Carvalho, Carlos; Dam, Niels A.; Lee, Jae Won

Working Paper

Real Rigidities and the Cross-Sectional Distribution of Price Stickiness: Evidence from Micro and Macro Data Combined

Texto para discussão, No. 634

Provided in Cooperation with:
Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro

Suggested Citation: Carvalho, Carlos; Dam, Niels A.; Lee, Jae Won (2014) : Real Rigidities and the Cross-Sectional Distribution of Price Stickiness: Evidence from Micro and Macro Data Combined, Texto para discussão, No. 634, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Departamento de Economia, Rio de Janeiro

This Version is available at:
https://hdl.handle.net/10419/176117

TEXTO PARA DISCUSSÃO

No. 634

Real Rigidities and the Cross-Sectional Distribution of Price Stickiness: Evidence from Micro and Macro Data Combined

Carlos Carvalho
Niels A. Dam
Jae Won Lee

DEPARTAMENTO DE ECONOMIA
www.econ.puc-rio.br

Real Rigidities and the Cross-Sectional Distribution of Price Stickiness: Evidence from Micro and Macro Data Combined*

Carlos Carvalho
PUC-Rio

Niels Arne Dam
Danmarks Nationalbank

Jae Won Lee
Seoul National University

December, 2014

Abstract

We use a standard sticky-price model to provide evidence on three mechanisms that can reconcile somewhat frequent price changes with large and persistent real effects of monetary shocks. To that end, we estimate a semi-structural model for the U.S. economy that allows for varying degrees of real rigidities, and cross-sectional heterogeneity in price stickiness. The model can extract some information about these two features of the economy from aggregate data, and discriminate between different distributions of price stickiness. Hence it can also speak to the debate about the role of sales and other temporary price changes in shaping the effects of monetary policy. Employing a Bayesian approach, we combine macroeconomic time-series data with information about empirical distributions of price stickiness derived from micro price data for the U.S. economy. Our estimates point to the presence of both large real rigidities and an important degree of heterogeneity in price stickiness. Moreover, cross-sectional distributions of price stickiness that factor out sales improve the empirical fit of the model. Our results suggest that bridging the gap between micro and macro evidence on nominal price rigidity may require the combination of several mechanisms.

JEL classification codes: E10, E30

Keywords: real rigidities, heterogeneity in price stickiness, sales, regular prices, micro data, macro data, Bayesian estimation

*This paper grew out of "The Cross-Sectional Distribution of Price Stickiness Implied by Aggregate Data" (Carvalho and Dam 2010). For helpful comments and suggestions we would like to thank Luis Álvarez, Marco Del Negro, Stefano Eusepi, Henrik Jensen, Ed Knotek, Oleksyi Kryvtsov, John Leahy, Frank Schorfheide, an anonymous referee, and, especially, Giorgio Primiceri. We also thank seminar participants at the ESWC 2010, AEA 2010, Banque de France conference "Understanding price dynamics: recent advances", ESEM 2009, SED 2009, NAMES 2009, Bank of Canada, ECB - WDN Workshop, Rutgers University, LAMES 2008, Danmarks Nationalbank, University of Copenhagen, Riksbank, and NY Fed. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Danmarks Nationalbank. E-mails: cvianac@econ.puc-rio.br, nad@nationalbanken.dk, jwlee7@snu.ac.kr.

1 Introduction

Our understanding of the real effects of monetary policy hinges, to a large extent, on the existence of some degree of nominal price rigidity. In the decade since the publication of the seminal Bils and Klenow (2004) paper, the availability of large amounts of micro price data has rekindled interest in this area, and allowed us to make progress. Yet, estimates of the extent of nominal price stickiness based on microeconomic data versus those based on aggregate data usually produce a conflicting picture.

According to Klenow and Malin's (2011) survey of the empirical literature based on micro data, prices change, on average, at least once a year – somewhat more often than we thought was the case prior to Bils and Klenow (2004). In contrast, making sense of estimates of the response of the aggregate price level to monetary shocks (from dynamic stochastic general equilibrium – DSGE – models, or vector autoregressions) requires much less frequent price adjustments.[1]

If nominal price rigidities are to continue to be the leading explanation for why monetary policy has large and persistent real effects, it is important that we deepen our understanding of mechanisms that can narrow the gap between the evidence of somewhat flexible individual prices and relatively sluggish aggregate prices – i.e., mechanisms that can produce a large "contract multiplier". If prices change frequently and each and every price change contributes to fully offset nominal disturbances, then nominal price rigidity cannot be the source of large and persistent monetary non-neutralities. Hence, a large contract multiplier requires that price adjustments, somehow, fail to perfectly neutralize monetary innovations.

In this paper, our goal is to contribute to bridge the gap between micro and macro evidence on the extent of nominal price rigidity. To that end, we estimate a standard macroeconomic model of price setting, and use it to speak to three mechanisms that can boost the contract multiplier. The first such mechanism are so-called "real rigidities", in the sense of Ball and Romer (1990). Large real rigidities reduce the sensitivity of individual prices to aggregate demand conditions, and thus serve as a source of endogenous persistence: for any given degree of price stickiness, partial adjustment of individual prices makes for a sluggish response of the aggregate price level to monetary shocks.

The other two mechanisms are motivated by the empirical evidence uncovered since Bils and Klenow (2004), and subsequent theoretical literature. Cross-sectional heterogeneity in price rigidity, to the extent documented in the micro data, can lead to much larger monetary non-neutralities than the average frequency of price changes would imply (Carvalho 2006, Nakamura and Steinsson 2010). The reason is that, while recurrent price changes by firms in more flexible sectors do not contribute as much

[1] See, for example, the survey by Maćkowiak and Smets (2008).

to offset monetary shocks, they do count for the frequency of price adjustment.[2] Heterogeneity can become an even more powerful mechanism when coupled with strategic complementarities in pricing decisions. In those circumstances, firms in the more sticky sectors become disproportionately important in shaping aggregate dynamics (relative to their sectoral weight), through their influence on pricing decisions of firms that change prices more frequently (Carvalho 2006).[3]

The third mechanism is the presence of sales and other temporary price changes. Guimaraes and Sheedy (2011) and Kehoe and Midrigan (2014) show that such price changes may help reconcile frequent micro adjustments with a sluggish aggregate price response to nominal disturbances.[4] A basic intuition for their results is that temporary price changes fail to offset monetary shocks well, since these shocks tend to induce permanent changes in the level of prices.[5]

We estimate a macroeconomic model that, while relatively standard, can provide some information about the three aforementioned mechanisms based on aggregate data. The price-setting block of the model is a multi-sector sticky-price economy that allows for heterogeneity in price stickiness, and can feature strategic complementarity or substitutability in pricing decisions. The remaining equations specify exogenous stochastic processes that drive firms' frictionless optimal prices. They provide the model with some flexibility to perform in empirical terms, and thus allow us to focus on the objects of interest in the price-setting block of the economy. Hence, we refer to our model as "semi-structural".[6]

We show that, at least in theory, the model is able to separately identify real and nominal rigidities, and tell apart economies with homogenous from those with heterogeneous price stickiness – based on aggregate data only. The model can also discriminate between different (non-degenerate) distributions of price rigidity, providing information on which one helps explain aggregate dynamics better. Hence, at least in theory, our analysis can also speak to the debate about how to treat temporary sales in micro price data.

[2] Carvalho and Schwartzman (2014) show how this intuition can be formalized in terms of a "selection effect" relative to the timing of price changes, which arises in the class of time-dependent pricing models.

[3] Nakamura and Steinsson (2010) conclude that this interaction is not important in their calibrated menu-cost model.

[4] Coibion et al. (2014) provide evidence that sales are essentially acyclical – which is consistent with the models in Guimaraes and Sheedy (2011) and Kehoe and Midrigan (2014). Kryvtsov and Vincent (2014), on the other hand, argue that sales do not help reconcile micro and macro evidence on price rigidity. They document a large degree of cyclicality of sales in the U.K. micro data, and develop a model that can explain their findings. In bad times, consumers intensify search for bargain prices and firms increase the frequency of sales. This "complementarity" between search effort and sales frequency breaks down the strategic substitutability of sales that would otherwise arise (as in Guimaraes and Sheedy 2011), and leads to cyclical sales.

[5] Information frictions can also lead to large contract multipliers. Not surprisingly, that literature picked up steam after the empirical literature based on micro price data flourished. Classic contributions include Caballero (1989), Reis (2006), and Maćkowiak and Wiederholt (2009), who obtain large monetary non-neutralities in models with information frictions in which prices change continuously. More recently, Bonomo, Carvalho, Garcia, and Malta (2014) obtain a large contract multiplier in an estimated model with menu costs and partially costly information.

[6] Several earlier papers in the literature combine structural equations with empirical specifications for other parts of the model (e.g., Sbordone 2002, Guerrieri 2006, and Coenen et al. 2007).

Identification of those features of price setting based on aggregate data is possible in our model because sectors that differ in price stickiness have different implications for the response of the macroeconomy to shocks at different frequencies. In particular, sectors where prices are more sticky are relatively more important in determining the low-frequency response to shocks; and vice-versa for more flexible sectors. These differences provide information about the cross-sectional distribution of price stickiness. Finally, separate identification of real and nominal rigidities comes from the way in which the aggregate price level depends on its own lags vis-a-vis lags of exogenous drivers of firms' frictionless optimal prices.

While our approach requires that we learn something about the mechanisms of interest from aggregate data, we argue that a more promising direction is to combine those data with information about the empirical cross-sectional distribution of price stickiness derived from micro price data.

In their favor, the micro data have the millions of observations used to compute measures of price rigidity. They also allow us to estimate separate distributions of price rigidity, with and without sales. Hence, on these grounds, one could imagine imposing alternative empirical distributions of price rigidity derived from micro data (e.g., with and without sales), estimating the other parameters of the model – in particular, the parameter associated with real rigidities – and comparing the performance of different estimated models in terms of fit and other dimensions of interest.

However, treating the estimates derived from micro data as the true "population moments" that matter for aggregate dynamics is not appropriate, in our view. First, it is possible that some price adjustments do not convey as much information about changes in macroeconomic conditions as others do. While this possibility is at the core of the debate about whether or not to exclude sales from price setting statistics for macro purposes, the argument applies more generally – for example, it also applies to the literature on price setting under information frictions. In that case, macro-based estimates should convey useful information about the price changes that do matter for aggregate dynamics. Second, and not less importantly, Eichenbaum et al. (2014) show that the BLS micro data underlying the CPI are plagued with measurement problems when it comes to computing statistics based on individual price changes. While Eichenbaum et al. (2014) focus on pitfalls involved in estimating the distribution of the size of price changes, the problems they document certainly add measurement error to available estimates of the cross-sectional distribution of price stickiness that use that data (e.g., Bils and Klenow 2004, Nakamura and Steinsson 2008, Klenow and Kryvtsov 2008).

Hence, we find it valuable to "let the aggregate data speak" about the cross-sectional distribution of price rigidity – something that our model can accomplish. At the same time, while we do not want to treat the statistics derived from micro data as error-free, we certainly do not want to ignore all the

information that they contain.

To strike a balance between extracting information from aggregate data and exploring what we know based on the micro data, we employ a full-information Bayesian approach. We use aggregate (time-series) data on nominal and real Gross Domestic Product (GDP) as observables, and incorporate the microeconomic information about the cross-sectional distribution of price stickiness through our prior on the model parameters that govern that distribution. Specifically, we parameterize that prior in a way that easily allows us to relate its moments to the analogous moments of empirical distributions of price stickiness. We focus on two empirical distributions: one that takes into account all price changes, including sales (derived from Bils and Klenow 2004); and another, based on "regular" price changes, that excludes sales and product substitutions (derived from Nakamura and Steinsson 2008).

We summarize our findings as follows. The estimated models can discriminate quite sharply between economies with heterogeneity in price stickiness and their homogeneous-firms counterparts. They also point to the existence of large real rigidities, which induce strong strategic complementarities in price setting.

Turning to the cross-sectional distribution of price rigidity, the information extracted from aggregate data accords quite well with the micro data. Specifically, the distribution estimated under an uninformative ("flat") prior has a correlation of 0.43 with the distribution that leaves sales in (Bils and Klenow 2004), and a correlation of 0.63 with the distribution that excludes sales and product substitutions (Nakamura and Steinsson 2008). Moreover, a formal statistical comparison between models with informative priors based on those two cross-sectional distributions provides some additional evidence in favor of the distribution based on regular prices.

Altogether, our results suggest that all three mechanisms that can boost the contract multiplier might have a role to play in our understanding of the effects of monetary policy.

1.1 Brief literature review

Our work is related to the literature that emphasizes the importance of heterogeneity in price rigidity for aggregate dynamics. However, our focus differs from that of existing papers. Most of the latter focus on the role of heterogeneity in boosting the contract multiplier in calibrated models (e.g., Carvalho 2006, Carvalho and Schwartzman 2008, Nakamura and Steinsson 2010, Carvalho and Nechio 2010, Dixon and Kara 2011). These papers do not address the question of whether such heterogeneity does in fact help sticky-price models fit the data better according to formal statistical criteria.

In terms of empirical work on the importance of heterogeneity in price stickiness, Imbs et al. (2011) study the aggregation of sectoral Phillips curves, and the statistical biases that can arise from using

estimation methods that do not account for heterogeneity. They rely on sectoral data for France, and find that the results based on estimators that allow for heterogeneity are more in line with the available microeconomic evidence on price rigidity. Lee (2009) and Bouakez et al. (2009) estimate multi-sector DSGE models with heterogeneity in price rigidity using aggregate and sectoral data. They also find results that are more in line with the microeconomic evidence than the versions of their models that impose the same degree of price rigidity for all sectors.[7] Taylor (1993) provides estimates of the distribution of the duration of wage contracts in various countries inferred solely from aggregate data, while Guerrieri (2006) provides estimates of the distribution of the duration of price spells in the U.S. based on aggregate data. Both models feature ex-post rather than ex-ante heterogeneity in nominal rigidities, as is the case in our model.[8] Coenen et al. (2007) estimate a model with (limited) ex-ante heterogeneity in *price contracts* using only aggregate data. They focus on the estimate of the Ball-Romer index of real rigidities and on the average duration of contracts implied by their estimates, which they emphasize is in line with the results in Bils and Klenow (2004).[9]

Jadresic (1999) is a precursor to some of the ideas in this paper. He estimates a model with ex-ante heterogeneous price spells using only aggregate data for the U.S. economy to study the joint dynamics of output and inflation. Similarly to our findings, his statistical results reject the assumption of identical firms. Moreover, he discusses the intuition behind the source of identification of the cross-sectional distribution of price rigidity from aggregate data in his model, which is the same as in our model. Despite these similarities, our paper differs from Jadresic's in several important dimensions. We use a different estimation method, and show the possibility of extracting information about the cross-sectional distribution of price rigidity from aggregate data in a more general context - in particular in the presence of pricing complementarities. Most importantly, the focus of our paper goes beyond assessing the empirical support for heterogeneity in price rigidity from aggregate data. We also investigate the similarities between our macro-based estimates and the available microeconomic evidence, and propose an approach to integrate the two sources of information on the distribution of price rigidity.

Finally, our results speak to the ongoing debate on the role of sales in macroeconomic models. That literature started out as a discussion about whether or not to exclude sales when computing price-setting statistics for macro purposes (Bils and Klenow 2004, Nakamura and Steinsson 2008, Klenow and Kryvtsov 2008). This initial debate was followed by a theoretical literature that provided macro-

[7] Bouakez et al. (2014) find similar results in an extension of their earlier paper to a larger number of sectors.

[8] Their frameworks are thus closer to the generalized time-dependent model of Dotsey et al. (1997) than to our model with ex-ante heterogeneity.

[9] Their estimated model features indexation to an average of past inflation and a (non-zero) constant inflation objective. Thus, strictly speaking their finding is that the average time between "contract reoptimizations" is comparable to the average duration of price spells documented by Bils and Klenow (2004).

economic models with sales and other temporary price changes (Guimaraes and Sheedy 2011, Kehoe and Midrigan 2014). More recently, the literature has focused on the cyclicality of sales and consumer behavior, both in theory and in the micro data (e.g., Coibion et al. 2014, Kryvtsov and Vincent 2014). We provide statistical evidence on the relative performance of macroeconomic models with different distributions of price rigidity that do and do not exclude sales (and product substitutions).

In Section 2 we present the semi-structural model and study the extent to which aggregate data contain information about the cross-sectional distribution of price stickiness and the parameter that controls the extent of real rigidities in the model. Section 3 describes our empirical methodology and data. In Section 4 we present our main results. We start with macro-based estimates obtained under an uninformative prior, assess the estimates against the empirical distributions from Bils and Klenow (2004) and Nakamura and Steinsson (2008), and perform model comparison with specifications that impose the same degree of price rigidity for all firms. We then provide our estimates that incorporate information from the micro data, and perform model comparison with different prior distributions. Section 6 reports our robustness analysis, and discusses the performance of the estimated model in light of additional micro price facts. The last section concludes.

2 The semi-structural model

There is a continuum of monopolistically competitive firms divided into K sectors that differ in the frequency of price changes. Firms are indexed by their sector, $k \in \{1, ..., K\}$, and by $j \in [0, 1]$. The distribution of firms across sectors is summarized by a vector $(\omega_1, ..., \omega_K)$ with $\omega_k > 0$, $\sum_{k=1}^{K} \omega_k = 1$, where ω_k gives the mass of firms in sector k. Each firm produces a unique variety of a consumption good, and faces a demand that depends negatively on its relative price.

In any given period, profits of firm j from sector k (henceforth referred to as "firm kj") are given by:

$$\Pi_t(k,j) = P_t(k,j)\, Y_t(k,j) - C\left(Y_t(k,j), Y_t, \xi_t\right),$$

where $P_t(k,j)$ is the price *charged* by the firm, $Y_t(k,j)$ is the quantity that it sells at the posted price (determined by demand), and $C\left(Y_t(k,j), Y_t, \xi_t\right)$ is the total cost of producing such quantity, which may also depend on aggregate output Y_t, and is subject to shocks (ξ_t). We assume that the demand faced by the firm depends on its relative price $\frac{P_t(k,j)}{P_t}$, where P_t is the aggregate price level in the economy, and on aggregate output. Thus, we write firm kj's profit as:

$$\Pi_t(k,j) = \Pi\left(P_t(k,j), P_t, Y_t, \xi_t\right),$$

and make the usual assumption that Π is homogeneous of degree one in the first two arguments, and single-peaked at a strictly positive level of $P_t(k,j)$ for any level of the other arguments.[10]

The aggregate price index combines sectoral price indices, $P_t(k)$'s, according to the sectoral weights, ω_k's:

$$P_t = \Gamma\left(\{P_t(k), \omega_k\}_{k=1,...,K}\right),$$

where Γ is an aggregator that is homogeneous of degree one in the $P_t(k)$'s. In turn, the sectoral price indices are obtained by applying a symmetric, homogeneous-of-degree-one aggregator Λ to prices charged by all firms in each sector:

$$P_t(k) = \Lambda\left(\{P_t(k,j)\}_{j\in[0,1]}\right).$$

We assume the specification of staggered price setting inspired by Taylor (1979, 1980). Firms set prices that remain in place for a fixed number of periods. The latter is sector-specific, and we save on notation by assuming that firms in sector k set prices for k periods. Thus, $\omega = (\omega_1, ..., \omega_K)$ fully characterizes the cross-sectional distribution of price stickiness that we are interested in. Finally, across all sectors, adjustments are staggered uniformly over time.

Before we continue, a brief digression about the Taylor pricing model is in order. As will become clear, this model allows us to tell apart real rigidities from nominal rigidities, and to infer the cross-sectional distribution of price stickiness implied by aggregate data. Hence, it serves our purposes well. However, strictly speaking, that model is at odds with the microeconomic evidence on the duration of price spells. Klenow and Kryvtsov (2008), for example, provide evidence that the duration of individual price spells varies at the *quote line* level. However, this evidence does not invalidate the use of the Taylor model for our purposes. In particular, in Section 5 we provide an alternative model in which the duration of price spells varies at the firm level, and yet the aggregate behavior of the model is identical to the one presented here. The alternative model can match additional micro facts documented in the literature. Hence, it provides a cautionary note on attempts to test specific models of price setting by confronting them with descriptive micro price statistics. For ease of exposition, we proceed with the standard Taylor pricing specification. But the reader should keep in mind that the aggregate implications that we are interested survive in models that can match the microeconomic evidence in many dimensions.

[10] This is analogous to Assumption 3.1 in Woodford (2003).

When setting its price $X_t(k,j)$ at time t, given that it sets prices for k periods, firm kj solves:

$$\max E_t \sum_{i=0}^{k-1} Q_{t,t+i} \Pi\left(X_t(k,j), P_{t+i}, Y_{t+i}, \xi_{t+i}\right),$$

where $Q_{t,t+i}$ is a (possibly stochastic) nominal discount factor. The first-order condition for the firm's problem can be written as:

$$E_t \sum_{i=0}^{k-1} Q_{t,t+i} \frac{\partial \Pi\left(X_t(k,j), P_{t+i}, Y_{t+i}, \xi_{t+i}\right)}{\partial X_t(k,j)} = 0. \tag{1}$$

Note that all firms from sector k that adjust prices at the same time choose a common price, which we denote $X_t(k)$.[11] Thus, for a firm kj that adjusts at time t and sets $X_t(k)$, the prices charged from t to $t+k-1$ satisfy:

$$P_{t+k-1}(k,j) = P_{t+k-2}(k,j) = ... = P_t(k,j) = X_t(k).$$

Given the assumptions on price setting, and uniform staggering of price adjustments, with an abuse of notation sectoral prices can be expressed as:

$$P_t(k) = \Lambda\left(\{X_{t-i}(k)\}_{i=0,...,k-1}\right).$$

Instead of postulating a fully specified model to obtain the remaining equations to be used in the estimation, we assume exogenous stochastic processes for nominal output ($M_t \equiv P_t Y_t$) and for the unobservable ξ_t process; hence, we refer to our model as semi-structural. Given our focus on estimation of parameters that characterize price-setting behavior in the economy in the presence of heterogeneity, our goal in specifying such exogenous time-series processes is to close the model with a set of equations that can provide it with flexibility relative to a fully-structural model. Such flexibility is useful because it allows us to draw conclusions about price setting that do not depend on details of structural models that are not the focus of our analysis.[12]

2.1 A loglinear approximation

We assume that the economy has a deterministic zero-inflation steady state characterized by $M_t = \overline{M}, \xi_t = \overline{\xi}, Y_t = \overline{Y}, Q_{t,t+i} = \beta^i$, and for all (k,j), $X_t(k,j) = P_t = \overline{P}$, and loglinearize (1) around it to

[11] In Section 5.2 we discuss how the model can be enriched with idiosyncratic shocks that can help it match some micro facts about the size of price changes without affecting any of its aggregate implications.

[12] Needless to say, the results are conditional on the particular model of price setting that we adopt. In Section 5 we discuss the extent to which our conclusions may generalize to alternative price-setting specifications.

obtain:[13]

$$x_t(k) = \frac{1-\beta}{1-\beta^k} E_t \sum_{i=0}^{k-1} \beta^i \left(p_{t+i} + \zeta \left(y_{t+i} - y_{t+i}^n \right) \right), \tag{2}$$

where lowercase variables denote log-deviations of the respective uppercase variables from the steady state. The parameter $\zeta > 0$ is the Ball and Romer (1990) index of real rigidities. The new variable Y_t^n is defined implicitly as a function of ξ_t by:

$$\left. \frac{\partial \Pi \left(X_t(k,j), P_t, Y_t^n, \xi_t \right)}{\partial X_t(k,j)} \right|_{X_t(k,j)=P_t} = 0.$$

In the loglinear approximation, y_t^n moves proportionately to $\log\left(\xi_t/\overline{\xi}\right)$. Strictly speaking, it is the level of output that would prevail in a flexible-price economy. In a fully specified model this would tie it down to preference and productivity shocks. Here we do not pursue a structural interpretation of the exogenous processes driving the economy.[14] Nevertheless, for ease of presentation we follow the literature and label y_t^n the "natural level of output."

The definition of nominal output yields:

$$m_t = p_t + y_t. \tag{3}$$

Finally, we postulate that the aggregators that define the overall level of prices P_t and the sectoral price indices give rise to the following loglinear approximations:[15]

$$p_t = \sum_{k=1}^{K} \omega_k p_t(k), \tag{4}$$

$$p_t(k) = \int_0^1 p_t(k,j)\, dj = \frac{1}{k} \sum_{j=0}^{k-1} x_{t-j}(k). \tag{5}$$

Large real rigidities (small ζ in equation (2)) reduce the sensitivity of prices to aggregate demand conditions, and thus magnify the non-neutralities generated by nominal price rigidity. In fully specified models, the extent of real rigidities depends on primitive parameters such as the intertemporal elasticity of substitution, the elasticity of substitution between varieties of the consumption good, the labor supply elasticity. It also depends on whether the economy features economy-wide or segmented factor markets,

[13] We write all such approximations as equalities, ignoring higher-order terms.

[14] We think such an interpretation is unreasonable because we take nominal output to be exogenous. In that context, an interpretation of y_t^n as being driven by preference and technology shocks would imply that these shocks have no effect on nominal output (i.e., that they have exactly offsetting effects on aggregate output and prices).

[15] This is what comes out of a fully-specified multi-sector model with the usual assumption of Dixit-Stiglitz preferences.

whether there is an explicit input-output structure etc.[16]

In the context of our model, ζ is itself a primitive parameter.[17] Following standard practice in the literature, we refer to economies with $\zeta < 1$ as displaying *strategic complementarities* in price setting. To clarify the meaning of this expression, replace (3) in (2) to obtain:

$$x_t(k) = \frac{1-\beta}{1-\beta^k} E_t \sum_{i=0}^{k-1} \beta^i \left(\zeta \left(m_{t+i} - y^n_{t+i} \right) + (1-\zeta) p_{t+i} \right). \tag{6}$$

That is, new prices are set as a discounted weighted average of current and expected future driving variables $\left(m_{t+i} - y^n_{t+i}\right)$ and prices p_{t+i}. $\zeta < 1$ implies that firms choose to set higher prices if the overall level of current and expected future prices is higher, all else equal. On the other hand, $\zeta > 1$ means that prices are *strategic substitutes*, in that under those same circumstances adjusting firms choose relatively lower prices.

2.2 Nominal (m_t) and natural (y^n_t) output

We postulate an $AR(p_1)$ process for nominal output, m_t:

$$m_t = \rho_0 + \rho_1 m_{t-1} + ... + \rho_{p1} m_{t-p_1} + \varepsilon^m_t, \tag{7}$$

and an $AR(p_2)$ process for the natural output level, y^n_t:

$$y^n_t = \delta_0 + \delta_1 y^n_{t-1} + ... + \delta_{p_2} y^n_{t-p_2} + \varepsilon^n_t, \tag{8}$$

where $\varepsilon_t = (\varepsilon^m_t, \varepsilon^n_t)$ is i.i.d. $N\left(0_{1\times 2}, \Omega^2\right)$, with $\Omega^2 = \begin{bmatrix} \sigma^2_m & 0 \\ 0 & \sigma^2_n \end{bmatrix}$.

2.3 State-space representation and likelihood function

We solve the semi-structural model (3)-(8) with Gensys (Sims, 2002), to obtain:

$$Z_t = \mathcal{C}(\theta) + G_1(\theta) Z_{t-1} + B(\theta) \varepsilon_t. \tag{9}$$

[16] For a detailed discussion of sources of real rigidities see Woodford (2003, chapter 3).

[17] The model features the same degree of real rigidity in all sectors. This is the case in essentially all of the literature on multi-sector sticky-price models (Nakajima et al. (2010) is, to our knowledge, the only paper to analyze a New Keynesian model with cross-sectional heterogeneity in real rigidities). This homogeneity follows from the fact that sectors are assumed to differ only in the degree of price stickiness. Other sources of sectoral heterogeneity might imply differences in the extent of real rigidities across sectors, with potentially interesting implications for aggregate dynamics. Inasmuch as our analysis is concerned, heterogeneity along this dimension might generate the need for additional observables – such as sectoral price data – in order to identify the underlying cross-sectional distribution. We leave this potentially interesting endeavour for future research.

where Z_t is a vector collecting all variables and additional "dummy" variables created to account for leads and lags and ε_t is as defined before. The vector θ collects the primitive parameters of the model:

$$\theta = \left(K, p_1, p_2, \beta, \zeta, \sigma_m, \sigma_n, \omega_1, \cdots, \omega_K, \rho_0, \cdots, \rho_{p_1}, \delta_0, \cdots, \delta_{p_2}\right).$$

In all estimations that follow we make use of the likelihood function $\mathcal{L}(\theta|Z^*)$, where Z^* is the vector of observed time series (i.e., a subset of Z). Given that our state vector Z_t includes many unobserved variables, such as the natural output level and sectoral prices, the likelihood function is constructed through application of the Kalman filter to the solved loglinear model (9). Letting H denote the matrix that singles out the observed subspace Z_t^* of the state vector Z_t (i.e., $Z_t^* = HZ_t$), our distributional assumptions can be summarized as:

$$\begin{aligned} Z_t|Z_{t-1} &\sim N\left(C(\theta) + G_1(\theta) Z_{t-1}, B(\theta)\Omega B(\theta)'\right), \\ Z_t^*|\left\{Z_\tau^*\right\}_{\tau=1}^{t-1} &\sim N\left(\mathcal{M}_{t|t-1}(\theta), V_{t|t-1}(\theta)\right), \end{aligned}$$

where $\mathcal{M}_{t|t-1}(\theta) \equiv HC(\theta) + HG_1(\theta)\hat{Z}_{t|t-1}$, $V_{t|t-1}(\theta) \equiv HB(\theta)\hat{\Sigma}_{t|t-1}B(\theta)'H'$, $\hat{Z}_{t|t-1}$ denotes the expected value of Z_t given $\left\{Z_\tau^*\right\}_{\tau=1}^{t-1}$, and $\hat{\Sigma}_{t|t-1}$ is the associated forecast-error covariance matrix.

2.4 Identification of the cross-sectional distribution from aggregate data

In estimating our multi-sector model we use only time-series data on aggregate nominal and real output as observables. It is thus natural to ask whether the structure of the model is such that these aggregate data reveal information about the cross-sectional distribution of price stickiness $\omega = (\omega_1, ..., \omega_K)$. As in Jadresic (1999), we start by looking at a simple case where it is easy to show that ω can be inferred from observations of those two aggregate time series. This helps develop the intuition for a more general case for which we also show identification. We then assess the small-sample properties of estimates of ω inferred from aggregate data through a Monte Carlo exercise. As in our estimation, we assume throughout that the discount factor, β, is known.

The key simplifying assumption to show identification in the first case is absence of pricing interactions: $\zeta = 1$. In that case, from (6) new prices $x_t(k)$ are set on the basis of current and expected future values of the two exogenous processes m_t and y_t^n. For simplicity and without loss of generality, assume

further that the latter follow the AR(1) processes:

$$m_t = \rho_1 m_{t-1} + \varepsilon_t^m, \text{ and} \tag{10}$$

$$y_t^n = \delta_1 y_{t-1}^n + \varepsilon_t^n. \tag{11}$$

Then, new prices are set according to:

$$x_t(k) = F(\beta, \rho_1, k)\, m_t - F(\beta, \delta_1, k)\, y_t^n,$$

where

$$F(\beta, a, k) \equiv \left(1 + \frac{1-\beta}{1-\beta^k}\frac{\beta a - (\beta a)^k}{1-\beta a}\right).$$

Replacing this expression for newly set prices in the sectoral price equation (5) and aggregating according to (4) produces the following expression for the aggregate price level:

$$p_t = \sum_{j=0}^{K-1}\sum_{k=j+1}^{K} F(\beta, \rho_1, k)\frac{\omega_k}{k} m_{t-j} - \sum_{j=0}^{K-1}\sum_{k=j+1}^{K} F(\beta, \delta_1, k)\frac{\omega_k}{k} y_{t-j}^n. \tag{12}$$

If we observe m_t and y_t - and thus p_t, estimates of the coefficients on m_{t-j} in (12) allow us to infer the sectoral weights ω. The reason is that $F(\beta, \rho_1, k)$ is "known", since ρ_1 can be estimated directly from (10). Thus, knowledge of the coefficient on the longest lag of m_{t-j} (attained when $j = K-1$) allows us to uncover ω_K. The coefficient on the next longest lag ($m_{t-(K-2)}$) depends on ω_{K-1} and ω_K, which reveals ω_{K-1}. We can thus recursively infer the sectoral weights from the coefficients $F(\beta, \rho_1, k)\frac{\omega_k}{k}$. Moreover, identification obtains with any estimation method that produces consistent estimates of these coefficients.[18]

Checking for identification of ω in the presence of pricing interactions ($\zeta \neq 1$) is slightly more involved. To gain intuition on why this is so, fix the case of pricing complementarities ($\zeta < 1$). Then, because of the delayed response of sticky-price firms to shocks, firms with flexible prices will only react partially to innovations to m_t and y_t^n on impact. They will eventually react fully to the shocks, but also with a delay.

It turns out that the "recursive identification" that applies when $\zeta = 1$ also works in this case. The reason is that, in equilibrium, pricing interactions manifest themselves through a dependence of the aggregate price level on its own lags. This is how they serve as a propagation mechanism. Specifically,

[18] Jadresic (1999) discusses identification in a similar context. The main differences are that he considers a regression based on a first-differenced version of the analogous equation in his model, and assumes $\rho_1 = 1$ and that the term corresponding to $\sum_{j=0}^{K-1}\sum_{k=j+1}^{K} F(\beta, \delta_1, k)\frac{\omega_k}{k}\Delta y_{t-j}^n$ is an i.i.d. disturbance.

the expression for the equilibrium price level becomes:

$$p_t = \sum_{j=1}^{K-1} a_j p_{t-j} + \sum_{j=0}^{K-1} b_j m_{t-j} - \sum_{j=0}^{K-1} b_j y^n_{t-j}, \tag{13}$$

where $a_1, ..., a_{K-1}$, b_0, ..., b_{K-1} are functions of the model parameters. Knowledge of the coefficients on the lags of the aggregate price level and on lagged nominal output again allows us to solve for the sectoral weights – and for ζ.[19]

The intuition behind the identification result in the absence of pricing interactions is clear: the impact of older developments of the exogenous processes on the current price level depends on prices that are sticky enough to have been set when the shocks hit. This provides information on the share of the sector with that given duration of price spells (and sectors with longer durations). More generally, in the presence of pricing interactions, fully forward-looking pricing decisions may also reflect past developments of the exogenous processes. This dependence manifests itself through lags of the aggregate price level. The intuition behind the mechanism that allows for identification extends in a natural way: sectors where prices are more sticky are relatively more important in determining the impact of older shocks to the exogenous processes on the current price level, and vice-versa for sectors where prices are more flexible. Moreover, the relative sizes of the coefficients on past prices and past nominal output in (13) pin down the index of real rigidities ζ.

These results on identification are of little practical use to us if the mechanism highlighted above does not work well in finite samples. To analyze this issue we rely on a Monte Carlo exercise. We generate artificial data on aggregate nominal and real output using parameter values that roughly resemble what we find when we estimate the model with actual data. Then, we estimate the parameters of the model by maximum likelihood. We conduct both a large- and a small-sample exercise. Details and results are reported in the Appendix.

The bottom line is that, for large samples, the estimates are quite close to the true parameter values, and fall within a relatively narrow range. For samples of the same size as our actual sample, we also find the aggregate data to be informative of the distribution of sectoral weights. However, in this case the estimates are less precise and somewhat biased. This finding underscores our case for incorporating prior information from the microeconomic evidence on price-setting, as we do in Section 4.3.

[19] In the Appendix we illustrate how the process works in a two-sector model.

3 Empirical methodology and data

With the challenges involved in bridging the gap between price-setting statistics based on micro data and time series of aggregate nominal and real output, the choice of empirical methodology is critical. We employ a Bayesian approach, as this allows us to integrate those two sources of information.

With some abuse of notation, the Bayesian principle can be shortly stated as:

$$f\left(\theta|Z^*\right) = f\left(Z^*|\theta\right) f\left(\theta\right) / f\left(Z^*\right) \propto \mathcal{L}\left(\theta|Z^*\right) f\left(\theta\right),$$

where f denotes density functions, Z^* is the vector of observed time series, θ is the vector of primitive parameters, and $\mathcal{L}\left(\theta|Z^*\right)$ is the likelihood function.

As observables, we use time series of aggregate nominal and real output. For constructing our prior distribution over the vector of sectoral weights, $f\left(\omega_1, ..., \omega_K\right)$, we derive empirical distributions from Bils and Klenow (2004) and Nakamura and Steinsson (2008), as discussed in detail in Subsection 3.1 below. In the next subsections we detail our prior distributions, sources of data, and estimation approach.

3.1 Prior over ω

We specify priors over the set of sectoral weights $\omega = (\omega_1, ..., \omega_K)$, which are then combined with the priors on the remaining parameters to produce the joint prior distribution for the set of all parameters of interest. We impose the combined restrictions of non-negativity and summation to unity of the ω's through a Dirichlet distribution, which is a multivariate generalization of the beta distribution. Notationally, $\omega \sim D\left(\alpha_1, ..., \alpha_K\right)$ with density function:

$$f_\omega\left(\omega|\alpha_1, ..., \alpha_K\right) \propto \prod_{k=1}^{K} \omega_k^{\alpha_k - 1},\ \forall \alpha_k > 0,\ \forall \omega_k \geq 0,\ \sum_{k=1}^{K} \omega_k = 1.$$

The Dirichlet distribution is well known in Bayesian econometrics as the conjugate prior for the multinomial distribution, and the hyperparameters $\alpha_1, ..., \alpha_K$ are most easily understood in this context, where they can be interpreted as the "number of occurrences" for each of the K possible outcomes that the econometrician assigns to the prior information.[20] Thus, for given $\alpha_1, ..., \alpha_K$, the parameter $\alpha_0 \equiv \sum_k \alpha_k$ captures, in some sense, the overall level of information in the prior distribution. The information about the cross-sectional distribution of price stickiness comes from the relative sizes of the α_k's. The latter also determine the marginal distributions for the ω_k's. For example, the expected value of ω_k is simply

[20] Gelman et al. (2003) offers a good introduction to the use of Dirichlet distribution as a prior distribution for the multinomial model.

α_k/α_0, whereas its mode equals $(\alpha_0 - K)^{-1}(\alpha_k - 1)$ (provided that $\alpha_i > 1$ for all i).

Whenever we want to estimate a cross-sectional distribution of price rigidity based solely on aggregate data, we can impose an uninformative ("flat") prior, in which all ω vectors in the K-dimensional unit simplex are assigned equal prior density. This corresponds to $\alpha_k = 1$ for all k – and thus $\alpha_0 = K$. This allows us to extract the information that the aggregate data contain about the cross-sectional distribution of price stickiness.

To incorporate microeconomic information in the estimation, we relate the relative sizes of the hyperparameters $(\alpha_1, ..., \alpha_K)$ to the empirical sectoral weights derived from the micro data, and choose the value $\alpha_0 > K$ to determine the tightness of the prior distribution around the empirical distribution. Specifically, let $\widehat{\omega}$ denote the set of sectoral weights from a given empirical distribution. We specify the relative sizes of the hyperparameters $(\alpha_1, ..., \alpha_K)$ so that the mode of the prior distribution for ω coincides with the empirical sectoral weights $\widehat{\omega}$. This requires setting $\alpha_k = 1 + \widehat{\omega}_k(\alpha_0 - K)$. The case of flat priors analyzed previously obtains when $\alpha_0 = K$. Henceforth, we refer to α_0/K as the degree of "prior informativeness".

3.2 Priors on remaining parameters

The remaining parameters of the model fall into three categories that we deal with in turn. Our goal in specifying their prior distributions is to avoid imposing any meaningful penalties on most parameter values – except for those that really seem extreme on an *a priori* basis. The first set comprises the ρ's and δ's, parameterizing the exogenous AR processes for nominal and natural output, respectively. These are assigned loose Gaussian priors with mean zero. We choose to fix the lag length at two for both processes, i.e. $p_1 = p_2 = 2$.[21] The second set of parameters consists of the standard deviations of the shocks to nominal (σ_m) and natural output (σ_n). These are strictly positive parameters to which we assign loose Gamma priors. The last parameter is the Ball-Romer index of real rigidity, ζ, which should also be non-negative. This is captured with a very loose Gamma prior distribution, with mode at unity and a 5-95 percentile interval equal to $(0.47, 16.9)$. Hence, any meaningful degree of pricing complementarity or substitutability should be a result of the estimation rather than of our prior assumptions. These priors are summarized in Table 1.[22]

[21] In principle we could specify priors over p_1, p_2 and estimate their posterior distributions as well. However, the computational cost of estimating all the models in the paper is already quite high, and we restrict ourselves to this specification with fixed number of lags. Our conclusions are robust to alternative assumptions about the number of lags (see Section 5).

[22] We do not include β in the estimation, and set $\beta = 0.99$.

Table 1: Prior distributions for remaining parameters

Parameter	Distribution	Mode	Mean	Std.dev.
ζ	$Gamma\,(1.2, 0.2)$	1.00	6.00	5.48
ρ_j, δ_j	$N\left(0, 5^2\right)$	0.00	0.00	5.00
σ_n, σ_m	$Gamma\,(1.5, 20)$	0.025	0.075	0.06

Note: The hyper-parameters for the Gamma distribution specify shape and inverse scale, respectively, as in Gelman et al. (2003).

3.3 Macroeconomic time series

We estimate the model using quarterly data on nominal and real output for the U.S. economy. These are measured as seasonally-adjusted GDP at, respectively, current and constant prices, from the Bureau of Economic Analysis. We take natural logarithms and remove a linear trend from the data. Whereas the assumptions underlying the model include one of an unchanged economic environment, the U.S. economy has undergone profound changes in the recent decades, including the so-called "Great Moderation" and the Volcker Disinflation. As a consequence, we choose not to confront the model with the full sample of post-war data. We use the period from 1979 to 1982 as a pre-sample, and evaluate the model according to its ability to match business cycle developments in nominal and real output in the period 1983-2007.[23]

3.4 Empirical distributions of price stickiness

We work with the statistics on the frequency of price changes for the 350 categories of goods and services ("entry level items") reported by Bils and Klenow (2004, henceforth BK), and with the 272 entry level items covered by Nakamura and Steinsson (2008, henceforth NS). In the latter case we use the statistics for regular prices (those excluding sales and product substitutions). We refer to the corresponding empirical distributions of price rigidity as distributions *with* (BK) and *without* (NS) *sales*.

Our goal is to map those statistics into an empirical distribution of sectoral weights over spells of price rigidity with different durations. We work at a quarterly frequency, and for computational reasons consider economies with at most 8 quarters of price stickiness. Sectors correspond to price spells which are multiples of one quarter. We denote an empirical cross-sectional distribution of price rigidity by $\{\widehat{\omega}_k\}_{k=1}^{8}$, where $\widehat{\omega}_1$ denotes the fraction of firms that change prices every quarter, $\widehat{\omega}_2$ the fraction with an expected duration of price spells between one (exclusive) and two quarters (inclusive), and so on. The sectoral weights are aggregated accordingly by adding up the corresponding CPI expenditure weights. We proceed in this fashion until the sector with 7-quarter price spells. Finally, we aggregate all the

[23] We make use of the pre-sample 1979-1982 by initializing the Kalman filter in the estimation stage with the estimate of Z_t and corresponding covariance matrix obtained from running a Kalman filter in the pre-sample. We use the parameter values in each draw. For the initial condition for the pre-sample, we use the unconditional mean and a large variance-covariance matrix.

remaining categories, which have mean durations of price rigidity of 8 quarters and beyond, into a sector with 2-year price spells. This gives rise to the empirical cross-sectional distributions of price stickiness that we use in our estimation, which are summarized in Table 2. We denote the sectoral weight for sector k obtained from this procedure by $\widehat{\omega}_k$. For each of the BK and NS distributions, we also compute the average duration of price spells, $\widehat{\overline{k}} = \sum_{k=1}^{8} \widehat{\omega}_k k$, and the cross-sectional standard deviation of the underlying distribution, $\widehat{\sigma}_k = \sqrt{\sum_{k=1}^{8} \widehat{\omega}_k \left(k - \widehat{\overline{k}}\right)^2}$.

Table 2: Empirical cross-sectional distributions of price stickiness

Parameter	With sales (BK)	Without sales (NS)
$\widehat{\omega}_1$	0.395	0.273
$\widehat{\omega}_2$	0.240	0.071
$\widehat{\omega}_3$	0.116	0.098
$\widehat{\omega}_4$	0.118	0.110
$\widehat{\omega}_5$	0.037	0.060
$\widehat{\omega}_6$	0.033	0.129
$\widehat{\omega}_7$	0.030	0.061
$\widehat{\omega}_8$	0.032	0.198
$\widehat{\overline{k}}^{(*)}$	2.54	4.23
$\widehat{\sigma}_k^{(*)}$	1.86	2.66

(*) In quarters. $\sum \widehat{\omega}_k$ might differ from unity due to rounding.

3.5 Simulating the posterior distribution

The joint posterior distribution of the model parameters is obtained through application of a Markov-chain Monte Carlo (MCMC) Metropolis algorithm. The algorithm produces a simulation sample of the parameters that converges to the joint posterior distribution under certain conditions.[24] We provide details of our specific estimation process in the Appendix. The outcome is a sample of one million draws from the joint posterior distribution of the parameters of interest, based on which we draw the conclusions that we start to report in the next section.

Having obtained a sample of the posterior distribution of parameters from any given model, we can estimate the marginal posterior density (henceforth MPD) of the data given the model as:

$$\text{MPD}_j = f\left(Z^* | \mathcal{M}_j\right) = \int \mathcal{L}\left(\theta | Z^*, \mathcal{M}_j\right) f\left(\theta | \mathcal{M}_j\right) d\theta, \tag{14}$$

and use it for model-comparison purposes. In (14), $\mathcal{M}_j$ refers to a specific configuration of the model and prior distribution, and $f\left(\theta | \mathcal{M}_j\right)$ denotes the corresponding joint prior distribution. Specifically,

[24] These conditions are discussed in Gelman et al. (2003, part III).

we approximate $\log(\text{MPD}_j)$ for each model using Geweke's (1999) modified harmonic mean. We use these estimates to evaluate the empirical fit of the models relative to one another. The MPD ratio of two model configurations constitutes the *Bayes factor*, and – when neither configuration is *a priori* considered more likely – the *posterior odds*. It indicates how likely the two models are relative to one another given the observed data Z^*.

4 Results

4.1 Macro-based estimates

Table 3 and Figure 1 report the results for the case of uninformative priors, in terms of marginal distributions for the parameters.[25] The empirical distributions of price rigidity from Table 2 are reproduced in the last columns, for ease of comparison. In what follows, we use the posterior means as the point estimates for the sectoral weights, reported in the third column of the table.[26]

The cross-sectional distributions that we infer from aggregate data conform quite well with the empirical ones. The macro-based estimates imply that approximately 28% of firms change prices every quarter; 43% change prices at least once a year; 13% change prices once every two years. The average duration of price spells is 13 months, and the standard deviation of the duration of price spells is approximately 8 months. These numbers are quite close to the empirical distribution without sales and product substitutions (last column of the table). The correlation between our macro-based estimates and those empirical weights is 0.63. The correlation of the estimates with the empirical distribution with sales and product substitutions is somewhat lower, at 0.43. This is a first, informal indication that the distribution that excludes sales and product substitutions helps the model fit aggregate dynamics better. Below we investigate this possibility by performing formal model comparisons using a standard measure of fit.

The index of real rigidities implies strong pricing complementarities. The posterior mean of ζ is 0.05 and the 95^{th} percentile equals 0.11, which falls within the 0.10-0.15 range that Woodford (2003) argues can be made consistent with fully specified models. As highlighted by Carvalho (2006), such complementarities interact with heterogeneity in price stickiness to amplify the aggregate effects of nominal rigidities in this type of sticky-price model.

[25] We use a Gaussian kernel density estimator to graph the marginal posterior density for each parameter. The priors on $\bar{k}$ and σ_k are based on 100,000 draws from the prior Dirichlet distribution.

[26] The results are almost insensitive to using alternative point estimates, such as the values at the joint posterior mode, or taking medians or modes from the marginal ditributions and renormalizing so that the weights sum to unity.

Table 3: Parameter estimates under a flat prior

	$\alpha_k = 1$ *for all* k ($\alpha_0 = 8$)			*Empirical distributions*	
				With sales	*W/o sales*
ζ	4.440 (0.466;16.863)	0.042 (0.015;0.111)	0.050	–	–
ω_1	0.094 (0.007;0.348)	0.264 (0.099;0.493)	0.276	0.395	0.273
ω_2	0.094 (0.007;0.348)	0.072 (0.007;0.212)	0.086	0.240	0.071
ω_3	0.094 (0.007;0.348)	0.020 (0.002;0.078)	0.027	0.116	0.098
ω_4	0.094 (0.007;0.348)	0.027 (0.002;0.107)	0.037	0.118	0.110
ω_5	0.094 (0.007;0.348)	0.144 (0.017;0.337)	0.156	0.037	0.059
ω_6	0.094 (0.007;0.348)	0.123 (0.011;0.345)	0.144	0.033	0.129
ω_7	0.094 (0.007;0.348)	0.120 (0.010;0.353)	0.143	0.030	0.061
ω_8	0.094 (0.007;0.348)	0.112 (0.010;0.323)	0.132	0.032	0.198
$\bar{k}$	4.501 (3.245;5.760)	4.394 (3.214;5.462)	4.37	2.54	4.25
σ_k	2.139 (1.584;2.678)	2.523 (2.112;2.893)	2.62	1.86	2.66
ρ_0	0.000 (−8.224;8.224)	0.000 (−0.001;0.001)	0.000	–	–
ρ_1	0.000 (−8.224;8.224)	1.426 (1.273;1.576)	1.426	–	–
ρ_2	0.000 (−8.224;8.224)	−0.446 (−0.593;−0.296)	−0.446	–	–
σ_m	0.059 (0.009;0.195)	0.005 (0.005;0.006)	0.005	–	–
δ_0	0.000 (−8.224;8.224)	0.002 (−0.002;0.007)	0.003	–	–
δ_1	0.000 (−8.224;8.224)	0.541 (0.270;0.763)	0.532	–	–
δ_2	0.000 (−8.224;8.224)	0.146 (−0.027;0.331)	0.149	–	–
σ_n	0.059 (0.009;0.195)	0.069 (0.030;0.172)	0.081	–	–

Note: The first two columns report the medians of, respectively, the marginal prior and posterior distributions; the third column gives the mean of the marginal posterior distribution; numbers in parentheses correspond to the 5^{th} and 95^{th} percentiles; the last columns reproduce the empirical distributions from Table 2.

4.2 Comparison with homogeneous-firms models

In this subsection we ask how sharply the data allow us to discriminate between multi-sector models with heterogeneity in price stickiness and one-sector models with homogeneous firms. To that end, we estimate one-sector models with price spells ranging from two to eight quarters. We keep the same prior distributions for all parameters besides the sectoral weights. A one-sector model with price spells of length k, say, can be seen as a restriction of the multi-sector model, with a degenerate distribution of sectoral weights ($\omega_k = 1$, $\omega_{k'} = 0$ for all $k' \neq k$).

We pick the best-fitting one-sector model according to the marginal density of the data given the models. The results are reported in Table 4 and Figure 2. The best-fitting model is the one in which all price spells last for 7 quarters. This seems unreasonable in light of the microeconomic evidence. Given the extent of nominal rigidity, not surprisingly the degree of pricing complementarity is smaller. The posterior distributions for the parameters of the nominal output process are quite similar to the ones obtained in the multi-sector models. Perhaps this should be expected given that this variable is one of the observables used in the estimation. In contrast, the distributions of the parameters of the unobserved driving process are different under the restriction of homogeneous firms. We defer a discussion of what might drive this result to the end of this subsection.

Table 4: Best-fitting homogeneous economy

	Prior	$K = 7$, $\omega_7 = 1$	
ζ	4.440 (0.466;16.863)	0.362 (0.193;0.830)	0.419
ρ_0	0.000 (−8.224;8.224)	0.000 (−0.001;0.001)	0.000
ρ_1	0.000 (−8.224;8.224)	1.430 (1.284;1.568)	1.428
ρ_2	0.000 (−8.224;8.224)	−0.454 (−0.590;−0.310)	−0.452
σ_m	0.059 (0.009;0.195)	0.005 (0.005;0.006)	0.005
δ_0	0.000 (−8.224;8.224)	0.003 (−0.003;0.011)	0.004
δ_1	0.000 (−8.224;8.224)	0.064 (−0.154;0.319)	0.071
δ_2	0.000 (−8.224;8.224)	0.135 (−0.027;0.327)	0.141
σ_n	0.059 (0.009;0.195)	0.216 (0.087;0.421)	0.230

Note: The first two columns report the medians of, respectively, the marginal prior and posterior distributions; the third column gives the mean of the marginal posterior distribution; numbers in parentheses correspond to the 5^{th} and 95^{th} percentiles.

The multi-sector model with $K = 8$ nests the best-fitting homogeneous-firms model. Thus, under measures of fit that do not "correct" for the number of parameters, the former model will necessarily perform at least as well as the latter model. To circumvent that problem we base our comparison on the marginal posterior density of the data given the models, which already accounts for the fact that

the multi-sector model has more parameters than the homogeneous-firms model.[27]

Table 5 reports the results for the multi-sector model with the flat prior for ω, and the best-fitting one-sector model. The fit of the multi-sector model is much better than that of the best-fitting one-sector model: the posterior odds in favor of the former model is of the order of $10^{11} : 1$.

Table 5: Model comparison - heterogeneous versus homogeneous economy

	Multi-sector model	Best-fitting 1-sector model
log MPD	808.03	781.33

Note: The logarithm of the marginal posterior density of the data given the models (log MPD) is approximated with Geweke's (1999) modified harmonic mean.

Our model-comparison criterion has the disadvantage that it does not provide any information on what drives the improved empirical fit of the multi-sector model. To shed some light on this question we compare model-implied dynamics for inflation and output to those of a restricted bivariate VAR including nominal and real output. In estimating the VAR we impose the same assumption used in the models, that nominal output is exogenous and follows an AR(2) process. We allow real output to depend on four lags of both itself and nominal output, and to be contemporaneously affected by innovations to nominal output. Estimation is by ordinary least squares. The multi-sector model is the one estimated under flat priors for ω, while the one-sector model is the one with the best fit. The parameter values are fixed at their posterior means. Since the impulse response functions are conditional on specific parameter values, the comparison does not correct for the larger number of parameters in the multi-sector model. Thus, it is only meant to provide some indication of the sources of the large differences in the posterior odds of the models.

The panel in Figure 3 shows the impulse response functions of output (y_t, left column) and inflation (π_t, right column) to positive innovations ε_t^m (top row) and ε_t^n (bottom row) of one standard deviation in size.[28] Relative to the one-sector model, the estimated multi-sector model does a better job at approximating the impulse response functions produced by the VAR at both short and medium horizons, in response to both shocks. Thus the overwhelming statistical support for heterogeneity does not seem to depend on any single feature of the dynamic response of macroeconomic variables to the shocks. Finally, these results suggest one explanation for why the estimated parameters associated with the unobserved driving process are different in the one-sector economy. While the multi-sector model can rely on the distribution of sectoral weights to balance the response of the economy to shocks at different

[27] The reason is that the vector of parameters is "integrated out" in (14).

[28] Following the notation of the semi-structural model, in the VAR ε_t^m denotes innovations to nominal output, and ε_t^n denotes the other (orthogonal) innovations.

horizons, the one-sector model lacks this mechanism. Given the facts that nominal output is observed and that its parameter estimates imply quite persistent dynamics in both economies, perhaps the one-sector economy needs to rely on the unobserved process as a more transient and volatile component that can help it do a better job at matching higher-frequency features of the data.

4.3 Combining micro and macro data in the estimation

We now turn to estimations that incorporate information from price-setting statistics derived from micro data. Table 6 and Figures 4-7 present the results for two sets of informative priors ($\alpha_0/K = 2$, 5) for each empirical distribution. The bottom row of Table 6 reports the log of the posterior marginal density of the various models. For the less informative set of priors ($\alpha_0/K = 2$) the two empirical distributions that inform the prior lead to models that perform similarly in terms of fit – and close to the model estimated under a flat prior. However, for the more informative set of priors ($\alpha_0/K = 5$), the model with prior based on the empirical distribution of price rigidity without sales fits the data better according to the posterior marginal density criterion – the difference of 4.4 log-points implies a posterior odds ratio of roughly 80 : 1 in favor of the model with prior distribution that excludes sales and product substitutions.

We can use such a comparison of posterior marginal densities of various estimated models for assessing the relative merits of the two sets of priors for the purpose of helping the model explain aggregate dynamics. To that end, we estimate a series of additional models with informative priors based on the two empirical distributions of price rigidity (with and without sales), progressively increasing the degree of prior informativeness (i.e., increasing α_0/K). Specifically, we estimate additional models with $\alpha_0/K = 10, 20, 100$, and 1000. In addition, we estimate models in which the distribution of price stickiness that forms the prior has equal weights in all sectors ("uniform prior"). We summarize the results in Figure 8. It shows clearly that the difference between the fit of estimated models increases as the priors become more informative. While the difference in fit between the models based on the prior distribution without sales and the uniform prior is not that large (it tends to approximately 3 log-points for very informative priors), the difference between models based on prior distributions with and without sales is more substantial. As the degree of prior informativeness increases, that difference approaches 6 log-points – which translates into a posterior odds ratio of roughly 400 : 1 in favor of the model with prior distribution that excludes sales and product substitutions.[29]

[29] Figure 8 also shows that, as we tighten the priors on the sectoral weights, the fit of models estimated under priors with sales (BK) and priors with a uniform distribution deteriorates. In turn, the fit of models estimated under priors that exclude sales remains essentially unchanged. This is consistent with our previous finding that sectoral weights estimated under flat priors are somewhat similar to the empirical distribution without sales.

Table 6: Parameter estimates with informative priors

	Inform. prior, $\alpha_0/K = 2$		*Inform. prior*, $\alpha_0/K = 5$		*Flat prior*	*Empirical distributions*	
	With sales	*W/o sales*	*With sales*	*W/o sales*		*With sales*	*W/o sales*
ζ	0.032 (0.01;0.08)	0.042 (0.02;0.11)	0.018 (0.01;0.05)	0.041 (0.02;0.10)	0.042 (0.02;0.11)	–	–
ω_1	0.324 (0.17;0.51)	0.277 (0.14;0.45)	0.425 (0.31;0.55)	0.309 (0.21;0.43)	0.264 (0.099;0.49)	0.395	0.273
ω_2	0.123 (0.04;0.24)	0.069 (0.01;0.18)	0.190 (0.11;0.29)	0.059 (0.02;0.13)	0.072 (0.01;0.212)	0.240	0.071
ω_3	0.035 (0.01;0.09)	0.033 (0.01;0.09)	0.059 (0.02;0.11)	0.051 (0.02;0.10)	0.020 (0.00;0.08)	0.116	0.098
ω_4	0.049 (0.01;0.13)	0.047 (0.01;0.12)	0.081 (0.03;0.15)	0.072 (0.03;0.14)	0.027 (0.00;0.11)	0.118	0.110
ω_5	0.106 (0.02;0.26)	0.109 (0.02;0.26)	0.056 (0.01;0.15)	0.066 (0.02;0.15)	0.144 (0.01;0.34)	0.037	0.059
ω_6	0.100 (0.01;0.27)	0.142 (0.04;0.31)	0.052 (0.01;0.15)	0.144 (0.07;0.25)	0.123 (0.01;0.35)	0.033	0.129
ω_7	0.090 (0.01;0.25)	0.086 (0.01;0.24)	0.042 (0.00;0.13)	0.058 (0.02;0.14)	0.120 (0.01;0.35)	0.030	0.061
ω_8	0.088 (0.01;0.24)	0.160 (0.05;0.32)	0.044 (0.01;0.13)	0.200 (0.11;0.31)	0.112 (0.01;0.32)	0.032	0.198
$\bar{k}$	3.776 (2.91;4.69)	4.367 (3.45;5.25)	2.811 (2.31;3.40)	4.262 (3.60;4.91)	4.394 (3.21;5.46)	2.54	4.25
σ_k	2.515 (2.17;2.85)	2.612 (2.28;2.91)	2.184 (1.81;2.56)	2.725 (2.50;2.93)	2.523 (2.11;2.89)	1.86	2.66
ρ_0	0.000 (−0.00;0.00)	0.000 (−0.00;0.00)	0.000 (−0.00;0.00)	0.000 (−0.00;0.00)	0.000 (−0.00;0.00)	–	–
ρ_1	1.425 (1.27;1.57)	1.427 (1.27;1.58)	1.424 (1.27;1.57)	1.429 (1.28;1.58)	1.426 (1.27;1.58)	–	–
ρ_2	−0.445 (−0.59;−0.30)	−0.447 (−0.59;−0.30)	−0.444 (−0.59;−0.29)	−0.449 (−0.60;−0.30)	−0.446 (−0.59;−0.30)	–	–
σ_m	0.005 (0.00;0.01)	0.005 (0.00;0.01)	0.005 (0.00;0.01)	0.005 (0.00;0.01)	0.005 (0.00;0.01)	–	–
δ_0	0.002 (−0.00;0.01)	0.002 (−0.00;0.01)	0.002 (−0.00;0.01)	0.002 (−0.00;0.01)	0.002 (−0.00;0.01)	–	–
δ_1	0.514 (0.30;0.72)	0.545 (0.32;0.75)	0.465 (0.28;0.65)	0.563 (0.38;0.75)	0.541 (0.27;0.76)	–	–
δ_2	0.176 (0.01;0.34)	0.151 (−0.01;0.32)	0.201 (0.06;0.34)	0.146 (−0.01;0.30)	0.146 (−0.03;0.33)	–	–
σ_n	0.068 (0.03;0.16)	0.066 (0.03;0.16)	0.072 (0.03;0.17)	0.062 (0.03;0.15)	0.069 (0.03;0.17)	–	–
log MPD	807.56	808.27	803.768	808.16	808.03		

Note: The first four columns report the posterior medians under informative priors, and the fifth column reproduces the posterior medians under a flat prior from Table 3; numbers in parentheses correspond to the 5^{th} and 95^{th} percentiles; the last two columns reproduce the empirical distributions from Table 2.

5 Robustness

Our findings are robust to different prior assumptions for the parameters ρ_i, δ_i, σ_m, σ_n and ζ, as well as different de-trending procedures and specifications for the exogenous time-series processes. In particular, they are robust to using a Baxter and King (1999) filter or first-differences instead of removing linear trends from the data, and to assuming AR(3) exogenous processes (i.e., $p_1 = p_2 = 3$). Also, unreported results with models with $K < 8$ suggest that one needs to allow for "enough" heterogeneity in order to avoid compromising the empirical performance of the model. In particular, the fit of models with $K = 4$ (as in Coenen et al. 2007) is much worse than models with $K = 6$ or 8. While the differences in empirical performance between models with $K = 6$ and $K = 8$ are not that large, the evidence against the specifications with $K = 4$ is quite strong: posterior odds ratios favor the models with $K = 6, 8$ by an order of $10^5 : 1$.

In the sections below we discuss the robustness of our findings to alternative models of price setting. In particular, we consider the Calvo (1983) model, and discuss a new model of price setting that produces the exact same results as our model, and yet can speak to a much larger set of empirical facts about price setting derived from micro data.

5.1 Results under the Calvo (1983) model

We also considered versions of the model with Calvo (1983) pricing. Mimicking our baseline analysis, the first step is to show that the model allows for identification of the cross-sectional distribution of price rigidity from aggregate data, and, given that result, that it also allows for separate identification of nominal and real rigidities. Indeed, all identification results go through, and the intuition is very similar to the one in the Taylor model. In the Appendix we provide a thorough proof of identification, including the case with strategic interactions in price-setting decisions (i.e., index of real rigidities $\zeta \neq 1$).

However, under Calvo pricing, not all of our conclusions are equally robust when it comes to relatively small samples. The reason is that, in the context of our semi-structural framework, identification of heterogeneity in price stickiness under Calvo pricing is "more difficult" than under Taylor pricing. Building on Monte Carlo analysis and analytical insights from simple versions of these two pricing models, we found that clear-cut identification of the distribution of price stickiness depends on whether the observable driving process has high variance relative to the unobservable process.

While this "restriction" applies to both price-setting specifications, the identification problem is more acute under Calvo pricing. Based on Monte Carlo analysis, we found that with our sample size and the relative variances for the two exogenous processes implied by our point estimates, the likelihood

criterion fails to provide a sharp discrimination between alternative (non-degenerate) distributions of price stickiness under Calvo pricing. This mirrors what we find in the data: under Calvo pricing, they do not allow too sharp a discrimination between different models with heterogeneity in price stickiness. In contrast, given the same sample size and relative variances for those two processes, the version of the model with Taylor pricing provides more information about the underlying distribution of price stickiness – as seen in previous sections.

However, despite that difficulty, our main findings *do* hold under the Calvo pricing model – at least qualitatively. First, on the comparison between models with heterogeneity in price stickiness and models with homogeneous firms, the estimated models provide clear evidence in favor of the former. Specifically, we find that a likelihood-ratio test of the homogeneous Calvo model against multi-sector versions of the model leads to rejection of the former at significance levels of less than 1%.[30] Second, all estimated models feature $\zeta < 1$, implying strategic complementarities in price setting. Finally, estimations under informative priors derived from the empirical distributions of price stickiness (as described in Section 3) also provide (qualitative) evidence in favor of the distribution that excludes sales and product substitutions.[31]

5.2 An alternative model

As we mentioned in Section 2, the Taylor model is, strictly speaking, at odds with the microeconomic evidence on the duration of price spells (e.g., Klenow and Kryvtsov 2008). This inconsistency may be viewed as a weakness of the Taylor model relative to alternatives – in particular the Calvo model, which naturally produces a non-degenerate distribution of the duration of price spells at the firm level.

However, this evidence does not invalidate the use of that model for our purposes. To show that this is the case, here we provide an alternative model in which the duration of price spells varies at the firm level. The model can match the empirical distribution of the duration of price spells. Yet, the aggregate behavior of the model is identical to the one presented in Section 2. Furthermore, this alternative model can match additional micro facts documented in the literature – in a similar fashion as the Calvo (1983) model.

There is a continuum of monopolistically competitive firms divided into N "economic" sectors (i.e.,

[30] Real and nominal rigidities are not separately identified in our Calvo model with homogeneous price stickiness. As a result, comparisons based on the log posterior marginal density are sensitive to the prior on the index of real rigidities (even though we use a very uninformative prior). Hence, in this case we find it more appropriate to use a criterion based only on the likelihood.

[31] That is, the log posterior density of the data given the model is always higher under informative priors based on the distribution that excludes sales. However, the difference is smaller than in the model with Taylor pricing – about 1.5 log points – and does not decay as noticeably when we increase the degree of prior informativeness within the same range as we did for the Taylor model.

not necessarily identified by price stickiness). Sectors are indexed by $n \in \{1, ..., N\}$. The distribution of firms across sectors is summarized by a vector $(\varphi_1, ..., \varphi_N)$ with $\varphi_n > 0$, $\sum_{n=1}^{N} \varphi_n = 1$, where φ_n gives the mass of firms in sector n. Each sector has a (sector-specific) *stationary* cross-sectional distribution of price stickiness. Before setting its price, a firm j in economic sector n makes a draw for the duration of its next price spell, and then sets its price optimally. Notice that the price will be chosen according to the same policy as in the Taylor model (i.e., the optimal price for a spell that will last for a known duration). This implies that, at a given time, firms within a given sector can be further divided into different "groups" depending on the duration of price spells that they draw.

The (also stationary) cross-sectional distribution of price stickiness for the entire economy can be constructed by aggregating across sectors. It is summarized by a vector $(\omega_1, ..., \omega_K)$ with $\omega_k \equiv \sum_{n=1}^{N} \varphi_n \omega_{n,k} \in (0, 1)$. It is easy to show that $\sum_{k=1}^{K} \omega_k = 1$:

$$\sum\nolimits_{k=1}^{K} \sum\nolimits_{n=1}^{N} \varphi_n \omega_{n,k} = \sum\nolimits_{n=1}^{N} \sum\nolimits_{k=1}^{K} \varphi_n \omega_{n,k} = \sum\nolimits_{n=1}^{N} \varphi_n \sum\nolimits_{k=1}^{K} \omega_{n,k} = 1.$$

The exact details of how each firm draws the duration for the new price spell – that is, how firms move around different "stickiness groups" within a sector – is inconsequential for the aggregate dynamics implied by this model. What matters is our assumption that the cross-sectional distribution of price stickiness of each sector is stationary (i.e. $\omega_{n,k}$ is time-invariant), which guarantees the stationarity of the economy-wide distribution of price stickiness. In the Appendix we provide an example with a flexible scheme for drawing durations within each sector, which allows for persistence in the duration of price spells at the firm level.

We can write the log-linear approximate model implied by this "Random Taylor" price-setting scheme as:

$$\begin{aligned}
x_t(k) &= \frac{1-\beta}{1-\beta^k} E_t \sum_{i=0}^{k-1} \beta^i \left(p_{t+i} + \zeta \left(y_{t+i} - y_{t+i}^n \right) \right), \\
p_t &= \sum_{n=1}^{N} \varphi_n p_t(n), \\
p_t(n) &= \sum_{k=1}^{K} \omega_{n,k} p_t(n,k), \\
p_t(n,k) &= \int_0^1 p_t(n,k,j)\, dj = \frac{1}{k} \sum_{j=0}^{k-1} x_{t-j}(k).
\end{aligned}$$

Note that $p_t(n,k)$ does not depend on n. Thus, we can rewrite the aggregate price index as:

$$\begin{aligned} p_t &= \sum_{n=1}^{N} \varphi_n p_t(n) = \sum_{n=1}^{N} \varphi_n \sum_{k=1}^{K} \omega_{n,k} p_t(n,k) = \sum_{n=1}^{N} \varphi_n \sum_{k=1}^{K} \omega_{n,k} \frac{1}{k} \sum_{j=0}^{k-1} x_{t-j}(k) \\ &= \sum_{k=1}^{K} \underbrace{\sum_{n=1}^{N} \varphi_n \omega_{n,k}}_{=\omega_k} \frac{1}{k} \sum_{j=0}^{k-1} x_{t-j}(k) = \sum_{k=1}^{K} \omega_k \tilde{p}_t(k), \end{aligned}$$

where

$$\tilde{p}_t(k) \equiv p_t(n,k) = \frac{1}{k} \sum_{j=0}^{k-1} x_{t-j}(k).$$

That is, despite time-variation in the duration of price spells at the firm level, the Random Taylor model implies the exact same aggregate dynamics as our multi-sector Taylor pricing model. Moreover, it is easy to augment the model with other features that leave aggregate dynamics intact, and yet allow it to match additional micro facts.[32]

Hence, this alternative model provides a cautionary note on attempts to test specific models of price setting by confronting them with descriptive price-setting statistics (e.g., Klenow and Kryvtsov 2008).

6 Conclusion

If prices change frequently and each and every price change contributes to fully offset nominal disturbances, then nominal price rigidity cannot be the source of large and persistent monetary non-neutralities. Hence, bridging the micro-macro gap on the extent of price rigidity requires that price adjustments somehow fail to perfectly neutralize monetary innovations (i.e., require a large contract multiplier).

In this paper we use a standard sticky-price model to provide some evidence on three mechanisms that can provide such a boost to the contract multiplier. The model allows for varying degrees of real rigidities, and cross-sectional heterogeneity in price stickiness. It can extract some information about these two features of the economy from aggregate data, and discriminate between different distributions of price stickiness. Hence the model can also speak to the debate about the role of sales and other temporary price changes in shaping aggregate dynamics in monetary economies. We employ a Bayesian approach and combine macroeconomic time-series data with information about empirical distributions of price stickiness (with and without sales) derived from micro price data for the U.S. economy.

[32] For brevity we do not present details of the argument here, and refer the interested reader to Carvalho and Dam (2010).

We find that the estimated models can discriminate quite sharply between economies with heterogeneity in price stickiness and their homogeneous-firms counterparts. They also point to the existence of large real rigidities, which induce strong strategic complementarities in price setting. Finally, in terms of the cross-sectional distribution of price rigidity, we find that the distribution based on regular prices helps the model fit the data better. Altogether, our results suggest that all three mechanisms that can boost the so-called contract multiplier might have a role to play in our understanding of the effects of monetary policy.

The experience with our semi-structural model suggests that combining microeconomic and macroeconomic data within a Bayesian framework can help us integrate our views on nominal price rigidity at the micro and macro levels. We find the results sufficiently compelling to warrant additional research – perhaps in the context of fully specified DSGE models, where one may draw different policy implications.

Finally, as a by-product, we develop a price-setting model that produces the same aggregate dynamics as our multi-sector model with Taylor pricing and, yet, can match various empirical facts on price setting – including the evidence of variation in the duration of price spells at the quote-line level. Hence the model provides a cautionary note on attempts to test specific models of price setting by confronting them with descriptive price-setting statistics (e.g., Klenow and Kryvtsov 2008).

References

[1] Ball, L. and D. Romer (1990), "Real Rigidities and the Non-Neutrality of Money," *Review of Economic Studies* 57: 183-203.

[2] Baxter, M. and R. King (1999), "Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series," *Review of Economics and Statistics*, 81: 575-593.

[3] Bils, M. and P. Klenow (2004), "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy* 112: 947-985.

[4] Bonomo, M., C. Carvalho, R. Garcia, and V. Malta (2014), "Persistent Monetary Non-neutrality in an Estimated Model with Menu Costs and Partially Costly Information," mimeo available at https://sites.google.com/site/cvianac2/papers.

[5] Bouakez, H., E. Cardia and F. Ruge-Murcia (2009), "The Transmission of Monetary Policy in a Multi-Sector Economy," *International Economic Review* 50: 1243-1266.

[6] _____ (2014), "Sectoral Price Rigidity and Aggregate Dynamics," *European Economic Review* 65: 1-22

[7] Carvalho, C. (2006), "Heterogeneity in Price Stickiness and the Real Effects of Monetary Shocks," *Frontiers of Macroeconomics* 2, Issue 1, Article 1.

[8] Carvalho, C. and N. Dam (2010), "The Cross-Sectional Distribution of Price Stickiness Implied by Aggregate Data," FRB of New York Staff Report No. 419.

[9] Carvalho, C. and F. Schwartzman (2008), "Heterogeneous Price Setting Behavior and Aggregate Dynamics: Some General Results," mimeo available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2182343.

[10] _____ (2014), "Selection and Monetary Non-Neutrality in Time-Dependent Pricing Models," FRB Richmond Working Paper #12-09.

[11] Coenen, G., A. Levin, and K. Christoffel (2007), "Identifying the Influences of Nominal and Real Rigidities in Aggregate Price-Setting Behavior," *Journal of Monetary Economics* 54: 2439-2466.

[12] Coibion, O., Y. Gorodnichenko, and G. H. Hong (2014), "The Cyclicality of Sales, Regular, and Effective Prices: Business Cycle and Policy Implications," forthcoming in the *American Economic Review*.

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

[14] Dotsey, M., R. King, and A. Wolman (1997), "State Dependent Pricing and Dynamics of Business Cycles," Federal Reserve Bank of Richmond Working Paper Series 97-02.

[15] M. Eichenbaum, N. Jaimovich, S. Rebelo, and J. Smith (2014), "How Frequent Are Small Price Changes?" *American Economic Journal: Macroeconomics* 6: 137-55.

[16] Gelman, A., J. Carlin, H. Stern and D. Rubin (2003), *Bayesian Data Analysis*, 2nd edition, Chapman & Hall/CRC.

[17] Geweke, J. (1999), "Using simulation methods for Bayesian econometric models: inference, development and communication", *Econometric Review* 18: 1-126.

[18] Guerrieri, L. (2006), "The Inflation Persistence of Staggered Contracts," *Journal of Money, Credit and Banking* 38: 483-494.

[19] Guimaraes, B. and K. Sheedy, (2011), "Sales and Monetary Policy," *American Economic Review* 101: 844-76.

[20] Imbs, J., Jondeau, E., and F. Pelgrin (2011), "Sectoral Phillips Curves and the Aggregate Phillips Curve," *Journal of Monetary Economics* 58: 328–344.

[21] Jadresic, E. (1999), "Sticky Prices: An Empirical Assessment of Alternative Models," IMF Working Paper # 99/72.

[22] Kehoe, P. and V. Midrigan (2014), "Prices Are Sticky After All," forthcoming in the *Journal of Monetary Economics*.

[23] Klenow, P. and B. Malin (2011), "Microeconomic Evidence on Price Setting," in: *Handbook of Monetary Economics 3A*, B. Friedman and M. Woodford (eds.), Elsevier.

[24] Klenow, P., and O. Kryvtsov (2008), "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" *Quarterly Journal of Economics* 123: 863-904

[25] Klenow, P. and J. Willis (2006), "Real Rigidities and Nominal Price Changes," Federal Reserve Bank of Kansas City Research Working Paper 06-03, available at http://www.kc.frb.org/Publicat/Reswkpap/rwpmain.htm.

[26] Kryvtsov, O., and N. Vincent (2014), "On the Importance of Sales for Aggregate Price Flexibility", mimeo available at https://sites.google.com/site/okryvtsov/.

[27] Lee, Jae Won (2009), "Heterogeneous Households, Real Rigidity, and Estimated Duration of Price Contracts in a Sticky Price DSGE Model," mimeo available at https://sites.google.com/site/jwlee7/research.

[28] Maćkowiak, B. and F. Smets (2008), "On Implications of Micro Price Data for Macro Models," ECB Working Paper Series 960.

[29] Nakajima, J., N. Sudo, and T. Tsuruga (2010), "How Well do the Sticky Price Models Explain the Disaggregated Price Responses to Aggregate Technology and Monetary Policy Shocks?" IMES Discussion Paper No. 2010-E-22.

[30] Nakamura, E. and J. Steinsson (2008), "Five Facts About Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics* 123: 1415-1464.

[31] _____ (2010), "Monetary Non-Neutrality in a Multi-Sector Menu Cost Model," *Quarterly Journal of Economics* 125: 961-1013.

[32] Sbordone, A. (2002), "Prices and Unit Labor Costs: A New Test of Price Stickiness," *Journal of Monetary Economics* 49: 265-292.

[33] Sims, C. (2002), "Solving Linear Rational Expectations Models," *Computational Economics* 20: 1-20.

[34] Taylor, J. (1979), "Staggered Wage Setting in a Macro Model," *American Economic Review* 69: 108-113.

[35] _____ (1980), "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy* 88: 1-23.

[36] _____ (1993), *Macroeconomic Policy in a World Economy: From Econometric Design to Practical Operation*, W. W. Norton.

[37] Woodford, M. (2003), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.

Appendix

A Identification

A.1 When $\zeta \neq 1$

When $\zeta \neq 1$ equation (12) becomes:

$$p_t = \sum_{j=1}^{K-1} a_j p_{t-j} + \sum_{j=0}^{K-1} b_j m_{t-j} - \sum_{j=0}^{K-1} b_j y^n_{t-j},$$

where $a_1, ..., a_{K-1}$, b_0, ..., b_{K-1} are functions of the model parameters. Checking for identification amounts to solving for these coefficients, and showing that $\omega_1, ..., \omega_K$, and ζ can be recovered from them.

Here we illustrate how the process works in a model with $K = 2$. Using the method of undetermined coefficients we can show that a_1, b_0, b_1 satisfy:

$$\begin{aligned}
a_1 &= \frac{\frac{\omega_2}{2}\frac{1-\beta}{1-\beta^2}(1-\zeta)}{1-\left(\left(\omega_1+\frac{\omega_2}{2}\frac{1-\beta}{1-\beta^2}(1+\beta)\right)(1-\zeta)+\left(\frac{\omega_2}{2}\frac{1-\beta}{1-\beta^2}\beta\right)(1-\zeta)a_1\right)}, \\
b_0 &= \frac{\left(\omega_1+\frac{\omega_2}{2}\frac{1-\beta}{1-\beta^2}(1+\beta)\right)\zeta+\left(\frac{\omega_2}{2}\frac{1-\beta}{1-\beta^2}\beta\right)(\zeta\rho+(1-\zeta)b_1)}{1-\left(\left(\omega_1+\frac{\omega_2}{2}\frac{1-\beta}{1-\beta^2}(1+\beta)\right)(1-\zeta)+\left(\frac{\omega_2}{2}\frac{1-\beta}{1-\beta^2}\beta\right)(1-\zeta)a_1\right)-\left(\frac{\omega_2}{2}\frac{1-\beta}{1-\beta^2}\beta\right)(1-\zeta)\rho}, \\
b_1 &= \frac{\frac{\omega_2}{2}\frac{1-\beta}{1-\beta^2}\zeta}{1-\left(\left(\omega_1+\frac{\omega_2}{2}\frac{1-\beta}{1-\beta^2}(1+\beta)\right)(1-\zeta)+\left(\frac{\omega_2}{2}\frac{1-\beta}{1-\beta^2}\beta\right)(1-\zeta)a_1\right)}.
\end{aligned}$$

The first equation is quadratic in a_1 and for each solution the other two equations yield b_0 and b_1 as a function of the model parameters. The stable solution for the first equation ($|a_1| \leq 1$) yields:

$$\begin{aligned}
a_1 &= \frac{(1+\beta)(2\zeta+(1-\zeta)\omega_2)+\sqrt{(1+\beta)^2((\zeta-1)\omega_2-2\zeta)^2-4\beta(\zeta-1)^2\omega_2^2}}{2\beta(1-\zeta)\omega_2}, \\
b_0 &= \frac{\zeta(\rho-1)(\rho\beta-1)\omega_2}{2(1+\beta)\rho\zeta+(\zeta-1)(\rho-1)(\rho\beta-1)\omega_2} \\
&\quad+\frac{\zeta(1+\beta)(1+\beta(1+2\rho(\zeta-1)))}{\beta(\zeta-1)(2(1+\beta)\rho\zeta+(\zeta-1)(\rho-1)(\rho\beta-1)\omega_2)} \\
&\quad-\zeta(1+\beta)\frac{(1+\beta)2\zeta+\sqrt{4(1+\beta)^2\zeta^2-4(1+\beta)^2(\zeta-1)\zeta\omega_2+(\beta-1)^2(\zeta-1)^2\omega_2^2}}{\beta(\zeta-1)^2\omega_2(2(1+\beta)\rho\zeta+(\zeta-1)(\rho-1)(\rho\beta-1)\omega_2)},
\end{aligned}$$

$$b_1=\zeta\frac{(1+\beta)(2\zeta+(1-\zeta)\omega_2)+\sqrt{(1+\beta)^2((\zeta-1)\omega_2-2\zeta)^2-4\beta(\zeta-1)^2\omega_2^2}}{2\beta(1-\zeta)^2\omega_2},$$

where we have used the fact that $\omega_1+\omega_2=1$. Finally, we can the combine the expressions for a_1 and b_1 to solve for ω_2 and ζ:

$$\begin{aligned}\omega_2 &= \frac{2(1+\beta)b_1}{(1-a_1)(1-\beta a_1)},\\ \zeta &= \frac{b_1}{a_1+b_1}.\end{aligned}$$

A.2 Monte Carlo exercise

We generate artificial data on aggregate nominal and real output using a model with $K=4$, and parameter values that roughly resemble what we find when we estimate a model of this size on actual U.S. data. Then, we estimate the parameters of the model by maximum likelihood.[33] We conduct both a large- (1000 observations) and a small-sample exercise (100 observations, as in our actual sample). Table 7 reports the results.

Table 7: Monte Carlo - maximum likelihood estimation

		Large sample				*Small sample*			
	True	*Mean*	5^{th} *perc.*	95^{th} *perc.*	*Ini. guess*	*Mean*	5^{th} *perc.*	95^{th} *perc.*	*Ini. guess*
ζ	0.10	0.106	0.059	0.15	1.00	0.179	0.022	0.415	1.00
ω_1	0.40	0.395	0.183	0.621	0.25	0.318	0.033	0.871	0.25
ω_2	0.10	0.100	0.000	0.257	0.25	0.096	0.000	0.376	0.25
ω_3	0.10	0.091	0.000	0.197	0.25	0.088	0.000	0.304	0.25
ω_4	0.40	0.414	0.233	0.570	0.25	0.498	0.064	0.801	0.25
ρ_0	0.00	0.000	0.000	0.000	0.000	0.000	−0.002	0.002	0.000
ρ_1	1.43	1.432	1.388	1.468	1.429	1.403	1.256	1.547	1.538
ρ_2	−0.45	−0.456	−0.499	−0.410	−0.455	−0.446	−0.579	−0.302	−0.577
σ_m	0.005	0.005	0.0048	0.0051	0.005	0.005	0.0043	0.0056	0.0058
δ_0	0.00	0.000	−0.001	0.001	0.000	0.000	−0.004	0.004	0.000
δ_1	0.35	0.336	0.091	0.513	1.066	0.231	−0.257	0.616	0.954
δ_2	0.15	0.146	0.049	0.258	−0.199	0.133	−0.073	0.326	−0.076
σ_n	0.05	0.053	0.033	0.083	0.0067	0.105	0.020	0.311	0.0062

The first column shows the true parameter values used to generate the data. The columns under "Large sample" report statistics across 75 artificial samples of 1000 observations each. The "Small sample" columns report statistics across 240 artificial samples of 100 observations each.[34] The "Ini. guess" column reports the average value of the initial guesses supplied for the optimization algorithm across

[33] We apply the same procedure that we use in the initial maximization stage of the Markov Chain Monte Carlo algorithm that we use to estimate the models with actual data, including the choice of initial values for the optimization algorithm (see Subsection 3.5).

[34] In each replication, the sample contains an additional 16 observations that we use as a pre-sample to initialize the Kalman filter, as we do in the actual estimation. The value of β is fixed at 0.99. The smaller number of replications for the large-sample exercise is simply due to its much higher computational cost.

the corresponding samples. Following the procedure that we use in the actual estimation algorithm, the initial guesses for ζ and $\omega_1 - \omega_4$ are the same across replications; the guesses for the remaining parameters in each replication are set equal to the ordinary least squares estimates based on nominal output (for the ρ's) and *actual* output (for the δ's).

B Details of the estimation algorithm

Our specific estimation strategy is as follows. We run two numerical optimization routines sequentially in order to maximize the posterior distribution. This determines the starting point of the Markov chain and provides a first crude estimate of the covariance matrix for our Random-Walk Metropolis Gaussian jumping distribution. The first optimization routine is `csminwel` by Chris Sims, while the second is `fminsearch` from Matlab's optimization toolbox. For the starting values, we set $\zeta = 1$ and $\omega_k = 1/K$; the values for the remaining parameters are set equal to the ordinary least squares estimates based on nominal output (for the ρ's) and *actual* output (for the δ's). Following the first optimization, we run additional rounds, starting from initial values obtained by perturbing the original initial values, and then the estimate of the first optimization round.

Before running the Markov chains we transform all parameters to have full support on the real line. We use the logarithmic transformation for each of $(\zeta, \sigma_m, \sigma_n)$, while $\omega_1, ..., \omega_K$ are transformed using a multivariate logistic function (see next subsection). Then we run a so-called *adaptive phase* of the Markov chain, with three sub-phases of 100, 200, and 600 thousand iterations, respectively. At the end of each sub-phase we discard the first half of the draws, update the estimate of the posterior mode, and compute a sample covariance matrix to be used in the jumping distribution in the next sub-phase. Finally, in each sub-phase we rescale the covariance matrix inherited from the previous sub-phase in order to get a *fine-tuned covariance matrix* that yields rejection rates as close as possible to 0.77.[35] Next we run the so-called *fixed phase* of the MCMC. We take the estimate of the posterior mode and sample covariance matrix from the adaptive phase, and run 5 parallel chains of 300,000 iterations each. Again, before making the draws that will form the sample we rescale such covariance matrix in order to get rejection rates as close as possible to 0.77. To initialize each chain we draw from a candidate normal distribution centered on the posterior mode estimate, with covariance matrix given by 9 times the fine-tuned covariance matrix. We check for convergence for the latter 2/3s of the draws of all 5 chains by calculating the potential scale reduction[36] (PSR) factors for each parameter and inspecting the

[35] This is the optimal rejection rate under certain conditions. See Gelman et al. (2003, p. 306).

[36] For each parameter, the PSR factor is the ratio of (square root of) an estimate of the marginal posterior variance to the average variance within each chain. This factor expresses the potential reduction in the scaling of the estimated marginal posterior variance relative to the true distribution by increasing the number of iterations in the Markov-chain algorithm.

histograms of all marginal distributions across the parallel chains. Upon convergence, the latter 2/3s of the draws of all 5 chains are combined to form a posterior sample of 1 million draws.

B.1 Transformation of the sectoral weights

We transform vectors $\omega = (\omega_1, ..., \omega_K)$ in the K-dimensional unit simplex into vectors $v = (v_1, ..., v_K)$ in $\mathbb{R}^K$ using the inverse of a restricted multivariate logistic transformation. We want to be able to draw v's and then use a transformation that guarantees that $\omega = h^{-1}(v)$ is in the K-dimensional unit simplex. For that purpose, we start with:

$$\omega_k = \frac{e^{v_k}}{\sum_{k=1}^{K} e^{v_k}}, k = 1, ..., K.$$

The transformation above guarantees the non-negativity and summation-to-unity constraints. However, without additional restrictions the mapping is not one-to-one. The reason is that all vectors v along the same *ray* give rise to the same ω. Therefore, we impose the restriction $v(K) = 0$ and in effect draw vectors $\widetilde{v} = (v_1, ..., v_{K-1})$ in $\mathbb{R}^{K-1}$. Thus, the transformation becomes $\widetilde{\omega} = \widetilde{h}^{-1}(\widetilde{v})$, with $\widetilde{\omega} = (\omega_1, ..., \omega_{K-1})$ and:

$$\begin{aligned} \omega_k &= \frac{e^{v_k}}{1 + \sum_{k=1}^{K-1} e^{v_k}}, k = 1, ..., K-1 \\ \omega_K &= \frac{1}{1 + \sum_{k=1}^{K-1} e^{v_k}}. \end{aligned}$$

If the density $f_\omega(\omega|\alpha)$ is that of the Dirichlet distribution with (vector) parameter α, the density of $\widetilde{v}$ is given by:

$$f_{\widetilde{v}}(\widetilde{v}|\alpha) = |J| f_\omega \left(\frac{e^{v_1}}{1 + \sum_{k=1}^{K-1} e^{v_k}}, ..., \frac{1}{1 + \sum_{k=1}^{K-1} e^{v_k}} | \alpha \right),$$

where $|J|$ is the determinant of the Jacobian matrix $\left[\frac{\partial \widetilde{h}^{-1}(\widetilde{v})}{\partial \widetilde{v}}\right]_{ij}$ given by:

$$\begin{bmatrix} \frac{\partial \omega_1}{\partial v_1} & \frac{\partial \omega_1}{\partial v_2} & \cdots & \frac{\partial \omega_1}{\partial v_{K-1}} \\ \frac{\partial \omega_2}{\partial v_1} & \frac{\partial \omega_2}{\partial v_2} & \cdots & \frac{\partial \omega_2}{\partial v_{K-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \omega_{K-1}}{\partial v_1} & \frac{\partial \omega_{K-1}}{\partial v_2} & \cdots & \frac{\partial \omega_{K-1}}{\partial v_{K-1}} \end{bmatrix},$$

with:

Hence, as the PSR factor approaches unity, it is a sign of convergence of the Markov-chain for the estimated parameter. See Gelman et al. (2003, p. 294 ff) for more information. For all specifications we require that the factor be below 1.01 for all parameters.

$$\begin{aligned}\frac{\partial \omega_k}{\partial v_k} &= \frac{e^{v_k}\left(1+\sum_{k=1}^{K-1}e^{v_k}\right)-e^{v_k}e^{v_k}}{\left(1+\sum_{k=1}^{K-1}e^{v_k}\right)^2} \\ &= \frac{e^{v_k}}{1+\sum_{k=1}^{K-1}e^{v_k}}-\frac{e^{v_k}e^{v_k}}{\left(1+\sum_{k=1}^{K-1}e^{v_k}\right)^2}.\end{aligned}$$

So:

$$J=-\begin{bmatrix}\frac{e^{v_1}}{1+\sum_{k=1}^{K-1}e^{v_k}}\\ \vdots \\ \frac{e^{v_{K-1}}}{1+\sum_{k=1}^{K-1}e^{v_k}}\end{bmatrix}\left[\frac{e^{v_1}}{1+\sum_{k=1}^{K-1}e^{v_k}},...,\frac{e^{v_{K-1}}}{1+\sum_{k=1}^{K-1}e^{v_k}}\right]+\begin{bmatrix}\frac{e^{v_1}}{1+\sum_{k=1}^{K-1}e^{v_k}} & 0 & ... & 0\\ 0 & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & 0\\ 0 & ... & 0 & \frac{e^{v_{K-1}}}{1+\sum_{k=1}^{K-1}e^{v_k}}\end{bmatrix}.$$

To recover the v_k's from ω simply set:

$$v_k=\log\left(\omega_k\right)-\log\left(\omega_K\right).$$

C Identification in a multi-sector Calvo (1983) model

Under the Calvo (1983) model, the optimal reset price equation becomes:

$$x_t\left(k\right)=\left(1-\beta\lambda_k\right)E_t\sum_{i=0}^{\infty}\left(\lambda_k\beta\right)^i\left(p_{t+i}+\zeta\left(y_{t+i}-y_{t+i}^n\right)\right), \tag{15}$$

and sectoral price equations become:

$$p_t\left(k\right)=\int_0^1 p_t\left(k,j\right)dj=\left(1-\beta\lambda_k\right)\sum_{i=0}^{\infty}\lambda_k^i x_{t-i}\left(k\right),$$

where $1-\lambda_k$ is the frequency of price changes in sector k. The remaining equations of the model are:

$$p_t=\sum_{k=1}^{K}\omega_k p_t\left(k\right),$$

and, for simplicity, AR(1) processes for m_t and y_t^n:

$$p_t+y_t=m_t=\rho_1 m_{t-1}+\sigma_m\widetilde{\varepsilon}_t^m$$

$$y_t^n = \delta_1 y_{t-1}^n + \sigma_n \widetilde{\varepsilon}_t^n.$$

The optimal reset price can also be written as:

$$x_t(k) = (1-\lambda_k\beta) E_t \sum_{i=0}^{\infty} (\lambda_k\beta)^i \left(\zeta\left(m_{t+i} - y_{t+i}^n\right) + (1-\zeta) p_{t+i}\right).$$

The equilibrium price level can be generally expressed as a weighted sum of the current and past exogenous variables:

$$p_t = \sum_{j=0}^{\infty} b_j m_{t-j} + \sum_{j=0}^{\infty} c_j y_{t-j}^n,$$

where $\{b_j, c_j\}_{j=1,...,\infty}$ are functions of the model parameters. If we observe m_t and y_t – and thus p_t – the coefficients on m_{t-j} can be estimated. Thus, checking for identification of the cross-sectional distribution amounts to asking if knowledge of the coefficients on lagged nominal output, $\{b_j\}$, allows us to pin down the sectoral weights ω – and ζ.[37]

We start with the case of strategic neutrality in price setting ($\zeta = 1$). Then, new prices are set according to:

$$x_t(k) = \left(\frac{1-\lambda_k\beta}{1-\rho_1\lambda_k\beta}\right) m_t - \left(\frac{1-\lambda_k\beta}{1-\delta_1\lambda_k\beta}\right) y_t^n.$$

Replacing this expression for newly set prices in the sectoral price equation and aggregating produces the following expression for the aggregate price level:

$$p_t = \sum_{j=0}^{\infty} \left(\sum_{k=1}^{K} F_j(\beta, \rho_1, \lambda_k)\,\omega_k\right) m_{t-j} - \sum_{j=0}^{\infty} \left(\sum_{k=1}^{K} F_j(\beta, \delta_1, \lambda_k)\,\omega_k\right) y_{t-j}^n, \tag{16}$$

where

$$F_j(\beta, \gamma, \lambda_k) \equiv \frac{(1-\lambda_k\beta)^2}{1-\gamma\lambda_k\beta}\lambda_k^j.$$

From estimates of the coefficients on m_{t-j} in (16), we can recover the sectoral weights ω. The reason is that $F_j(\beta, \rho_1, \lambda_k)$ is "known", since ρ_1 can be estimated directly from observation of m_t. Notice that each of the coefficients on m_{t-j} is a weighted sum of the sectoral weights. Thus we can infer $(\omega_1, ..., \omega_K)$ by solving a system of linear equations:

$$b_j = \sum_{k=1}^{K} F_j(\beta, \rho_1, \lambda_k)\,\omega_k, \quad j = 0, 1, 2, .., \tag{17}$$

where $F_j(\beta, \rho_1, \lambda_k)$ and b_j are known constants for every j.

[37] In analogy with the Taylor model, we assume throughout that the discount factor, β, and the Calvo stickiness parameter for each sector, $\lambda_k \equiv \frac{k-1}{k}$, are known.

The intuition behind the identification result is the same as before: sectors where prices are more sticky are relatively more important in determining the impact of older shocks to the exogenous processes on the current price level, and vice-versa for sectors where prices are more flexible. This structure of the model provides a source of identification of sectoral weights. Indeed it is straightforward to show that:

$$\frac{\partial F_j(\beta,\rho_1,\lambda)}{\partial j}<0 \quad \text{and} \quad \frac{\partial F_j(\beta,\rho_1,\lambda)}{\partial j\partial\lambda}>0.$$

The first partial derivative indicates that each sector's influence on the coefficient on m_{t-j} diminishes as j increases because, as time passes, more firms in each sector adjust their prices in response to past innovations. In addition, the decay process happens faster in sectors with more price flexibility. This makes the linear system (17) non-degenerate, and allow us to uniquely pin down the sectoral weights $(\omega_1,...,\omega_K)$ from knowledge of $(b_0,b_1,b_2,...)$.

The same intuition carries over to the more general case in which pricing interactions are present, although proving identification of ω in this case is a bit more involved. New prices $x_t(k)$ depend not only on the two exogenous processes m_t and y_t^n, but also on current and expected future values of the aggregate price level. Consequently, each of the coefficients on m_{t-j} is no longer a linear combination of $(\omega_1,...,\omega_K)$ but a nonlinear function of the the sectoral weights as well as the real rigidity index ζ:[38]

$$b_j = G_j(\omega_1,\omega_2,...,\omega_K,\zeta), \quad j=0,1,2,..,$$

which is a nonlinear counterpart of (17). Identifying ω and ζ therefore involves solving a system of non-linear equations. Similarly to the previous case, however, the relative importance of $(\omega_1,\omega_2,...,\omega_K,\zeta)$ in the coefficient on m_{t-j} varies with j, which essentially allows us to pin down those parameters.

For illustration, consider a simple economy divided into two sectors ($K=2$), one of which is characterized by fully flexible prices ($\lambda_1=0$). To make our analysis even simpler, assume further that m_t follows the i.i.d. process ($\rho_1=...=\rho_{p1}=0$) and ignore y_t^n ($y_t^n=0\ \forall t$). In this case, the aggregate price level can be obtained as:

$$\begin{aligned} p_t &= \sum_{j=0}^{\infty} b_j m_{t-j} = \left[\omega_1(\omega_1+\tilde{\omega}_2)^{-1}+\tilde{\omega}_2(\omega_1+\tilde{\omega}_2)^{-2}\kappa f(\omega_1+\tilde{\omega}_2)\right] m_t \\ &\quad +\tilde{\omega}_2(\omega_1+\tilde{\omega}_2)^{-2}\kappa\sum_{j=1}^{\infty} f(\omega_1+\tilde{\omega}_2)^{j+1} m_{t-j}, \end{aligned} \tag{18}$$

[38] Besides the sectoral weights and real rigidity index, function G_j depends also on the (known) parameters $(\beta,\lambda_1,...,\lambda_K)$ and the autoregressive parameter ρ_1, which can be estimated directly from observation of m_t. This dependence is omitted to save on notation.

where κ is a constant that depends on β and λ_2, $\tilde{\omega}_2 \equiv \frac{\omega_2}{\zeta}$ is a "*scaled* ω_2" – the mass of firms in sector 2 adjusted by the index of real rigidity – and f is a function that is strictly increasing:

$$f\left(\omega_1 + \tilde{\omega}_2\right) \equiv \frac{\left(1 + \beta + \kappa\left(\omega_1 + \tilde{\omega}_2\right)^{-1}\right) - \sqrt{\left(1 + \beta + \kappa\left(\omega_1 + \tilde{\omega}_2\right)^{-1}\right)^2 - 4\beta}}{2\beta}.$$

Thus, unlike the previous case without pricing interactions, each of the coefficients on m_{t-j} is now a nonlinear function of ω and ζ, as mentioned earlier.

To check for identification, notice that the ratio of the coefficient on m_{t-j} to that of m_{t-j-1} satisfies:

$$\frac{b_{j+1}}{b_j} = f\left(\omega_1 + \tilde{\omega}_2\right) \text{ for } j \geq 1,$$

which allows us to identify the sum, $\omega_1 + \tilde{\omega}_2$, as $f^{-1}\left(\frac{b_{j+1}}{b_j}\right)$. In turn, we can identify ω_1 and $\tilde{\omega}_2$ by solving two linear equations:

$$b_0 = \omega_1 \left[f^{-1}\left(\frac{b_{j+1}}{b_j}\right)\right]^{-1} + \tilde{\omega}_2 \left[f^{-1}\left(\frac{b_{j+1}}{b_j}\right)\right]^{-2} \left(\frac{b_{j+1}}{b_j}\right) \kappa, \quad j \geq 1, \tag{19}$$

$$b_j = \tilde{\omega}_2 \left[f^{-1}\left(\frac{b_{j+1}}{b_j}\right)\right]^{-2} \left(\frac{b_{j+1}}{b_j}\right)^{j+1} \kappa, \quad j \geq 1. \tag{20}$$

Our analysis – especially the last two equations (19) and (20) – illustrates that the same economic mechanism works in the presence of pricing interactions. The flexible-price sector (sector 1) is relatively more important in the response of the price level to more recent shocks (b_0 in this example), and vice versa for the sticky-price sector. Also similarly to the previous case, we can see from (19) and (20) that while the coefficients on ω_1 and $\tilde{\omega}_2$ decrease as j increases, the coefficient on the latter decays more slowly, which allows for identification of ω_1 and $\tilde{\omega}_2$.

Finally, it remains to recover ω_2 and ζ from the scaled weight, $\tilde{\omega}_2$. At first pass, this may seem a hard task since, intuitively, as the degree of real rigidities increases (i.e. as ζ gets smaller), firms with flexible prices will react more gradually to innovations to m_t and y_t^n than they would under no pricing interactions ($\zeta = 1$). Thus, even with a smaller size of sticky-price sector (a lower ω_2), the model can generate similar dynamics of the aggregate price level. In this sense, nominal and real rigidities are substitutes. However, in the model with heterogeneity they are *not* perfect substitutes, as they are in the standard version of the Calvo model with homogeneous price stickiness. To see that, notice that a change in ω_2 affects also the weights of the other sectors (ω_1 in the present example), as sectoral weights must sum to one. Hence, it affects aggregate dynamics differently from an equivalent variation in ζ that would leave $\tilde{\omega}_2$ unchanged. In the simple case considered here, we can easily obtain ω_2 from

$\omega_1+\omega_2=1$, and then ζ from $\tilde{\omega}_2$. More generally, ω and ζ enter the coefficients on m_{t-j} asymmetrically, which allows us to distinguish the role of real rigidities from that of nominal stickiness in the dynamics of the aggregate price level.

We now provide some more detail on identification, and solve for sectoral weights and real rigidities in a model with price stickiness in all sectors. To that end, it is useful to derive a Phillips curve relation for each sector using the price-setting equations under Calvo pricing:

$$\pi_t(k) = g(\lambda_k)\zeta\left(y_t - y_t^n\right) + \beta E_t\left[\pi_{t+1}(k)\right] - g(\lambda_k)\left(p_t(k) - p_t\right), \quad \text{for } k = 1, 2, ..., K, \tag{21}$$

where $\pi_t(k) \equiv p_t(k) - p_{t-1}(k)$ and $g(\lambda_k) \equiv \frac{(1-\lambda_k)(1-\lambda_k\beta)}{\lambda_k}$.

Here we illustrate how the sectoral weights are identified in a model with $K = 2$, and real rigidities. Define the "relative price" as:

$$p_t^R \equiv p_t(1) - p_t(2).$$

We can rewrite (21) using the relative price as:

$$\pi_t(1) = g(\lambda_1)\zeta\left(y_t - y_t^n\right) + \beta E_t\left[\pi_{t+1}(1)\right] - \omega_2 g(\lambda_1) p_t^R, \tag{22}$$

$$\pi_t(2) = g(\lambda_2)\zeta\left(y_t - y_t^n\right) + \beta E_t\left[\pi_{t+1}(2)\right] + \omega_1 g(\lambda_2) p_t^R. \tag{23}$$

To help with the exposition, we start by looking again at the special case where one of the sectors is characterized by fully flexible prices ($\lambda_1 = 0$), and m_t follows an i.i.d. process, while $y_t^n = 0$. We then proceed to consider a general two-sector model where λ_1 can be nonzero and m_t and y_t^n can follow more general stationary processes.

C.1 A special case: $\lambda_2 > \lambda_1 = 0$.

First, note that the aggregate price level is given by:

$$p_t = \omega_1 p_t(1) + \omega_2 p_t(2) = \omega_1\left\{\left(1-\zeta\right)p_t + \zeta m_t\right\} + \omega_2 p_t(2).$$

Solving for the price level, we obtain:

$$p_t = \frac{\omega_1}{\Sigma_\omega} m_t + \frac{\tilde{\omega}_2}{\Sigma_\omega} p_t(2) \qquad \text{where } \Sigma_\omega \equiv \omega_1 + \tilde{\omega}_2. \tag{24}$$

Equation (24) indicates the aggregate price level is a weighted average of m_t and $p_t(2)$ – the average price of firms in the sticky-price sector. In the absence of pricing interactions ($\zeta = 1$), $\Sigma_\omega = 1$ and

$p_t = \omega_1 m_t + \omega_2 p_t(2)$. In general, however, as the degree of strategic complementarities increases (i.e., as ζ decreases), $\frac{\omega_1}{\Sigma_\omega}$ decreases while $\frac{\tilde{\omega}_2}{\Sigma_\omega}$ increases, reflecting the fact that the aggregate price level becomes more sluggish because flexible-price firms' pricing decisions are disproportionately influenced by prices set by firms in sector 2 (this is the "strategic interaction effect" in Carvalho 2006).

When $\lambda_1 = 0$, $g(\lambda_1)$ is infinite. By dividing (22) by $g(\lambda_1)$, we can show that:

$$p_t^R = \frac{\zeta}{\omega_2} y_t = \frac{\zeta}{\omega_2} (m_t - p_t). \tag{25}$$

To obtain the equilibrium level of $p_t(2)$, rewrite (23) using (25):

$$\pi_t(2) = g(\lambda_2)\zeta y_t + \beta E_t [\pi_{t+1}(2)] + g(\lambda_2)\zeta \frac{\omega_1}{\omega_2} y_t = \beta E_t [\pi_{t+1}(2)] + g(\lambda_2)\frac{\zeta}{\omega_2} (m_t - p_t),$$

which can be expressed as a second-order linear difference equation of $p_t(2)$, employing (24):

$$\beta E_t [p_{t+1}(2)] - \left(1 + \beta + g(\lambda_2)\Sigma_\omega^{-1}\right) p_t(2) + p_{t-1}(2) = g(\lambda_2)\Sigma_\omega^{-1} m_t.$$

The solution for the linear difference equation above is:

$$p_t(2) = f(\Sigma_\omega)\, p_{t-1}(2) + g(\lambda_2)\Sigma_\omega^{-1} \sum_{j=0}^{\infty} f(\Sigma_\omega)^{j+1} E_t [m_{t+j}], \tag{26}$$

where

$$f(\Sigma_\omega) = \frac{\left(1 + \beta + g(\lambda_2)\Sigma_\omega^{-1}\right) - \sqrt{\left(1 + \beta + g(\lambda_2)\Sigma_\omega^{-1}\right)^2 - 4\beta}}{2\beta}.$$

It is straightforward to show that there is a one-to-one mapping between Σ_ω and f, and moreover that f is strictly increasing in Σ_ω:

$$\frac{\partial f}{\partial \Sigma_\omega} > 0 \quad \text{with} \quad \lim_{\Sigma_\omega \to 0} f = 0 \quad \text{and} \quad \lim_{\Sigma_\omega \to \infty} f = 1.$$

Since m_t is i.i.d., (26) can be simplified to:

$$\begin{aligned} p_t(2) &= f(\Sigma_\omega)\, p_{t-1}(2) + g(\lambda_2)\Sigma_\omega^{-1} f(\Sigma_\omega)\, m_t \\ &= g(\lambda_2)\Sigma_\omega^{-1} \sum_{j=0}^{\infty} f(\Sigma_\omega)^{j+1} m_{t-j}. \end{aligned}$$

Combine this with (24) to get:

$$p_t = \frac{\omega_1}{\Sigma_\omega} m_t + \frac{\widetilde{\omega}_2}{\Sigma_\omega^2} g(\lambda_2) \sum_{j=0}^{\infty} f\left(\Sigma_\omega\right)^{j+1} m_{t-j},$$

which is exactly equation (18) presented before. Once again, it is tedious yet straightforward to recover $\{\zeta, \omega_1, \omega_2\}$ from $\{b_j\}$:

$$\begin{aligned} \omega_2 &= \frac{b_0 + \Sigma_\omega^{-2} g(\lambda_2) \frac{b_2}{b_1} \left(1 - \Sigma_\omega\right) - \Sigma_\omega^{-1}}{\Sigma_\omega^{-2} g(\lambda_2) \frac{b_2}{b_1} - \Sigma_\omega^{-1}}, \\ \omega_1 &= 1 - \omega_2, \\ \zeta &= \frac{\omega_2}{\Sigma_\omega - \omega_1}, \end{aligned}$$

where $\Sigma_\omega = f^{-1}\left(\frac{b_2}{b_1}\right)$.

As a final note, this model nests a one-sector model with price spells of expected length k as a trivial case. For a *known* k, $\omega_k = 1$ while $\omega_{k'} = 0$ for all $k' \neq k$. In turn, the real rigidity index ζ is trivially identified as $\Sigma_\omega^{-1} = \frac{1}{f^{-1}\left(\frac{b_2}{b_1}\right)}$.

C.2 A more general case: $\lambda_2 > \lambda_1 > 0$

We can derive a Phillips curve for the aggregate price level by aggregating (22) and (23):

$$\pi_t = \left\{g(\lambda_1)\omega_1 + g(\lambda_2)\omega_2\right\} \zeta \left(y_t - y_t^n\right) + \beta E_t\left[\pi_{t+1}\right] - \left\{g(\lambda_1) - g(\lambda_2)\right\} \omega_1 \omega_2 p_t^R. \tag{27}$$

Notice that the relative price p_t^R affects aggregate dynamics. To obtain the equilibrium relative price, we subtract (23) from (22) to get:

$$\pi_t^R = \left\{g(\lambda_1) - g(\lambda_2)\right\} \zeta \left(y_t - y_t^n\right) + \beta E_t\left[\pi_{t+1}^R\right] - \left\{g(\lambda_1)\omega_2 + g(\lambda_2)\omega_1\right\} p_t^R, \tag{28}$$

where $\pi_t^R = p_t^R - p_{t-1}^R$. We then rewrite the system of two equations (27) and (28) as:

$$\beta E_t\left[p_{t+1}\right] - \left(\beta + 1 + A\right) p_t + p_{t-1} = B p_t^R - A m_t + A y_t^n \tag{29}$$

$$\beta E_t\left[p_{t+1}^R\right] - \left(\beta + 1 + D\right) p_t^R + p_{t-1}^R = C p_t - C m_t + C y_t^n, \tag{30}$$

where A, B, C and D have positive values and are functions of the parameters $\{\omega_k\}$ and ζ, as well as known constants, $g(\lambda_1)$ and $g(\lambda_2)$ (with $g(\lambda_1) > g(\lambda_2)$):

$$\begin{aligned} A &= \{g(\lambda_1)\omega_1 + g(\lambda_2)\omega_2\}\,\zeta \\ B &= \{g(\lambda_1) - g(\lambda_2)\}\,\omega_1\omega_2 \\ C &= \{g(\lambda_1) - g(\lambda_2)\}\,\zeta \\ D &= \{g(\lambda_1)\omega_2 + g(\lambda_2)\omega_1\}\,. \end{aligned}$$

Equations (29) and (30) jointly determine the dynamics of $\left\{p_t, p_t^R\right\}$, given exogenous process for $\{m_t, y_t^n\}$.

Unlike the two special cases discussed in the main text, it is difficult to solve the model analytically. Our strategy here is to take two *potentially* different sets of values of $\{\zeta, \omega_1, \omega_2\}$, denoted $\{\zeta^*, \omega_1^*, \omega_2^*\}$ and $\left\{\zeta^{\#}, \omega_1^{\#}, \omega_2^{\#}\right\}$, and then examine if they can produce the same impulse response function of the aggregate price level to a one-unit increase in m_t.

Let

$$\begin{aligned} A^* &= \{g(\lambda_1)\omega_1^* + g(\lambda_2)\omega_2^*\}\,\zeta^*, & A^{\#} &= \left\{g(\lambda_1)\omega_1^{\#} + g(\lambda_2)\omega_2^{\#}\right\}\zeta^{\#}, \\ B^* &= \{g(\lambda_1) - g(\lambda_2)\}\,\omega_1^*\omega_2^*, & B^{\#} &= \{g(\lambda_1) - g(\lambda_2)\}\,\omega_1^{\#}\omega_2^{\#}, \\ C^* &= \{g(\lambda_1) - g(\lambda_2)\}\,\zeta^*, & C^{\#} &= \{g(\lambda_1) - g(\lambda_2)\}\,\zeta^{\#}, \\ D^* &= \{g(\lambda_1)\omega_2^* + g(\lambda_2)\omega_1^*\}\,, & D^{\#} &= \left\{g(\lambda_1)\omega_2^{\#} + g(\lambda_2)\omega_1^{\#}\right\}. \end{aligned}$$

Introduce some notation for the impulse response functions:

$$\begin{aligned} \mathcal{P}_j &= \frac{\partial E_t\left[p_{t+j}\right]}{\partial m_t}, \\ \mathcal{P}_j^R &= \frac{\partial E_t\left[p_{t+j}^R\right]}{\partial m_t}, \\ \mathcal{M}_j &= \frac{\partial E_t\left[m_{t+j}\right]}{\partial m_t}. \end{aligned}$$

Note that $\frac{\partial E_t\left[y_{t+j}^n\right]}{\partial m_t} = 0$ for all j, and hence can be ignored. We then construct linear difference equations for $\mathcal{P}_j$ and $\mathcal{P}_j^R$ from (29) and (30):

$$\beta\mathcal{P}_{j+1} - (\beta + 1 + A)\,\mathcal{P}_j + \mathcal{P}_{j-1} = B\mathcal{P}_j^R - A\mathcal{M}_j, \tag{31}$$

$$\beta\mathcal{P}_{j+1}^R - (\beta + 1 + D)\,\mathcal{P}_j^R + \mathcal{P}_{j-1}^R = C\mathcal{P}_j - C\mathcal{M}_j. \tag{32}$$

Equation (32) implies:

$$\begin{aligned}
\mathcal{P}_{j+1}^R &= \frac{(\beta+1+D)}{\beta}\mathcal{P}_j^R - \frac{1}{\beta}\mathcal{P}_{j-1}^R + \frac{C}{\beta}\mathcal{P}_j - \frac{C}{\beta}\mathcal{M}_j \\
&= \frac{(\beta+1+D)}{\beta}\left[\frac{(\beta+1+D)}{\beta}\mathcal{P}_{j-1}^R - \frac{1}{\beta}\mathcal{P}_{j-2}^R + \frac{C}{\beta}\mathcal{P}_{j-1} - \frac{C}{\beta}\mathcal{M}_{j-1}\right] \\
&\quad -\frac{1}{\beta}\left[\frac{(\beta+1+D)}{\beta}\mathcal{P}_{j-2}^R - \frac{1}{\beta}\mathcal{P}_{j-3}^R + \frac{C}{\beta}\mathcal{P}_{j-2} - \frac{C}{\beta}\mathcal{M}_{j-2}\right] + \frac{C}{\beta}\mathcal{P}_j - \frac{C}{\beta}\mathcal{M}_j \\
&= \left[\frac{(\beta+1+D)}{\beta}\right]^2\mathcal{P}_{j-1}^R - \frac{2}{\beta}\frac{(\beta+1+D)}{\beta}\mathcal{P}_{j-2}^R + \frac{1}{\beta^2}\mathcal{P}_{j-3}^R \\
&\quad +\frac{C}{\beta}\mathcal{P}_j + \frac{(\beta+1+D)}{\beta}\frac{C}{\beta}\mathcal{P}_{j-1} - \frac{1}{\beta}\frac{C}{\beta}\mathcal{P}_{j-2} \\
&\quad -\frac{C}{\beta}\mathcal{M}_j - \frac{(\beta+1+D)}{\beta}\frac{C}{\beta}\mathcal{M}_{j-1} - \frac{1}{\beta}\frac{C}{\beta}\mathcal{M}_{j-2}.
\end{aligned}$$

We can rewrite the equation above as:

$$\mathcal{P}_j^R = \frac{C}{\beta}\mathcal{P}_{j-1} + \frac{(\beta+1+D)}{\beta}\frac{C}{\beta}\mathcal{P}_{j-2} - \frac{C}{\beta}\mathcal{M}_{j-1} - \frac{(\beta+1+D)}{\beta}\frac{C}{\beta}\mathcal{M}_{j-2} + \Omega(\mathcal{P}_{j-3}, \mathcal{M}_{j-3}), \tag{33}$$

where $\Omega(\mathcal{P}_{j-3}, \mathcal{M}_{j-3})$ is a linear function of $\{\mathcal{P}_{j-3}, \mathcal{P}_{j-4}, \ldots, \mathcal{M}_{j-3}, \mathcal{M}_{j-4}, \ldots\}$. Plug (33) into (31) to get:

$$\begin{aligned}
\mathcal{P}_{j+1} &= \frac{(\beta+1+A)}{\beta}\mathcal{P}_j + \left(\frac{BC}{\beta^2} - \frac{1}{\beta}\right)\mathcal{P}_{j-1} + \frac{(\beta+1+D)\,BC}{\beta^3}\mathcal{P}_{j-2} \\
&\quad -\frac{A}{\beta}\mathcal{M}_j - \frac{BC}{\beta^2}\mathcal{M}_{j-1} - \frac{(\beta+1+D)\,BC}{\beta^3}\mathcal{M}_{j-2} \\
&\quad +\frac{B}{\beta}\Omega(\mathcal{P}_{j-3}, \mathcal{M}_{j-3}).
\end{aligned}$$

Notice that $\mathcal{P}_{-1} = \mathcal{P}_{-2} = \cdots = \mathcal{M}_{-1} = \mathcal{M}_{-2} = \cdots = 0$.

Consider the impulse responses of the aggregate price level to a one-unit increase in m_t under $\{\zeta^*, \omega_1^*, \omega_2^*\}$ and $\left\{\zeta^{\#}, \omega_1^{\#}, \omega_2^{\#}\right\}$, and assume $\mathcal{P}_j$ and $\mathcal{M}_j$ are equal for all j between the two parameterization. The impulse response of p_{t+1} is given as:

$$\frac{\partial E_t\left[p_{t+1}\right]}{\partial m_t} = \mathcal{P}_1 = \frac{(\beta+1+A)}{\beta}\mathcal{P}_0 - \frac{A}{\beta}\mathcal{M}_0 = \frac{(\beta+1+A)}{\beta}\mathcal{P}_0 - \frac{A}{\beta}. \quad (\because \mathcal{M}_0 = 1)$$

It follows that

$$\frac{(\beta+1+A^*)}{\beta}\mathcal{P}_0 - \frac{A^*}{\beta} = \frac{\left(\beta+1+A^{\#}\right)}{\beta}\mathcal{P}_0 - \frac{A^{\#}}{\beta},$$

which holds if and only if

$$A^* = A^{\#}. \tag{34}$$

Consider the impulse response of p_{t+2}:

$$\frac{\partial E_t\left[p_{t+2}\right]}{\partial m_t} = \mathcal{P}_2 = \frac{(\beta+1+A)}{\beta}\mathcal{P}_1 + \left(\frac{BC}{\beta^2} - \frac{1}{\beta}\right)\mathcal{P}_0 - \frac{A}{\beta}\mathcal{M}_1 - \frac{BC}{\beta^2}.$$

It then follows that

$$\left(\frac{B^*C^*}{\beta^2} - \frac{1}{\beta}\right)\mathcal{P}_0 - \frac{B^*C^*}{\beta^2} = \left(\frac{B^\#C^\#}{\beta^2} - \frac{1}{\beta}\right)\mathcal{P}_0 - \frac{B^\#C^\#}{\beta^2},$$

which holds if and only if

$$B^*C^* = B^\#C^\#. \tag{35}$$

Finally, consider the impulse response of p_{t+3}:

$$\begin{aligned}\frac{\partial E_t\left[p_{t+3}\right]}{\partial m_t} &= \mathcal{P}_3 = \frac{(\beta+1+A)}{\beta}\mathcal{P}_2 + \left(\frac{BC}{\beta^2} - \frac{1}{\beta}\right)\mathcal{P}_1 + \frac{(\beta+1+D)\,BC}{\beta^3}\mathcal{P}_0 \\ &\quad -\frac{A}{\beta}\mathcal{M}_2 - \frac{BC}{\beta^2}\mathcal{M}_1 - \frac{(\beta+1+D)\,BC}{\beta^3}.\end{aligned}$$

This implies

$$\frac{(\beta+1+D^*)\,B^*C^*}{\beta^3}\mathcal{P}_0 - \frac{(\beta+1+D^*)\,B^*C^*}{\beta^3} = \frac{\left(\beta+1+D^\#\right)B^\#C^\#}{\beta^3}\mathcal{P}_0 - \frac{\left(\beta+1+D^\#\right)B^\#C^\#}{\beta^3}.$$

It is easy to verify that this last equality holds if and only if

$$D^* = D^\#. \tag{36}$$

Finally notice that (34), (35) and (36) hold if and only if $\zeta^* = \zeta^\#$, $\omega_1^* = \omega_1^\#$ and $\omega_2^* = \omega_2^\#$. In other words, the model produces the same dynamics of the aggregate price level if and only if $\zeta^* = \zeta^\#$, $\omega_1^* = \omega_1^\#$ and $\omega_2^* = \omega_2^\#$. Therefore, we conclude that the parameters $\{\zeta, \omega_1, \omega_2\}$ are identified from the impulse response function of the aggregate price level to shocks (and hence they are identified by the likelihood implied by the model).

D A Random Taylor model with persistent durations of price spells

We can model the random draws for the duration of price spells in our Random Taylor model using a Markov chain with sector-specific transition matrices:

$$Q_n = \begin{pmatrix} q_n(1,1) & ... & q_n(1,K) \\ \vdots & \ddots & \vdots \\ q_n(K,1) & ... & q_n(K,K) \end{pmatrix},$$

where

$$\begin{aligned} q_n\left(k,k'\right) &= \Pr\left(duration\ new\ spell = k' | duration\ old\ spell = k\right), \\ \text{with } q_n\left(k,k'\right) &\in [0,1] \text{ and } \sum_{k'=1}^{K} q_n\left(k,k'\right) = 1. \end{aligned}$$

In other words, $q_n(k,k')$ is the probability that a firm in sector n that has fixed its price for k periods up to time $t-1$ will fix its new price for the next k' periods starting from time t.

The stationarity assumption will impose some restrictions on the transition matrices Q_n. At time t and in sector n, the mass of price-changing firms that had a price spell of duration k (up to $t-1$) and is about to set a new price is $\frac{\varphi_n \omega_{n,k}}{k}$. To preserve the stationarity, it must be the case that:

$$\left(\begin{array}{ccccccc} \frac{\omega_{n,1}}{1} & ... & \frac{\omega_{n,k}}{k} & ... & \frac{\omega_{n,K}}{K} \end{array} \right) Q_n = \left(\begin{array}{ccccccc} \frac{\omega_{n,1}}{1} & ... & \frac{\omega_{n,k}}{k} & ... & \frac{\omega_{n,K}}{K} \end{array} \right).$$

Although this restriction imposes some structure on how firms switch among various stickiness groups, the condition above is still fairly flexible – there might be multiple transition matrices that satisfy the condition for a given $(\omega_{n,1}, ..., \omega_{n,K})$. Hence, this model will have greater flexibility to match price duration data, such as estimates of the hazard of price adjustment. Once again, this shows how little informative about aggregate dynamics the usual price-setting statistics derived from micro price data can be.

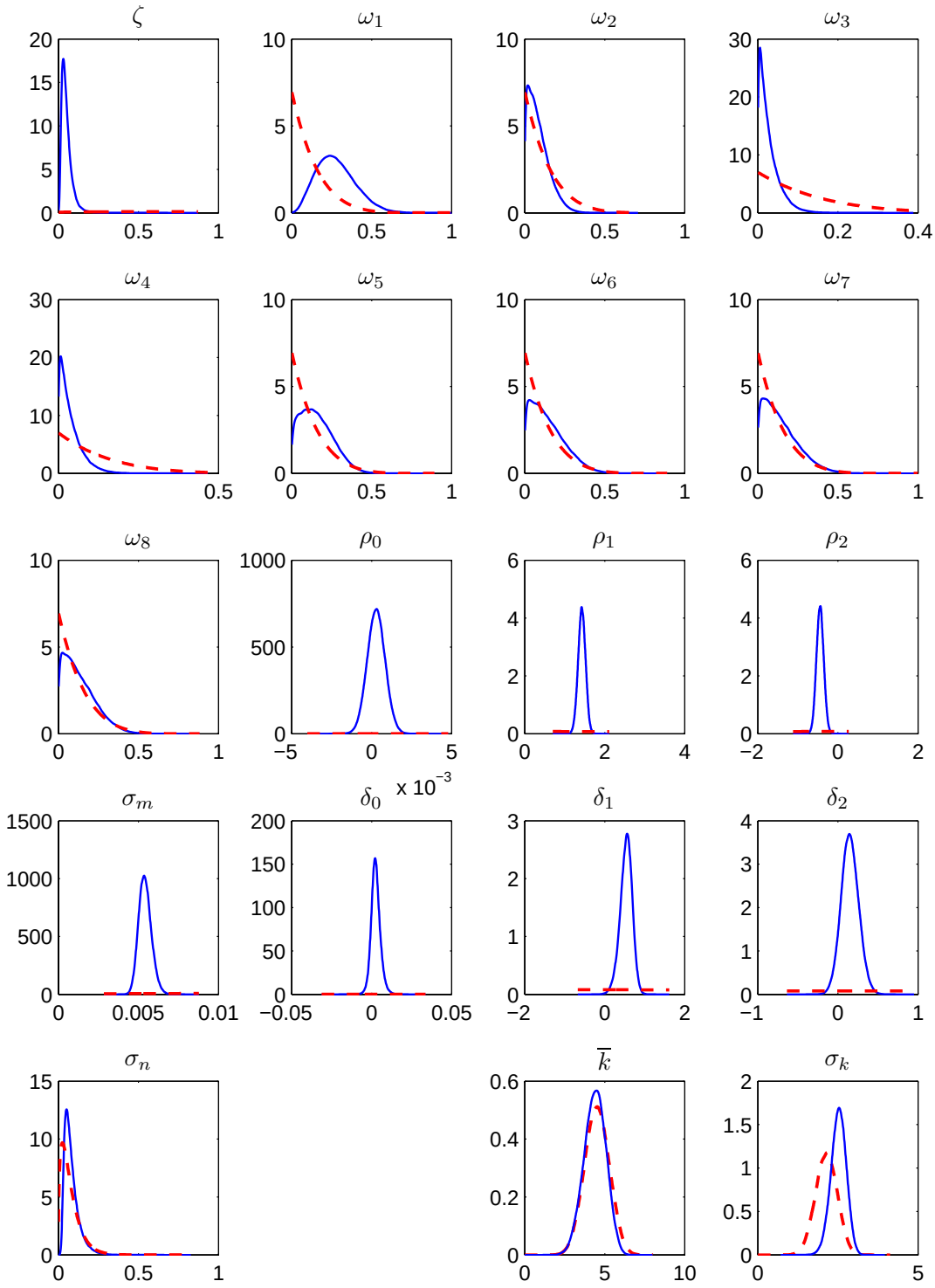

Figure 1: Marginal prior (dashed line) and posterior (solid line) distributions, flat prior

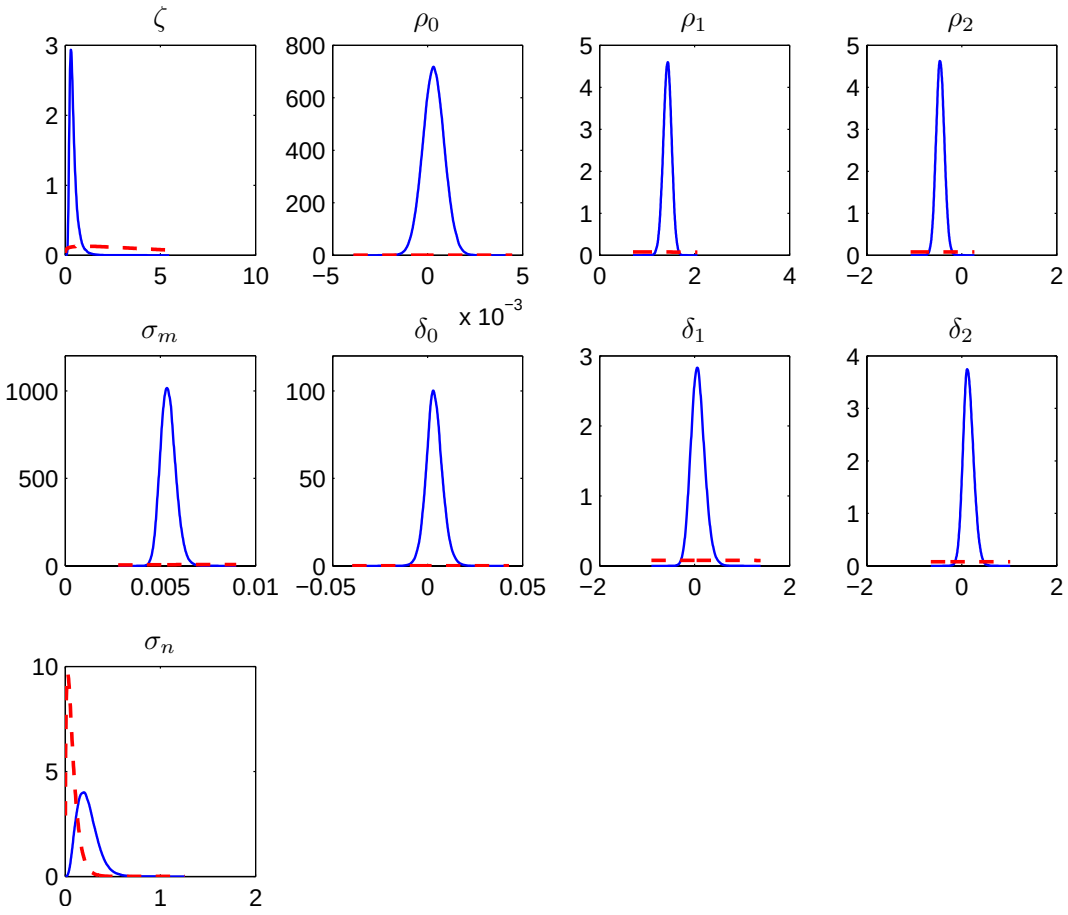

Figure 2: Marginal prior (dashed line) and posterior (solid line) distributions, one-sector model with 7-quarter price spells

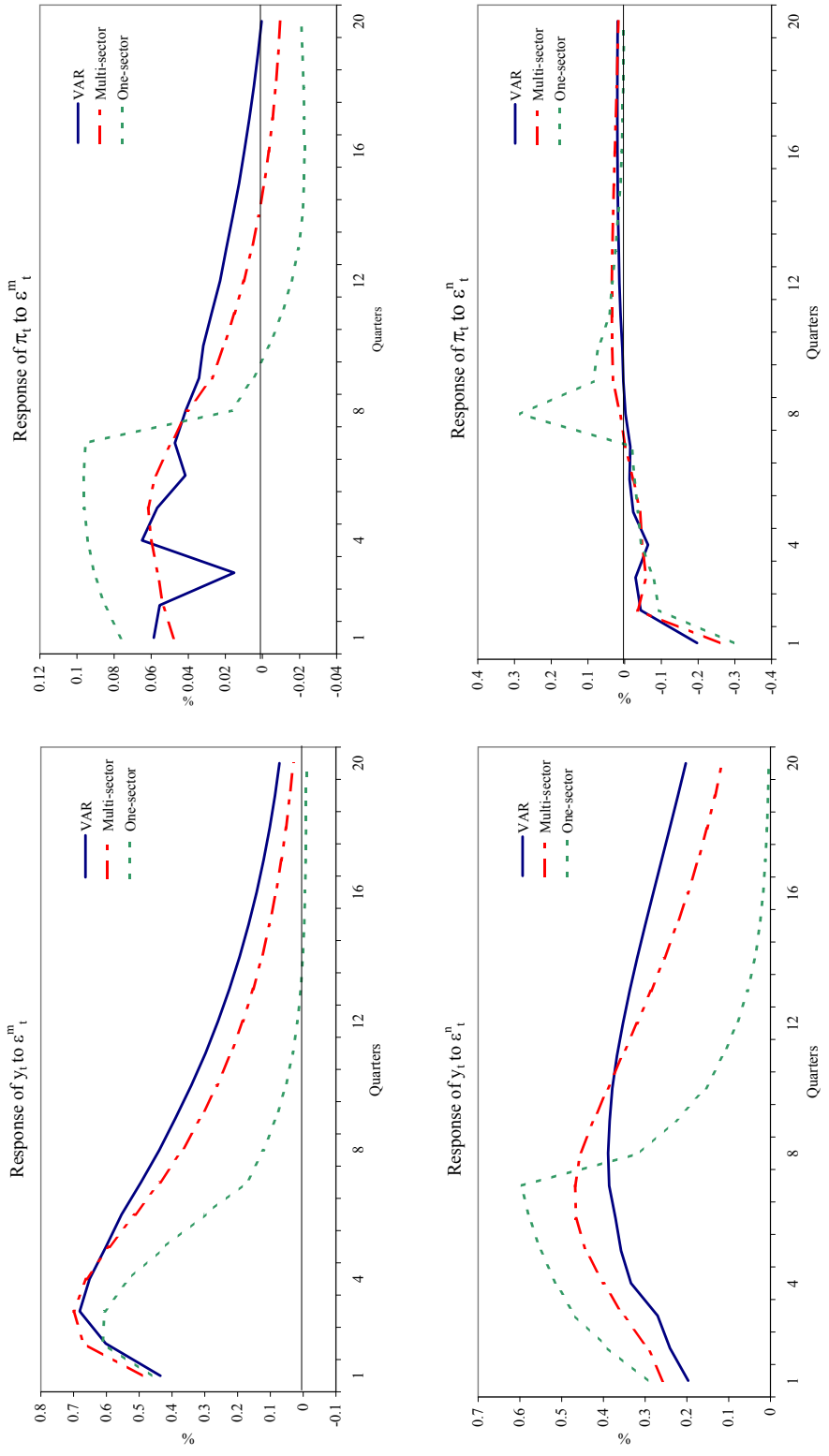

Figure 3: Impulse response functions of models and bivariate

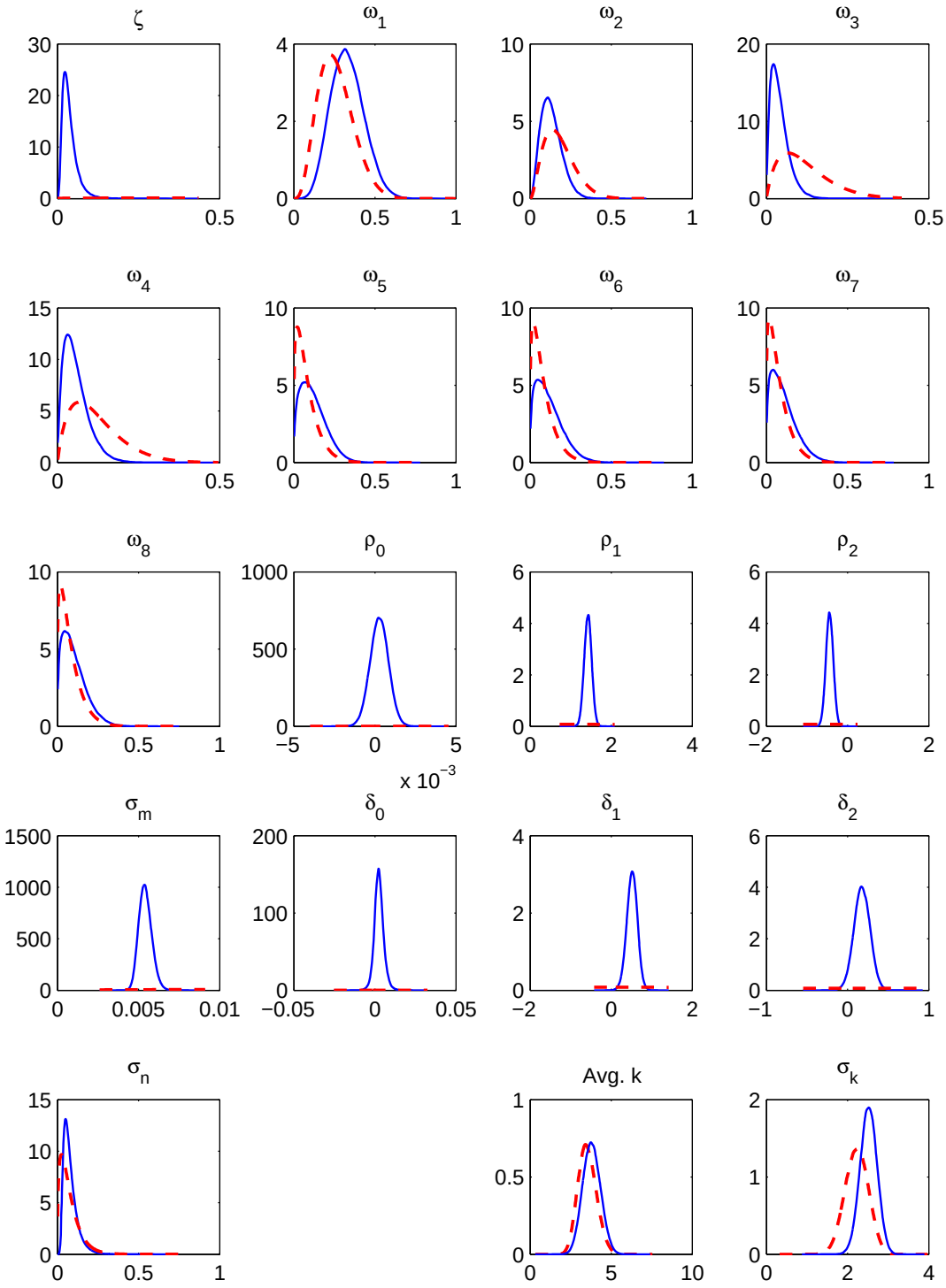

Figure 4: Marginal prior (dashed line) and posterior (solid line) distributions, prior with sales, prior informativeness = 16

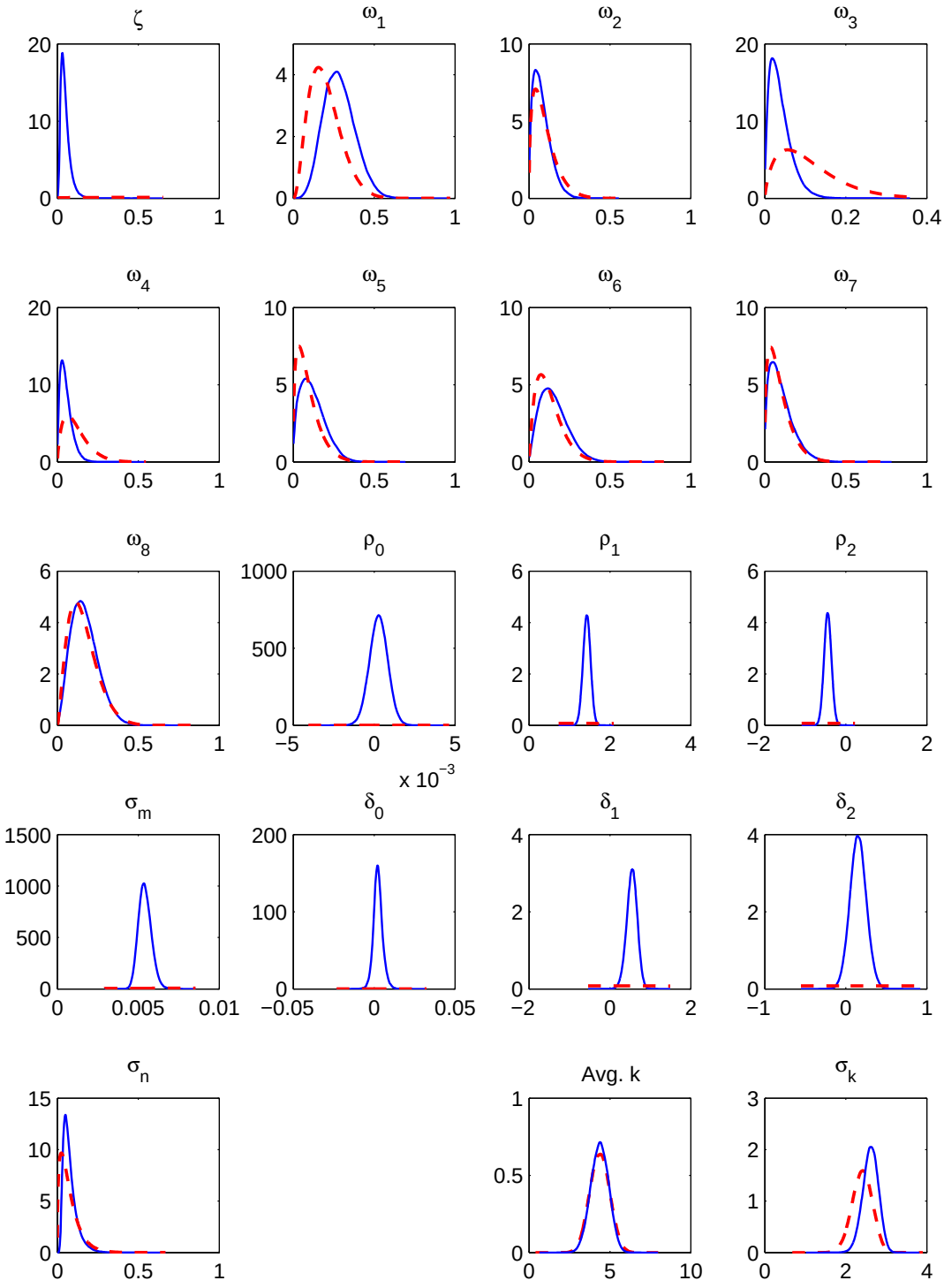

Figure 5: Marginal prior (dashed line) and posterior (solid line) distributions, prior without sales, prior informativeness = 16

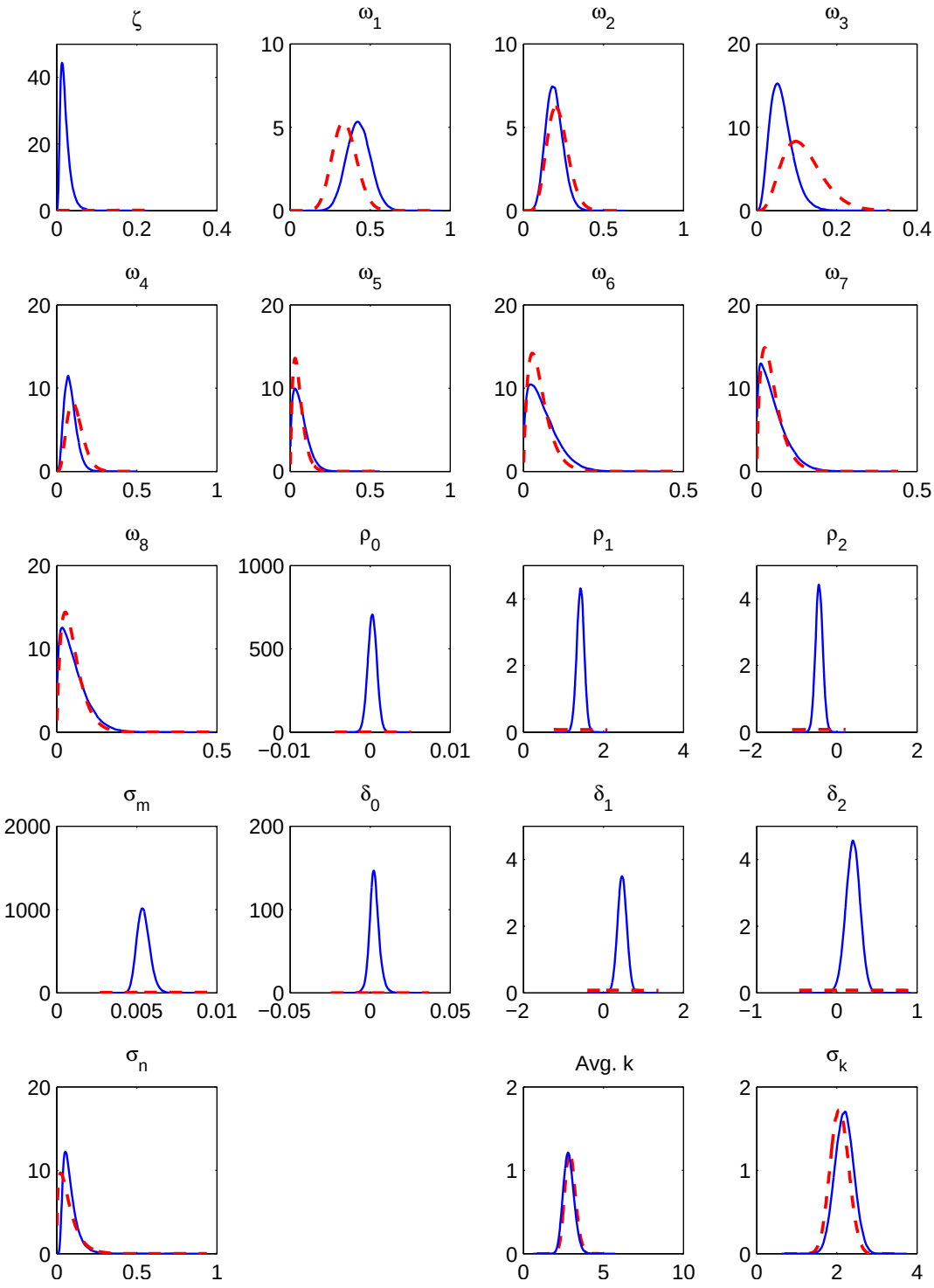

Figure 6: Marginal prior (dashed line) and posterior (solid line) distributions, prior with sales, prior informativeness = 40

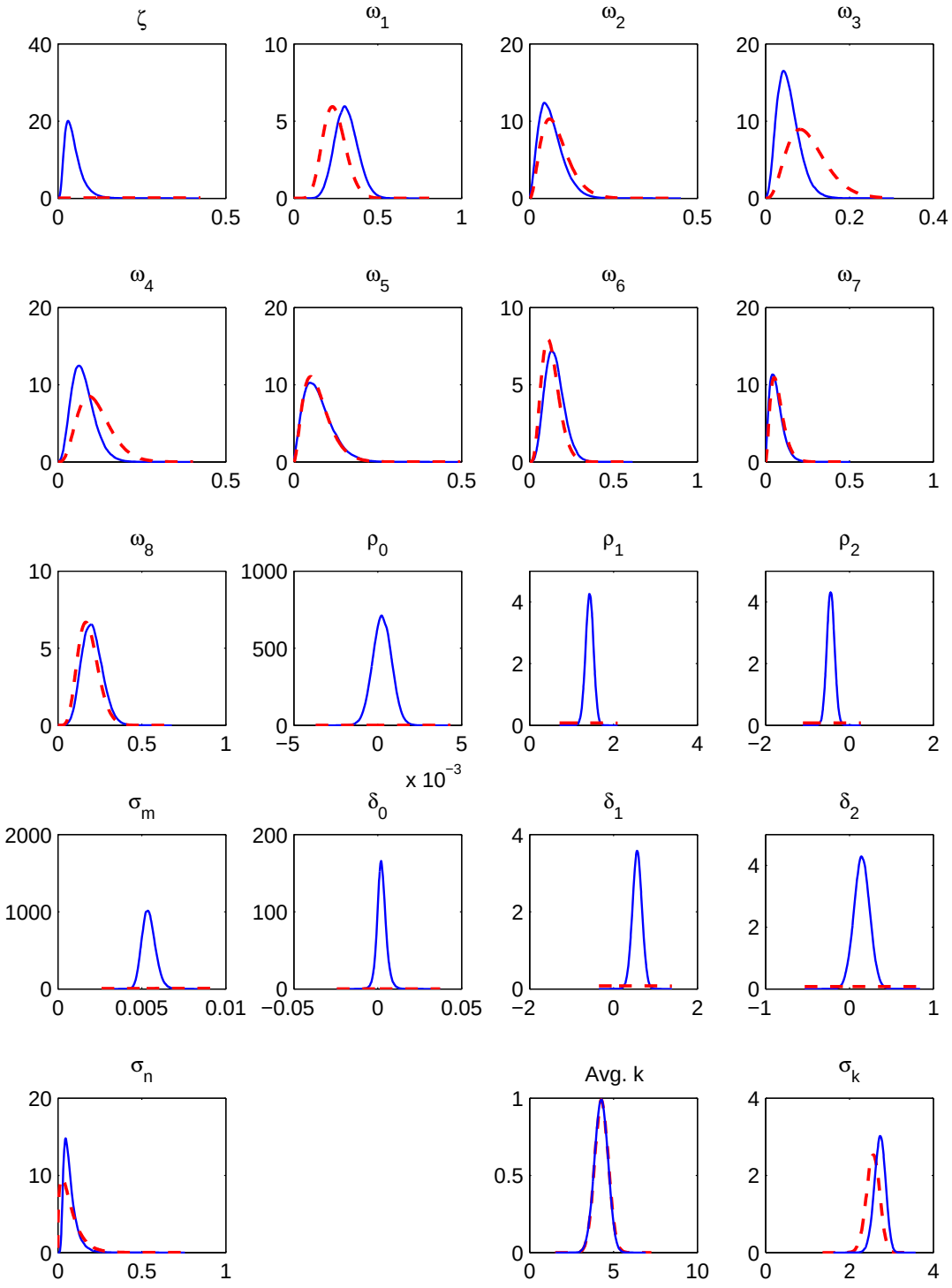

Figure 7: Marginal prior (dashed line) and posterior (solid line) distributions, prior without sales, prior informativeness = 40

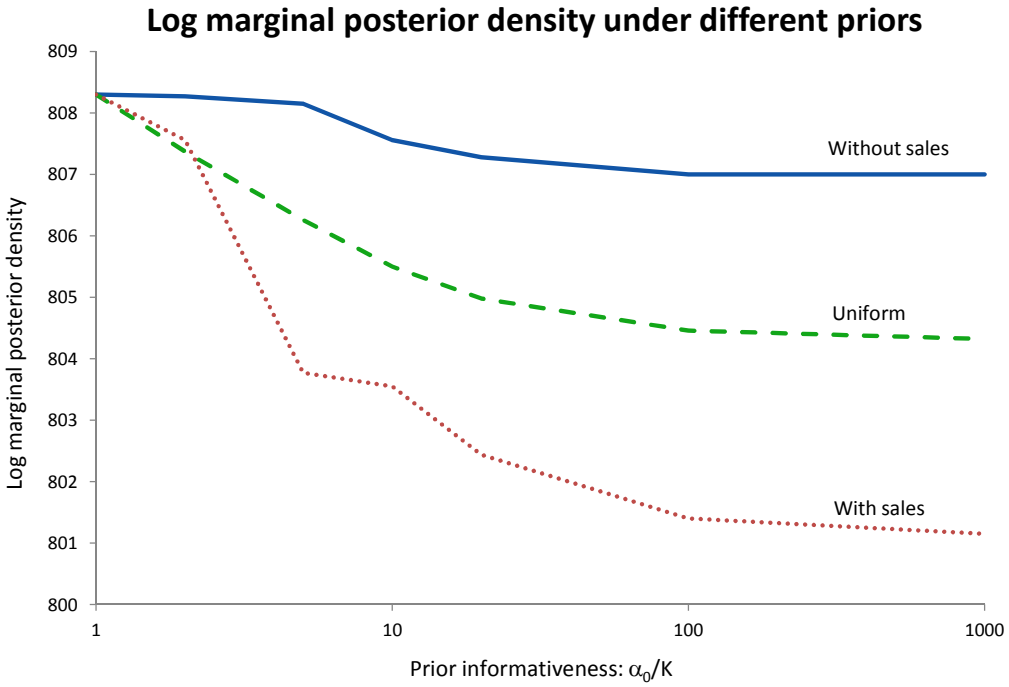

Figure 8: Log posterior marginal density of various models as a function of prior informativeness

Departamento de Economia PUC-Rio
Pontifícia Universidade Católica do Rio de Janeiro
Rua Marques de Sâo Vicente 225 - Rio de Janeiro 22453-900, RJ
Tel.(21) 31141078 Fax (21) 31141084
www.econ.puc-rio.br
flavia@econ.puc-rio.br