	-30.419	−0.858	−19.822
dur_7	-1.257	-27.556	−0.999	−20.612
dur_8	-1.051	-22.200	−0.958	−18.369
dur_9	-1.499	-24.761	−1.101	−17.926
dur_{10}	-1.583	-23.177	−1.281	−17.776
dur_{11}	-1.590	-20.836	−1.163	−15.125
dur_{12}	-1.346	-17.656	−1.240	−13.719
dur_{13}	-1.569	-16.827	−1.511	−12.000
dur_{14}	-1.803	-15.659	−1.675	−10.686
dur_{15}	-1.660	-13.594	−2.736	−8.964
dur_{16}	-1.744	-12.098	—	—
dur_{17}	-2.070	-10.534	—	—
dur_{18}	-2.309	-8.859	—	—
dur_{19}	-1.983	-7.615	—	—
dur_{20}	-2.491	-7.393	—	—
2nd quarter	-0.099	-6.108	−0.123	−7.440
3nd quarter	-0.012	-0.795	0.026	1.631
4nd quarter	-0.376	-23.406	−0.232	−14.379
1992	0.110	4.208	—	—
1994	-0.094	-5.259	—	—
1995	-0.058	-3.470	—	—
1996	0.051	2.393	—	—
1997	-0.489	-27.519	0.202	10.912
1999	—	—	−0.097	−6.653
2000	—	—	−0.481	−29.521

REFERENCES

Aucremmane, L., and Dhyne E., (2005), "Time-dependent versus state-dependent pricing: A panel data approach to the determimants of Belgian consumer price changes", European Central Bank, W.P. No. 462;

Bover, O., Arellano, M., and Bentolila, S. (2002), “Unemployment duration, benefit duration and the business cycle”, *The Economic Journal*, 112, 223-265.

Calvo, G.A. (1983), “Staggered Prices in a Utility-Maximizing Framework,” *Journal of Monetary Economics*, 12, 383-98.

Cecchetti, S. (1986), “The frequency of price adjustment, A study of the newsstand prices of magazines”, *Journal of Econometrics*, 31, 255-274.

Cox, D.R. (1969), “Some Sampling Problems in Technology”, *in* Johnson, N.L. and Smith, H. (eds.) *New Developments in Survey Sampling*, Wiley, New York.

Dias, M., Dias, D., and Neves, P.D. (2004), “Stylised features of price setting behaviour in Portugal: 1992-2001”, European Central Bank, W.P. No. 332.

Dotsey, M., King, R. and Wolman, A. (1999), “State Dependent Pricing and the General Equilibrium Dynamics of Money and Output”, *Quarterly Journal of Economics*, 114, 655-690.

Fougère, D., Le Bihan H., and Sevestre P., (2005), “Heterogeneity in Price Stickiness: Microeconometric Investigation”, mimeo;

Heckman, J. and Singer, B. (1984), “Econometric Duration Analysis”, *Journal of Econometrics*, Annals 1984-1, 63-132.

Jenkins, S. P. (1995), “Easy estimation methods for discrete-time duration models”, *Oxford Bulletin of Economics and Statistics*, 57, 129-138.

Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Econometric Society Monographs, 17, Cambridge University Press.

Portugal, P., Addison, J. T. (2000), “Short and long-term unemployment”, *Economics Letters*, 66, 107-112.

Prentice, R. and Gloeckler J. (1978), “Regression analysis of grouped survival data with application to breast cancer data”, Biometrics, 34, 57-67;

Ridder, G. (1987), “The sensitivity of duration models to misspecified unobserved heterogeneity and duration dependence”, Working Paper, University of Groningen.

Sheshinski E., Weiss, Y. (1977), "Inflation and costs of adjustment", *The Review of Economic Studies*, 44, 287-303.

Sheshinski E., Weiss, Y. (1983), "Optimum pricing policy under stochastic inflation", *The Review of Economic Studies*, L, 513-529.

Taylor, John B. (1980). "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy*, 8, 1-23.

European Central Bank working paper series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (http://www.ecb.int)

482 "Forecasting macroeconomic variables for the new member states of the European Union" by A. Banerjee, M. Marcellino and I. Masten, May 2005.

483 "Money supply and the implementation of interest rate targets" by A. Schabert, May 2005.

484 "Fiscal federalism and public inputs provision: vertical externalities matter" by D. Martínez-López, May 2005.

485 "Corporate investment and cash flow sensitivity: what drives the relationship?" by P. Mizen and P. Vermeulen, May 2005.

486 "What drives productivity growth in the new EU member states? The case of Poland" by M. Kolasa, May 2005.

487 "Computing second-order-accurate solutions for rational expectation models using linear solution methods" by G. Lombardo and A. Sutherland, May 2005.

488 "Communication and decision-making by central bank committees: different strategies, same effectiveness?" by M. Ehrmann and M. Fratzscher, May 2005.

489 "Persistence and nominal inertia in a generalized Taylor economy: how longer contracts dominate shorter contracts" by H. Dixon and E. Kara, May 2005.

490 "Unions, wage setting and monetary policy uncertainty" by H. P. Grüner, B. Hayo and C. Hefeker, June 2005.

491 "On the fit and forecasting performance of New-Keynesian models" by M. Del Negro, F. Schorfheide, F. Smets and R. Wouters, June 2005.

492 "Experimental evidence on the persistence of output and inflation" by K. Adam, June 2005.

493 "Optimal research in financial markets with heterogeneous private information: a rational expectations model" by K. Tinn, June 2005.

494 "Cross-country efficiency of secondary education provision: a semi-parametric analysis with non-discretionary inputs" by A. Afonso and M. St. Aubyn, June 2005.

495 "Measuring inflation persistence: a structural time series approach" by M. Dossche and G. Everaert, June 2005.

496 "Estimates of the open economy New Keynesian Phillips curve for euro area countries" by F. Rumler, June 2005.

497 "Early-warning tools to forecast general government deficit in the euro area: the role of intra-annual fiscal indicators" by J. J. Pérez, June 2005.

498 "Financial integration and entrepreneurial activity: evidence from foreign bank entry in emerging markets" by M. Giannetti and S. Ongena, June 2005.

499 “A trend-cycle(-season) filter” by M. Mohr, July 2005.

500 “Fleshing out the monetary transmission mechanism: output composition and the role of financial frictions” by A. Meier and G. J. Müller, July 2005.

501 “Measuring comovements by regression quantiles” by L. Cappiello, B. Gérard, and S. Manganelli, July 2005.

502 “Fiscal and monetary rules for a currency union” by A. Ferrero, July 2005

503 “World trade and global integration in production processes: a re-assessment of import demand equations” by R. Barrell and S. Dées, July 2005.

504 “Monetary policy predictability in the euro area: an international comparison” by B.-R. Wilhelmsen and A. Zaghini, July 2005.

505 “Public good issues in TARGET: natural monopoly, scale economies, network effects and cost allocation” by W. Bolt and D. Humphrey, July 2005.

506 “Settlement finality as a public good in large-value payment systems” by H. Pagès and D. Humphrey, July 2005.

507 “Incorporating a “public good factor” into the pricing of large-value payment systems” by C. Holthausen and J.-C. Rochet, July 2005.

508 “Systemic risk in alternative payment system designs” by P. Galos and K. Soramäki, July 2005.

509 “Productivity shocks, budget deficits and the current account” by M. Bussière, M. Fratzscher and G. J. Müller, August 2005.

510 “Factor analysis in a New-Keynesian model” by A. Beyer, R. E. A. Farmer, J. Henry and M. Marcellino, August 2005.

511 “Time or state dependent price setting rules? Evidence from Portuguese micro data” by D. A. Dias, C. R. Marques and J. M. C. Santos Silva, August 2005.