Federal Reserve Bank of New York
Staff Reports

The Cross-Sectional Distribution of Price Stickiness Implied by Aggregate Data

Carlos Carvalho
Niels Arne Dam

Staff Report no. 419
December 2009
Revised August 2010

The Cross-Sectional Distribution of Price Stickiness Implied by Aggregate Data
Carlos Carvalho and Niels Arne Dam
Federal Reserve Bank of New York Staff Reports, no. 419
December 2009; Revised August 2010
JEL classification: E10, E30

Abstract

Using only aggregate data as observables, we estimate multisector sticky-price models for twelve countries, allowing the degree of price stickiness to vary across sectors. We use a specification that allows us to extract information about the underlying cross-sectional distribution from aggregate data. Identification is possible because sectors play different roles in determining the response of aggregate variables to shocks at different frequencies: sectors where prices are more sticky are relatively more important in determining the low-frequency response. We find that the inferred distributions of price stickiness conform quite well with empirical distributions constructed from the available microeconomic evidence on price setting. We then explore our Bayesian approach to combine the aggregate time-series data with the microeconomic information on the distributions of price rigidity, and re-estimate the models for the United States, Denmark, and Japan. Our results show that allowing for this type of heterogeneity is critically important to understanding the joint dynamics of output and prices, and it constitutes a step toward reconciling the extent of nominal price rigidity implied by aggregate data with the evidence from price micro data.

Key words: heterogeneity, price stickiness, micro data, macro data, Bayesian estimation

---

Carvalho: Federal Reserve Bank of New York (e-mail: carlos.carvalho@ny.frb.org). Dam: Danmarks Nationalbank (e-mail: nad@nationalbanken.dk). This paper was previously distributed under the title "Estimating the Cross-Sectional Distribution of Price Stickiness from Aggregate Data." For helpful comments and suggestions, the authors thank Luis Alvarez, Marco Del Negro, Stefano Eusepi, Henrik Jensen, Ed Knotek, Oleksyi Kryvtsov, John Leahy, Frank Schorfheide, and seminar participants at the American Economic Association 2010 Meeting, the Banque de France conference "Understanding Price Dynamics: Recent Advances," the 2009 Econometric Society European Meeting, the 2009 Society for Economic Dynamics Meeting, the 2008 North American Meeting of the Econometric Society, the Bank of Canada, the European Central Bank – Wage Dynamics Network Workshop, Rutgers University, the 2008 Latin American Meeting of the Econometric Society, Danmarks Nationalbank, the University of Copenhagen, Sveriges Riksbank, and the Federal Reserve Bank of New York.

1 Introduction

Comparisons of estimates of important economic parameters based on microeconomic or disaggregated data versus those based on aggregate data often produce a conflicting picture. Perhaps the prime example involves estimates of the elasticity of individual labor supply, which are typically smaller than estimates of the elasticity of the aggregate labor supply.[1] Other examples are the elasticity of substitution between domestic and foreign goods - which is substantially higher when estimated from disaggregated data, and the parameters determining the extent of habit formation in consumption - which may differ significantly depending on whether microeconomic or aggregate data are employed in the estimation.[2] These discrepancies are usually associated with a large amount of heterogeneity in estimates of the parameters based on disaggregated data, which contrasts with the (explicit or implicit) homogeneity assumption that underlies most of the estimates based on aggregate data. These differences have fostered debates about how to calibrate models with representative agents, and also the development of heterogeneous-agents models that can reconcile the empirical findings.[3]

Estimates of the extent of nominal price rigidity reveal the same kind of tension. A recent empirical literature that uses price micro data from various sources documents that, on average, prices change at least once a year (for a recent survey, see Klenow and Malin 2009). In contrast, estimates of the frequency of price changes from dynamic stochastic general equilibrium (DSGE) models and aggregate data imply much less frequent adjustment.[4] Likewise, evidence on the response of the aggregate price level to various shocks in vector autoregressions (VARs) points to sluggish adjustment. This discrepancy between micro- and macro-based evidence on price stickiness has shaped many of the developments in the field of Monetary Economics since the emergence of so-called new Keynesian Economics in the 1980s.

In this paper we take a step towards reconciling the apparent disconnect between micro- and macro-based estimates of nominal price rigidity. Our approach is motivated by the microeconomic evidence of substantial heterogeneity in the degree of price rigidity across sectors emphasized by Bils and Klenow (2004) and subsequent papers, and by recent work showing that such heterogeneity matters for aggregate dynamics (Carvalho 2006, Nakamura and Steinsson 2010, and others).

We employ a Bayesian approach to estimate sticky-price models for twelve countries allowing for

[1] For a recent discussion of some of these contrasting empirical findings, see Shimer (2009).
[2] See, respectively, Imbs and Mejean (2009) and Ravina (2007).
[3] See, for example, Browning et al. (1999) and Chang and Kim (2006).
[4] See, for example, the recent survey by Maćkowiak and Smets (2008).

sectoral heterogeneity in price stickiness, and using only aggregate data on nominal and real Gross Domestic Product (GDP) as observables. We start by asking whether these data provide support for this type of heterogeneity. We can address this question because we use models that allow us to extract information about the underlying cross-sectional distribution from aggregate data. Identification is made possible by the fact that in the models different sectors are relatively more important in determining the response of aggregate variables 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-price sectors. To assess the empirical support for heterogeneity we also estimate versions of the model that impose the same degree of price rigidity for all firms in the economy.

The results discriminate sharply between the models with heterogeneity in price stickiness and their homogeneous-firms counterparts, and provide strong support for this type of heterogeneity. In a comparison with the *best-fitting* homogeneous model, the posterior odds in favor of the heterogeneous models is of the order of $10^{11} : 1$. Moreover, the homogeneous specification favored by the data implies an extent of price rigidity that, at 7 quarters, seems unrealistic in light of the microeconomic evidence.

With our macro-based estimates in hand, we then ask whether the implied cross-sectional distribution of price rigidity is consistent with the available microeconomic evidence for the various countries. We find that it is. In particular, for countries with detailed enough cross-sectional statistics on price rigidity (U.S., Denmark and Japan) the macro-based estimates often capture nuanced features of the empirical distributions. The positive answers to our two questions provide empirical support for heterogeneity in price stickiness as an important step toward reconciling the extent of nominal price rigidity implied by aggregate data with the evidence from price micro data.

Similarities notwithstanding, there are differences between our macro-based estimates of price rigidity and the available microeconomic evidence. Thus, a legitimate question is why should we care about macro-based estimates if the microeconomic evidence usually relies on direct observation of millions of individual price paths - and is thus model-free? One obvious answer is hinted at by our estimation of the model for countries with limited availability of microeconomic evidence - or lack thereof in some cases: if our macro-based estimates line up reasonably well with the microeconomic evidence when the latter is available, perhaps we can rely on them in the absence of that evidence. The second answer is actually related to the model-free nature of the microeconomic evidence. It is possible that some price adjustments do not convey information about changes in macroeconomic conditions, while others do.[5] In that case, macro-based estimates should convey useful information about the price changes that do matter for aggregate dynamics.

Taking advantage of our Bayesian methodology, we argue that a promising approach to integrating our micro and macro views on price setting is to combine macroeconomic observables with microeconomic information contained in empirical cross-sectional distributions of price rigidity, in the context of models with explicit heterogeneity along this dimension. We parameterize our prior over the set of parameters that characterize the cross-sectional distribution in the model in a way that easily allows us to relate its moments to an empirical distribution of price rigidity. We then reestimate the models for the U.S., Denmark, and Japan, still using only aggregate data on nominal and real output as observables, and discuss how the results change relative to our purely macro-based estimates.

Until recently, most of the efforts devoted to reconciling micro- and macro-based evidence on price rigidity focused on mechanisms that can slow down the adjustment of the aggregate price level to shocks, in the context of models in which all prices are equally sticky. These mechanisms are usually referred to as "real rigidities" (Ball and Romer 1990). As long as price changes are not perfectly synchronized, strong enough real rigidities make it optimal for firms to undertake a series of partial adjustments to shocks that affect the aggregate price level, rather than a single adjustment. This delayed response is due to an interdependence in pricing decisions that is often referred to as a "strategic complementarity" in price setting.

For a given frequency of price adjustments, a model with heterogeneity in price stickiness can produce richer - in particular, more sluggish - aggregate dynamics than an otherwise identical homogeneous-firms model with the same frequency of price changes, even in the absence of pricing complementarities (Carvalho 2006, Carvalho and Schwartzman 2008, Nakamura and Steinsson 2010). Yet, the two mechanisms can, and do coexist in our estimated model - which implies that pricing decisions are strategic complements. Coupled with heterogeneity in price rigidity, this complementarity leads firms in the more sticky sectors to become disproportionately important (relative to their sectoral weight) in shaping aggregate dynamics, through their influence on pricing decisions of firms that change prices more frequently.[6]

[5] This possibility is at the core of the debate about whether to exclude price changes due to sales when computing measures of price rigidity (see Nakamura and Steinsson 2008 and Klenow and Kryvtsov 2008 for empirical results, and Kehoe and Midrigan 2008 and Guimaraes and Sheedy 2009 for models that incorporate temporary price changes).

[6] This potential amplification mechanism arising from the interaction between heterogeneity in price rigidity and pricing complementarities is discussed in detail in Carvalho (2006) and Carvalho and Schwartzman (2008). Nakamura and Steinsson (2010) conclude that this interaction is not important in their calibrated menu-cost model.

While we focus on the microeconomic evidence on heterogeneity in price rigidity, the recent literature has established a few additional empirical regularities regarding price setting in the U.S. and other economies. After assessing how our estimated models fare in light of the evidence on heterogeneity, we discuss their ability to match additional micro price facts (at least qualitatively). We conclude that such models can match four out of the eight facts documented by Klenow and Kryvtsov (2008). In addition, we provide details on how to change the models to match an additional three of their micro facts without any changes to their aggregate implications.

Our work is related to the recent 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 as an amplification mechanism for "monetary non-neutralities" in calibrated models (e.g., Carvalho 2006, Nakamura and Steinsson 2010, Carvalho and Nechio 2010).[7] These papers do not address the question of whether such heterogeneity does in fact help existing sticky-price models fit the data better according to formal statistical criteria. We do that by estimating our multi-sector model, exploring the model-implied mapping between the distribution of price stickiness and the dynamics of macroeconomic observables. As part of this process, we provide estimates of the distribution of price rigidity based on macroeconomic data.

In terms of empirical work on the importance of heterogeneity in price stickiness, Imbs et al. (2007) 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. (2009a) 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.[8] 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 (de facto) assume ex-post rather than ex-ante heterogeneity in nominal rigidities.[9] 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

[7] For brevity, and given our empirical focus, we refer the reader to Carvalho and Schwartzman (2008) for references to theoretical contributions and/or quantitative work based on calibrated models.

[8] Bouakez et al. (2009b) find similar results in an extension of their earlier paper to a larger number of sectors.

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

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).[10]

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 Jadersic'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. We also provide macro-based estimates for twelve countries other than the U.S. 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.

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. Section 3 describes our empirical methodology and data. In Section 4 we present our macro-based estimates of multi-sector models for twelve countries, and perform model comparison with specifications that impose the same degree of price rigidity for all firms. In Section 5 we combine microeconomic information and macroeconomic data in the estimation. Section 6 discusses the performance of the model in light of additional micro price facts. We discuss robustness issues and directions for future research in Section 7, before concluding.

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

[10] 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).

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 $C(Y_t(k,j), Y_t, \xi_t)$ 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.[11]

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

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

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.

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

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

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

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 $X_t(k)$.[12] 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:

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

The structure of the supply side of our model is a multi-sector economy with Taylor (1979, 1980) staggered price setting, in which the extent of price rigidity varies across different sectors. Instead of postulating a fully specified economy 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".[13] 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 are less dependent on details of the demand side of the model, which is not the focus of our analysis.[14]

[12] In Section 6 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.

[13] Several earlier papers combine structural equations with empirical specifications for other parts of the model. Sbordone (2002), Guerrieri (2006) and Coenen et al. (2007) are recent examples.

[14] Needless to say, the results are conditional on the particular model of price setting that we adopt. In Section 7 we discuss the extent to which our conclusions may generalize to alternative price-setting specifications.

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 obtain:[15]

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

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 technological shocks. Here we do not pursue a structural interpretation of the exogenous processes driving the economy.[16] 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:[17]

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

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

Large real rigidities (small ζ in equation (2)) reduce the sensitivity of prices to aggregate demand

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

[16] 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).

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

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, whether there is an explicit input-output structure etc.[18] In the context of our model, ζ is itself a primitive parameter. Following standard practice in the literature, we refer to economies with $\zeta < 1$ as ones displaying *strategic complementarities* in price setting. To clarify the meaning of this expression, replace (3) in (2) to obtain:

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

That is, new prices are set as a discounted weighted average of current and expected future driving variables $\left(m_{t+i} - y_{t+i}^n\right)$ 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_t^n) output

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

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

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

$$y_t^n = \delta_0 + \delta_1 y_{t-1}^n + ... + \delta_{p2} y_{t-p_2}^n + \varepsilon_t^n, \tag{8}$$

where $\varepsilon_t = (\varepsilon_t^m, \varepsilon_t^n)$ is i.i.d. $N\left(0_{1\times 2}, \Omega^2\right)$, with $\Omega^2 = \begin{bmatrix} \sigma_m^2 & 0 \\ 0 & \sigma_n^2 \end{bmatrix}$.

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

2.3 State-space representation and likelihood function

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

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

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 $\{Z_\tau^*\}_{\tau=1}^{t-1}$, 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 only use 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 can 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:

$$\begin{aligned} m_t &= \rho_1 m_{t-1} + \varepsilon_t^m, \text{ and} && (10) \\ y_t^n &= \delta_1 y_{t-1}^n + \varepsilon_t^n. && (11) \end{aligned}$$

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.[19]

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. This illustrates why the straightforward recursive identification that

[19] 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.

applies when $\zeta = 1$ no longer works.

It turns out that, in equilibrium, pricing interactions manifest themselves through a dependence of the aggregate price level on its own lags. This is how pricing interactions serve as a propagation mechanism. Specifically, the expression for the equilibrium price level becomes:

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

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.[20]

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 use to us if the mechanism highlighted above does not work well in practice, especially 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 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. 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,

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

in this case the estimates are slightly biased and less precise. This finding underscores our case for incorporating prior information from the microeconomic evidence on price-setting, as we do in Section 5.

3 Empirical methodology and data

With the challenges involved in bridging the gap between microeconomic information on firms' pricing behavior and time series of aggregate nominal and real output, the choice of empirical methodology is of critical importance. We employ a Bayesian approach as this allows us to eventually integrate microeconomic information on the distribution of price rigidity with those macroeconomic time series. With some abuse of notation, the Bayesian principle can be shortly stated as:

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

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

We use our sources of information in two ways. First, in Section 4, we estimate the cross-sectional distribution of price stickiness for twelve countries using time-series data on aggregate nominal and real output under a prior distribution that is "flat" (uninformative) over the domain of the vector of sectoral weights. The macro-based estimates are then compared with available microeconomic evidence on price rigidity. The countries are: Australia, Canada, Denmark, France, Italy, Japan, Korea, Norway, Sweden, Switzerland, the U.K., and the U.S.

In a second estimation, in Section 5, for the three countries with a comprehensive cross-section of micro-based statistics on price rigidity (U.S., Denmark, and Japan), we combine prior information from the microeconomic data with the information obtained from the likelihood function jointly implied by the model and macroeconomic observables. 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, $\omega \sim D(\alpha_1, ..., \alpha_K)$ with density function:

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

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.[21] Thus, for given $\alpha_1, ..., \alpha_K$, $\alpha_0 \equiv \sum_k \alpha_k$ 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 α_k/α_0, whereas its mode equals $(\alpha_0 - K)^{-1}(\alpha_k - 1)$ (provided that $\alpha_i > 1$ for all i).

As mentioned previously, our first estimation does not make use of the microeconomic information. It imposes a "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 case allows us to extract the information that the aggregate data contain about the cross-sectional distribution of price stickiness. Subsequently, when we incorporate the microeconomic information in the estimation, we relate the relative sizes of the hyperparameters $\alpha_1, ..., \alpha_K$ to the empirical sectoral weights for each of the U.S., Denmark, and Japan, and choose the value $\alpha_0 > K$ to determine the tightness of the prior distribution around the empirical distribution. We leave the details of how we construct the informative priors to Section 5.

3.2 Priors on remaining parameters

The remaining model parameters 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. $p_1 = p_2 = 2$.[22] The second set of parameters consists of the standard deviations

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

[22] 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

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.[23]

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
σ_n, σ_m	Gamma$(1.5, 20)$	0.025	0.075	0.06
ρ_j, δ_j	$\mathrm{N}(0, 5^2)$	0.00	0.00	5.00

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. These are measured as seasonally-adjusted GDP at, respectively, current and constant prices. We take natural logarithms and remove a linear trend from the data. The sample period is driven by the choice for the U.S. economy. 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.[24] Data for the U.S. are from the Bureau of Economic Analysis. Estimations for the other eleven countries use the analogous data taken from the OECD database (http://stats.oecd.org).[25] Data for Sweden are only available since 1980, and data for Italy are available since 1981. In these cases we shorten the pre-sample and start the actual

specification with fixed number of lags. Our conclusions are robust to alternative assumptions about the number of lags (see Section 7).

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

[24] 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.

[25] With the exception of Denmark and the U.S., the countries were chosen precisely on the basis of availability of these data in the OECD database. In the case of Denmark, the quarterly data starting in 1979 were provided to us by Danmarks Nationalbank.

sample in 1983Q1, as for the U.S. economy.

3.4 Empirical distributions of price stickiness

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 that change prices every other quarter, and so on.

For each country, our goal is to map the available statistics on the frequency of price changes into an empirical distribution of sectoral weights $\widehat{\omega}$. We aggregate the items for which the statistics are reported (usually categories of goods and services) so that the ones which have an average implied duration of price spells between zero and one quarter (inclusive) are assigned to the first sector; the ones with an average duration between one (exclusive) and two quarters (inclusive) are assigned to the second sector, and so on. The sectoral weights are aggregated accordingly by adding up the corresponding expenditure weights associated with the items, which are usually reported along with the pricing statistics. We proceed in this fashion until the sector with 7-quarter price spells. Finally, we aggregate all the remaining items, which have average implied durations of price rigidity that exceed 7 quarters, into the sector with 2-year price spells. This gives rise to the empirical cross-sectional distributions of price stickiness that we use to either assess the results obtained under flat priors, or as information to be incorporated in the estimation. For each country we also compute the average duration of price spells, $\widehat{\overline{k}} = \sum_{k=1}^{8} \widehat{\omega}_k k$, and the standard deviation of the underlying distribution, $\widehat{\sigma}_k = \sqrt{\sum_{k=1}^{8} \widehat{\omega}_k \left(k - \widehat{\overline{k}}\right)^2}$.

For the U.S., we extract our microeconomic information about the cross-sectional distribution of price stickiness from Nakamura and Steinsson (2008). Following the seminal work of Bils and Klenow (2004), they analyze the frequency of price changes in the U.S. economy using quite disaggregated datasets from the Bureau of Labor Statistics, which underlie the construction of price indices. We work with the statistics on the frequency of regular price changes (i.e. excluding those associated with sales and product substitutions) that they report for 272 categories of goods and services contained in the Consumer Price Index.

Among the other eleven countries, we are aware of studies reporting price-setting statistics based on micro data for Canada (Harchaoui et al. 2008), Denmark (Hansen and Hansen 2006), France (Baudry et al. 2007), Italy (Fabiani et al. 2006), Japan (Higo and Saita 2007), Norway (Wulfsberg

2009), Switzerland (Kaufmann 2008), and the U.K. (Bunn and Ellis 2009). With the exception of Canada, we have access to the cross-sectional price-setting statistics for these countries.[26]

Unfortunately, Denmark and Japan are the only cases where the level of cross-sectional detail is comparable to the studies for the U.S. economy. For Denmark, Hansen and Hansen (2006) include statistics for 391 goods and services categories at the COICOP 5-digit level. For Japan, Higo and Saita (2007) provide statistics for 513 goods and services categories at a very disaggregated level. The papers for Switzerland, France and Italy also provide statistics at the COICOP 5-digit level, but for 139, 136 and 48 categories, respectively. Wulfsberg (2009) provides statistics for Norway for 89 4-digit COICOP classes. Finally, Bunn and Ellis (2009) provide much more aggregated statistics for the U.K..

The empirical distributions for the U.S., Denmark and Japan are summarized in Table 2. For the countries with coarser, but still minimally detailed microeconomic information we focus on the average duration of price spells and cross-sectional standard deviation of the sectoral distribution of price spells (we report these statistics in Table 6, Subsection 4.2).

Table 2: Empirical distributions of price stickiness

	$\widehat{\omega}_1$	$\widehat{\omega}_2$	$\widehat{\omega}_3$	$\widehat{\omega}_4$	$\widehat{\omega}_5$	$\widehat{\omega}_6$	$\widehat{\omega}_7$	$\widehat{\omega}_8$	$\widehat{\overline{k}}^{(*)}$	$\widehat{\sigma}_k^{(*)}$
U.S.	0.273	0.071	0.098	0.110	0.059	0.129	0.061	0.198	4.25	2.66
Denmark	0.136	0.179	0.045	0.091	0.124	0.067	0.108	0.250	4.77	2.60
Japan	0.311	0.204	0.043	0.045	0.070	0.046	0.039	0.244	3.87	2.85

(*) 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.[27] 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

[26] We thank Philip Bunn, Silvia Fabiani, Niels Lynggård Hansen, Daniel Kaufmann, Herve Le Bihan, Yumi Saita, and Fredrik Wulfsberg for providing us with the statistics from their respective papers.

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

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, 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 Macro-based estimates

4.1 U.S., Denmark, and Japan

Tables 3-5 and Figures 1-3 report the results for, respectively, the U.S., Denmark, and Japan, in terms of marginal distributions for the parameters.[28] The empirical distributions of price rigidity from Table 2 are reproduced in the last column of each table for ease of comparison.

The distributions that we infer from aggregate data conform quite well with the empirical ones. In what follows, we use the posterior means as the point estimates for the sectoral weights, reported in the third column of each table.[29] The macro-based estimates for the U.S. 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 (cross-sectional) distribution of price spells is approximately 8 months. These numbers are quite close to the empirical distribution. The correlation between estimated and empirical sectoral weights is 0.62. The index of real rigidities implies strong pricing complementarities. The posterior mean of ζ is 0.05 and the 95$^{\text{th}}$ 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 in Carvalho (2006), in this type of model such complementarities interact with heterogeneity in price stickiness to amplify the aggregate effects of nominal rigidities.

[28] 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.

[29] 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 for the U.S. economy - flat prior

	Model with $K = 8$, $\alpha_k = 1$ *for all* k			*Empirical distribution*
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	0.273
ω_2	0.094 (0.007;0.348)	0.072 (0.007;0.212)	0.086	0.071
ω_3	0.094 (0.007;0.348)	0.020 (0.002;0.078)	0.027	0.098
ω_4	0.094 (0.007;0.348)	0.027 (0.002;0.107)	0.037	0.110
ω_5	0.094 (0.007;0.348)	0.144 (0.017;0.337)	0.156	0.059
ω_6	0.094 (0.007;0.348)	0.123 (0.011;0.345)	0.144	0.129
ω_7	0.094 (0.007;0.348)	0.120 (0.010;0.353)	0.143	0.061
ω_8	0.094 (0.007;0.348)	0.112 (0.010;0.323)	0.132	0.198
$\bar{k}$	4.501 (3.245;5.760)	4.394 (3.214;5.462)	4.37	4.25
σ_k	2.139 (1.584;2.678)	2.523 (2.112;2.893)	2.62	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 column reproduces the empirical distribution from Table 2.

Table 4: Parameter estimates for the Danish economy - flat prior

	Model with $K = 8$, $\alpha_k = 1$ for all k			*Empirical distribution*
ζ	4.440 (0.466;16.863)	0.190 (0.060;0.586)	0.240	–
ω_1	0.094 (0.007;0.348)	0.283 (0.109;0.521)	0.295	0.136
ω_2	0.094 (0.007;0.348)	0.051 (0.004;0.170)	0.064	0.179
ω_3	0.094 (0.007;0.348)	0.034 (0.003;0.125)	0.045	0.045
ω_4	0.094 (0.007;0.348)	0.034 (0.003;0.133)	0.047	0.091
ω_5	0.094 (0.007;0.348)	0.065 (0.005;0.239)	0.086	0.124
ω_6	0.094 (0.007;0.348)	0.074 (0.006;0.264)	0.097	0.067
ω_7	0.094 (0.007;0.348)	0.176 (0.017;0.458)	0.198	0.108
ω_8	0.094 (0.007;0.348)	0.147 (0.014;0.402)	0.169	0.250
$\bar{k}$	4.501 (3.245;5.760)	4.507 (3.214;5.731)	4.58	4.77
σ_k	2.139 (1.584;2.678)	2.668 (2.203;3.044)	2.78	2.60
ρ_0	0.000 (−8.224;8.224)	0.001 (−0.001;0.003)	0.001	–
ρ_1	0.000 (−8.224;8.224)	0.774 (0.607;0.938)	0.774	–
ρ_2	0.000 (−8.224;8.224)	0.183 (0.020;0.348)	0.183	–
σ_m	0.059 (0.009;0.195)	0.013 (0.011;0.014)	0.013	–
δ_0	0.000 (−8.224;8.224)	0.001 (−0.004;0.007)	0.001	–
δ_1	0.000 (−8.224;8.224)	0.297 (0.063;0.505)	0.292	–
δ_2	0.000 (−8.224;8.224)	0.299 (0.147;0.449)	0.298	–
σ_n	0.059 (0.009;0.195)	0.065 (0.032;0.171)	0.079	–

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 column reproduces the empirical distribution from Table 2.

In the case of Denmark, 30% of firms change prices every quarter; 54% change prices at least once a year; 17% change prices once every two years. The average duration of price spells is 13.7 months, and the standard deviation of the (cross-sectional) distribution of price spells is approximately 8.3 months. With the exception of the split of weights between the first two sectors, the estimated distribution is quite close to the empirical distribution. If we construct a 7-sector distribution by combining sectors one and two, the correlation with the empirical distribution is 0.85. The same correlation for the estimated 8-sector distribution is 0.34. The index of real rigidities is noticeably higher than in the U.S., implying weaker pricing complementarities.

Table 5: Parameter estimates for the Japanese economy - flat prior

	Model with $K = 8$, $\alpha_k = 1$ for all k			*Empirical distribution*
ζ	4.440 (0.466;16.863)	0.034 (0.014;0.085)	0.040	–
ω_1	0.094 (0.007;0.348)	0.295 (0.134;0.523)	0.307	0.311
ω_2	0.094 (0.007;0.348)	0.072 (0.006;0.212)	0.085	0.204
ω_3	0.094 (0.007;0.348)	0.059 (0.005;0.177)	0.070	0.043
ω_4	0.094 (0.007;0.348)	0.048 (0.004;0.164)	0.061	0.045
ω_5	0.094 (0.007;0.348)	0.043 (0.003;0.158)	0.057	0.070
ω_6	0.094 (0.007;0.348)	0.052 (0.004;0.192)	0.069	0.046
ω_7	0.094 (0.007;0.348)	0.131 (0.012;0.363)	0.151	0.039
ω_8	0.094 (0.007;0.348)	0.187 (0.025;0.415)	0.199	0.244
$\bar{k}$	4.501 (3.245;5.760)	4.291 (3.055;5.470)	4.28	3.87
σ_k	2.139 (1.584;2.678)	2.732 (2.351;3.044)	2.82	2.85
ρ_0	0.000 (−8.224;8.224)	0.000 (−0.002;0.002)	0.000	–
ρ_1	0.000 (−8.224;8.224)	1.378 (1.193;1.580)	1.381	–
ρ_2	0.000 (−8.224;8.224)	−0.401 (−0.609;−0.213)	−0.405	–
σ_m	0.059 (0.009;0.195)	0.011 (0.010;0.012)	0.011	–
δ_0	0.000 (−8.224;8.224)	0.005 (−0.004;0.014)	0.005	–
δ_1	0.000 (−8.224;8.224)	0.485 (0.210;0.737)	0.480	–
δ_2	0.000 (−8.224;8.224)	0.183 (−0.020;0.390)	0.184	–
σ_n	0.059 (0.009;0.195)	0.146 (0.060;0.295)	0.158	–

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 column reproduces the empirical distribution from Table 2.

In the case of Japan, 31% of firms change prices every quarter; 52.3% change prices at least once a year; 20% change prices once every two years. The average duration of price spells is 12.8 months, and the standard deviation of the (cross-sectional) distribution of price spells is approximately 8.5 months. The correlation between the estimated and empirical sectoral weights is 0.8. The estimated extent of pricing complementarities is similar to the U.S.[30]

4.2 Other countries

Due to the limited cross-sectional information available for the other countries, we focus on a comparison between the average duration of price spells and cross-sectional standard deviation of the sectoral distribution of price spells implied by our macro-based estimates, and their empirical counterparts. The results are summarized in Table 6.

Countries are ranked (roughly) in descending order in terms of available level of detail about the cross-section of interest. Perhaps with the exception of the cross-sectional standard deviation of the duration of price spells in France, the results are broadly in line with the available microeconomic evidence. In particular, our macro-based estimates pick up the fact that France exhibits noticeably less price stickiness than the average in the Euro area (Dhyne et al. 2006). For completeness, in the Appendix we report a full set of parameter estimates for the remaining countries.

4.3 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 $k' \neq k$). For brevity, we focus on the U.S. economy, for which the disconnect between micro and macro estimates of price rigidity is well documented.

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 7 and Figure 4. 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

[30] An interesting question that we leave for future research is to analyze the differences across countries in estimated parameters. For example, why is the estimated degree of real rigidities in Denmark smaller than in the U.S. and Japan? Similarly, why are the estimated dynamics for nominal aggregate demand different?

Table 6: Estimates of moments of the cross-sectional distribution of price stickiness - flat prior

	Model with $K = 8$, $\alpha_0 = 8$		*Empirical distribution*	
	$\bar{k}$	σ_k	$\bar{k}$	σ_k
Switzerland[1)]	4.91 (3.71;5.98)	2.39 (1.81;2.68)	4.87	2.52
France	3.86 (2.72;5.03)	2.86 (2.35;3.09)	3.20	1.60
Italy[2)]	4.47 (3.20;5.63)	2.67 (2.07;2.93)	4.70	2.35
Norway[3)]	4.30 (3.20;5.41)	2.35 (1.72;2.71)	4.50	2.33
U.K.	4.46 (3.22;5.71)	2.65 (1.98;3.02)	–	–
Canada	4.65 (3.34;5.88)	2.78 (2.19;3.05)	–	–
Australia	4.98 (3.65;6.16)	2.68 (2.10;2.93)	–	–
Korea	4.52 (3.24;5.74)	2.72 (2.14;2.99)	–	–
Sweden	4.49 (3.23;5.73)	2.67 (1.99;3.02)	–	–

Note: Model-based estimates for the moments k and σ_k are computed using the means of the marginal posterior distributions as the point estimates of the sectoral weights; numbers in parentheses correspond to the 5^{th} and 95^{th} percentiles. The last two columns report the empirical moments computed from the available microeconomic evidence: 1) statistics based on the durations reported in Kaufmann (2008, Table 7); 2) statistics based on the implied median durations reported in Fabiani et al. (2006, Table A3.1); 3) durations for individual classes are scaled up proportionately so that the cross-sectional (weighted) average duration computed from the monthly statistics matches the mean duration reported by Wulfsberg (2009, Table 1) for the sub-period 1990-2004; this adjustment makes Wulfsberg's statistics more comparable to those of the other countries, which are based on more recent samples covering periods of lower inflation.

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 7: Best-fitting model with homogeneous firms

	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$^{\text{th}}$ and 95$^{\text{th}}$ 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.[31]

Table 8 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 $10^{11}:1$.

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

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

Table 8: Model comparison, U.S. 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.

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 5 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.[32] 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 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.

[32] 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.

5 Combining micro and macro data in the estimation

In this section we use the empirical cross-sectional distributions reported in Table 2 to incorporate microeconomic information in our Bayesian estimation through the use of "informative" priors. For that purpose, we specify the relative sizes of the hyperparameters $(\alpha_1, ..., \alpha_K)$ of the Dirichlet prior over the sectoral weights ω for each country so that the mode of the prior distributions coincide with the empirical sectoral weights $\widehat{\omega}$. That requires setting $\alpha_k = 1 + \widehat{\omega}_k (\alpha_0 - K)$. Incidentally, note that the flat-priors case analyzed previously obtains when $\alpha_0 = K$. As discussed in Subsection 3.1, we control the tightness of the prior by varying the parameter α_0.

Table 9 and Figures 6-8 present the results for the U.S., Denmark (DK), and Japan (JP), in terms of marginal distributions for the parameters assuming $\alpha_0 = 80$.[33] As expected, the posterior distributions for the sectoral weights now look even more similar to the empirical distributions. As an indication of the effect on the estimation of incorporating the microeconomic information, the correlation between estimated and empirical sectoral weights is now above 0.97 for all three countries (using the posterior means as the point estimates of the sectoral weights).

Incorporating the microeconomic information produces only small changes to the posterior distributions of the remaining parameters, in most cases. This was somewhat expected, since the distribution of the duration of price spells inferred purely from aggregate data already conforms well with the empirical distributions, especially for the U.S. and Japan. The largest changes in the estimates of the remaining parameters occur in the case of Denmark, which has the largest discrepancies between purely macro-based estimates of ω and empirical sectoral weights. This estimation exercise illustrates the potential for incorporating at least some of the vast amounts of microeconomic information about pricing behavior produced by the empirical literature into the estimation of macroeconomic models of price setting.

6 Consistency with other dimensions of the micro data

Our results show that allowing for heterogeneity in price rigidity goes a long way toward reconciling the degrees of price rigidity implied by micro and macro data. However, so far we have not discussed how the estimated multi-sector models perform when confronted with other features of price setting documented in the recent empirical literature. Here we provide such an assessment. Specifically,

[33] In previous versions of the paper we reported results with different degrees of prior tightness. The results reported here are representative of what we learn from alternative configurations.

Table 9: Parameter estimates for the U.S., Denmark (DK), and Japan (JP) - informative priors

	Models with $K = 8$, $\alpha_0 = 80$						*Empirical distributions*		
	U.S.		*DK*		*JP*		*U.S.*	*DK*	*JP*
ζ	0.041 (0.02;0.10)	0.047	0.228 (0.09;0.59)	0.269	0.029 (0.02;0.06)	0.033	–	–	–
ω_1	0.282 (0.21;0.37)	0.283	0.158 (0.10;0.23)	0.161	0.307 (0.23;0.39)	0.309	0.273	0.136	0.311
ω_2	0.071 (0.03;0.12)	0.074	0.152 (0.10;0.22)	0.154	0.179 (0.12;0.25)	0.18	0.071	0.179	0.204
ω_3	0.069 (0.04;0.11)	0.071	0.040 (0.02;0.08)	0.043	0.043 (0.02;0.09)	0.046	0.098	0.045	0.043
ω_4	0.087 (0.05;0.14)	0.090	0.078 (0.04;0.13)	0.081	0.046 (0.02;0.09)	0.049	0.110	0.091	0.045
ω_5	0.071 (0.03;0.13)	0.074	0.118 (0.07;0.18)	0.121	0.065 (0.03;0.12)	0.068	0.059	0.124	0.070
ω_6	0.138 (0.08;0.21)	0.141	0.069 (0.03;0.12)	0.072	0.048 (0.02;0.09)	0.051	0.129	0.067	0.046
ω_7	0.067 (0.03;0.12)	0.071	0.116 (0.06;0.19)	0.119	0.046 (0.02;0.10)	0.050	0.061	0.108	0.039
ω_8	0.193 (0.13;0.27)	0.195	0.247 (0.17;0.33)	0.249	0.244 (0.17;0.33)	0.246	0.198	0.250	0.244
$\bar{k}$	4.28 (3.81;4.75)	4.28	4.79 (4.31;5.25)	4.78	3.97 (3.47;4.45)	3.97	4.25	4.77	3.87
σ_k	2.68 (2.51;2.83)	2.70	2.62 (2.45;2.78)	2.64	2.85 (2.67;3.00)	2.86	2.66	2.60	2.85
ρ_0	0.000 (−0.00;0.00)	0.000	0.001 (−0.00;0.00)	0.001	0.000 (−0.00;0.00)	0.000	–	–	–
ρ_1	1.429 (1.28;1.58)	1.428	0.779 (0.61;0.94)	0.778	1.364 (1.19;1.56)	1.366	–	–	–
ρ_2	−0.449 (−0.60;−0.30)	−0.448	0.179 (0.01;0.35)	0.179	−0.386 (−0.58;−0.21)	−0.389	–	–	–
σ_m	0.005 (0.00;0.01)	0.005	0.013 (0.01;0.02)	0.013	0.011 (0.01;0.02)	0.011	–	–	–
δ_0	0.002 (−0.00;0.01)	0.002	0.002 (−0.00;0.01)	0.002	0.005 (−0.00;0.02)	0.005	–	–	–
δ_1	0.544 (0.37;0.72)	0.543	0.122 (−0.06;0.31)	0.125	0.375 (0.20;0.57)	0.378	–	–	–
δ_2	0.150 (0.00;0.30)	0.150	0.360 (0.24;0.49)	0.361	0.289 (0.13;0.44)	0.288	–	–	–
σ_n	0.066 (0.03;0.16)	0.077	0.091 (0.04;0.21)	0.104	0.145 (0.07;0.27)	0.154	–	–	–

Note: For each country the two columns report, respectively, the medians and means of the marginal posterior distributions; numbers in parentheses correspond to the 5^{th} and 95^{th} percentiles; the last three columns reproduce the empirical distributions from Table 2.

we focus on the facts documented by Klenow and Kryvtsov (2008) for the U.S., and discuss the ability of the estimated model to match them - at least qualitatively. It turns out that of the eight price-setting facts documented by Klenow and Kryvtsov,[34] four are matched by our model in its current form. Moreover, we can adapt the model to match an additional three facts without any effect on its implications for aggregate dynamics.

Specifically, our estimated model matches the empirical facts when it comes to i) frequent price changes; ii) many small price changes; iii) the intensive margin dominates the variance of inflation; and iv) the frequency of price increases and decreases moves with inflation. The last feature of the model probably deserves some comment. In standard time-dependent pricing models the total frequency of price changes is exogenous and constant, and as such it obviously does not covary with inflation. However, in these models the frequency of price increases usually covaries positively with inflation, whereas the frequency of price decreases covaries negatively.

In contrast with fact ii), the model in its current form fails to match fact v): the large average absolute size of price changes. This is because the model only includes aggregate shocks, which are too small to produce enough price changes that are as large as in the micro data. We can easily change this feature of the model to match fact v), by adding firm-level heterogeneity in the form of idiosyncratic shocks that induce a non-degenerate distribution of prices among adjusting firms in any given sector. As long as these shocks cancel out when averaged over a continuum of firms, they do not have any first-order effect on aggregate dynamics. The reason is that with such shocks the (loglinear) optimal price setting equation (2) is replaced by:

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

where $\varepsilon_t(k,j)$ is proportional to the idiosyncratic shock to firm kj.[35] Thus, (2) can be interpreted as the average price set by firms from sector k that change prices at time t. As a result, all of the aggregate results are identical under this specification. As long as the idiosyncratic shocks cancel out in the cross-section, they can be used to match moments of the distribution of the size of price changes.[36]

Other two facts that the current model fails to match are: vi) variable durations of price spells;

[34] Table VIII in Klenow and Kryvtsov (2008).

[35] Here we implicitly assume that idiosyncratic shocks follow a first-order Markov process.

[36] This additional dimension of the price micro data might impose restrictions on the nature of the real rigidities that are assume in the model. The reason is that some types of real rigidity might affect the coefficient of proportionality between $\varepsilon_t(k,j)$ and the underlying idiosyncratic shock. See, e.g., Klenow and Willis (2006).

and vii) flat hazard rates for price adjustment. The source of the mismatch is the assumption that all firms in sector k always set prices for k periods. It turns out that this assumption can be replaced with one that allows the model to match these two facts without affecting any of its aggregate implications.

Start from a given multi-sector economy with Taylor pricing, as in our estimated models. Rather than having sector-specific time-invariant durations of price spells, the idea is to add some random variation in price spells in each sector, as follows. Each time a firm from sector k is about to set a new price, it draws the duration for the new price spell from a distribution that has mean k (periods). Note that this is *not* the assumption underlying either the Calvo (1983) model, or models with generalized adjustment hazards (e.g., Dotsey et al. 1997). Here the durations of the spells are known by firms when they set their prices. The duration draws are independent across all firms. Any economy-wide distribution of price spells can be obtained through an appropriate choice of sectoral weights and sector-specific distributions of spells. Thus, although each firm now has time-varying spells of price rigidity, at any point in time the cross-sectional distributions of spells of rigidity and of the corresponding prices set by firms are exactly the same as in the original multi-sector Taylor economy. This produces identical aggregate dynamics in the two models. This modification takes care of fact vi).

As for fact vii), if one has access to the empirical distributions of spells used to estimate the empirical hazard rates of price adjustments, it is clear that one can pick the sector-specific distributions of durations to produce the same hazard estimates.[37] Note, however, that the pooling of observed spells leads the econometrician to estimate a misspecified model for the adjustment hazard. In reality, in the model with the duration draws laid out in the previous paragraph, the true (unobserved) hazard function for each spell is still of the Taylor variety: the firm knows the duration of each price spell at its start. In contrast, the hazard estimated from the pooled observations would be non-zero at more than just one point.

Finally, we believe that the extended versions of the model described in this section would sill have difficulties matching the eighth fact documented by Klenow and Kryvstov (2008): the size of price changes does not increase with the duration of price spells. In the estimated model with only aggregate shocks, these are somewhat persistent, and as a result the innovations that are incorporated at each price change "build up" over time. On average, this leads to larger price changes after longer price spells. The introduction of more transient idiosyncratic shocks should attenuate

[37]The reason is that the modification spelled out in the previous paragraph does not restrict features of the sectoral distributions of spells, other than their first moments.

this effect, since these shocks add a component to firms' desired prices that does not become so dispersed over time. Thus, in the micro data produced by a model with such relatively transient idiosyncratic shocks, the size of price changes would appear to be less related to the duration of price spells relative to the model with only aggregate shocks, especially if the latter are small relative to the idiosyncratic shocks. This might lead to an apparent lack of an empirical relationship between the duration of price spells and the size of price changes in small (short time-series) samples.[38] In fact, the same small-sample problem might apply to the regressions estimated by Klenow and Kryvtsov (2008) on actual micro data. Whether our conjecture is correct is a question we leave for future research.

7 Robustness, and directions for future research

7.1 Robustness[39]

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 $p_1 = p_2 = 3$. In all cases that we analyzed we found overwhelming support for the models with heterogeneity.

We also considered versions of the models with Calvo (1983) pricing. In that case, not all of our conclusions are equally robust. 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 crucially on whether the observable driving process has high variance relative to the unobservable process. While this applies to both price-setting specifications, the identification problem is more acute under Calvo pricing. The reason is that, in terms of implications for the aggregate dynamics of output and prices, the differences between specifications with varying degrees of price stickiness are starker under Taylor pricing than under Calvo pricing. Based on Monte Carlo analysis, we found that with the sample size that we have and the relative variances for the two exogenous processes implied by our point estimates, the likelihood criterion fails to provide

[38] The cumulative effect of aggregate innovations would still be present in this context, and we conjecture that with enough artificial micro data one should be able to detect that relationship.

[39] Due to the high computational cost of estimating the models, in our robustness exercises we focus on the U.S. economy.

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 clear discrimination between models with heterogeneity in the frequency of price changes. 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.

However, despite that obstacle, one of the main findings of the paper *does* hold under the Calvo pricing model: the (log) posterior density of the data given specifications with heterogeneity in price stickiness is roughly 5-7 points higher than under the best specification with homogeneous firms. In addition, using classical methods we find that a likelihood-ratio test of the homogeneous Calvo model against a two-sector version of the model leads to rejection of the former at significance levels of less than 1%.[40]

7.2 Directions for future research

Our experience based on specifications with Taylor and Calvo pricing models suggests that the shape of the hazard function for price adjustments assumed in the price-setting model is important in determining how precisely the cross-sectional distribution of price stickiness can be inferred from aggregate data. An interesting way to address this question would be to specify a generalized pricing model with a more flexible price adjustment hazard than in the Calvo and Taylor models, and take the model to the data allowing for sectoral heterogeneity in the hazards.[41] The question, then, would be how to use the microeconomic evidence on price setting to inform the priors over the nature of such adjustment hazard functions. One alternative would be to use a parsimonious parametric family of hazard functions, say a two-parameter family characterized by level and slope of the hazard. Then, one could use the microeconomic evidence on the frequency of price changes and on the shape of adjustment hazard functions estimated from microeconomic data (e.g., Klenow and Kryvtsov 2008) to form priors over those two parameters, and estimate the model using aggregate data as observables, as we do.

We wish to emphasize that our results *do not* imply that identification of more nuanced features

[40] The likelihood-ratio statistic ranges from roughly 10.5 to 14 (depending on the specification for the exogenous time-series processes), whereas the 0.1% and 1% critical values for the $\chi^2(1)$ distribution are, respectively, 10.83 and 6.64.

[41] Dotsey et al. (1997) proposed such a generalized price-setting model, assuming that all firms are ex-ante identical. Similar specifications have been used subsequently by Wolman (1999) and Mash (2004), among others. Coenen et al. (2007), Guerrieri (2006), Sheedy (2007), and Yao (2009) estimate models with generalized price-setting hazards using aggregate data. To our knowledge the only paper to allow for generalized adjustment hazards and ex-ante heterogeneity in models of price setting is Carvalho and Schwartzman (2008).

of the distribution of price stickiness from aggregate data is infeasible under Calvo pricing. In fact, as we show in the Appendix, in this version of the model the sectoral weights are also identified in the formal sense. However, our findings do suggest that, in practice, additional structure is needed for estimation. It is possible that the restrictions obtained by moving to a fully specified model and using additional observables in the estimation will impose more "discipline" on the latent stochastic processes and thus attenuate the identification problems we encountered in our semi-structural framework with Calvo pricing. In addition, making use of sectoral data as well, along the lines of Lee (2009), and Bouakez et al. (2009a,b), seems promising.

Finally, in the previous section we described a variant of the model that we estimate that can - at least qualitatively - match additional micro price facts while preserving the exact same aggregate implications as those of our estimated models. Moreover, that variant has the potential to perform at least as well as the Calvo (1983) model in terms of matching the empirical facts documented by Klenow and Kryvtson (2008). An interesting open (quantitative) question is how well such an extended model performs when confronted with those facts.

8 Conclusion

We estimate small semi-structural models for twelve countries, allowing the extent of price rigidity to vary across sectors. We provide estimates of the underlying cross-sectional distribution based only on aggregate data, and estimates that incorporate prior microeconomic information from the recent empirical price-setting literature. Perhaps surprisingly, we find that the macro-based estimates accord well with the latter evidence. More generally, we find overwhelming empirical support for specifications with heterogeneity in price stickiness, over ones in which all prices are equally sticky.

We find the results sufficiently compelling to warrant further research. In particular, it would be interesting to evaluate the consequences of allowing for heterogenous pricing behavior when estimating fully specified DSGE models on aggregate data. The experience with our semi-structural model suggests that combining microeconomic information and macroeconomic data within a Bayesian framework can help us integrate our views on price setting at the microeconomic and macroeconomic levels. Quantitative normative analysis in models with heterogeneity in price stickiness, along the lines of Eusepi et al. (2009), might also benefit from such a combination.

References

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

[2] Baudry, L., H. Le Bihan, P. Sevestre and S. Tarrieu (2007), "What Do Thirteen Million Price Records Have to Say about Consumer Price Rigidity?" *Oxford Bulletin of Economics and Statistics* 69: 139-183.

[3] 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.

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

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

[6] ____________ (2009b), "Sectoral Price Rigidity and Aggregate Dynamics," mimeo available at http://www.cireq.umontreal.ca/personnel/ruge.html.

[7] Browning, M., L. Hansen and J. Heckman (1999), "Micro data and general equilibrium models," *Handbook of Macroeconomics*, Volume 1, Part 1: 543-633.

[8] Bunn, P. and C. Ellis (2009), "Price-setting behaviour in the United Kingdom: a microdata approach," Bank of England Quarterly Bulletin 2009 Q1.

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

[10] Carvalho, C. and F. Schwartzman (2008), "Heterogeneous Price Setting Behavior and Aggregate Dynamics: Some General Results," mimeo available at http://www.newyorkfed.org/research/economists/carvalho/papers.html.

[11] Chang, Y. and S. Kim (2006), "From Individual to Aggregate Labor Supply: A Quantitative Analysis Based on a Heterogeneous Agent Macroeconomy," *International Economic Review* 47: 1 - 27.

[12] 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.

[13] Dhyne, E., L. Álvarez, H. Le Bihan, G. Veronese, D. Dias, J. Hoffman, N. Jonker, P. Lünnemann, F. Rumler and J. Vilmunen (2006), "Price Changes in the Euro Area and the United States: Some Facts from Individual Consumer Price Data," *Journal of Economic Perspectives* 20: 171-192.

[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] Enomoto, H. (2007), "Multi-Sector Menu Cost Model, Decreasing Hazard, and Phillips Curve," Bank of Japan Working Paper Series No. 07-E-3.

[16] Eusepi, S., B. Hobijn and A. Tambalotti (2009), "CONDI: A Cost-of-Nominal-Distortions Index," Federal Reserve Bank of New York Staff Reports 367.

[17] Fabiani S., A. Gattulli, R. Sabbatini and G. Veronese (2006), "Consumer Price Setting In Italy," *Giornale degli Economisti e Annali di Economia* 65: 31-74.

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

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

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

[21] Guimarães, B. and K. Sheedy, (2009), "Sales and Monetary Policy," forthcoming in the *American Economic Review.*

[22] Hansen, B. and N. Hansen (2006), "Price Setting Behavior in Denmark: A Study of CPI Micro Data 1997-2005," Danmarks Nationalbank Working Paper 39.

[23] Harchaoui, T., C. Michaud and J. Moreau (2008), "Consumer Price Changes in Canada, 1995-2006," Yearbook on Productivity 2007, Statistics Sweden.

[24] Higo, M. and Y. Saita (2007), "Price Setting in Japan: Evidence from CPI Micro Data," Bank of Japan Working Paper Series No. 07-E-20.

[25] Imbs, J., Jondeau, E., and F. Pelgrin (2007), "Aggregating Phillips Curves," ECB Working Paper Series no. 785.

[26] Imbs, J. and I. Mejean (2009), "Elasticity Optimism," CEPR Discussion Paper 7177.

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

[28] Kaufmann, D. (2008), "Price-Setting Behaviour in Switzerland: Evidence from CPI Micro Data," Swiss National Bank Working Papers 2008-15.

[29] Kehoe, P. and V. Midrigan (2008), "Temporary Price Changes and the Real Effects of Monetary Policy," Federal Reserve Bank of Minneapolis Research Department Staff Report 413.

[30] Klenow, P. and B. Malin (2009), "Microeconomic Evidence on Price Setting," mimeo prepared for the *Handbook of Monetary Economics*.

[31] 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

[32] 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.

[33] Lee, Jae Won (2009), "Heterogeneous Households in a Sticky Price Model," mimeo available at http://econweb.rutgers.edu/jwlee/.

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

[35] Mash, R. (2004), "Optimising Microfoundations for Inflation Persistence," University of Oxford Economics Series Working Papers #183.

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

[37] ____________ (2010), "Monetary Non-Neutrality in a Multi-Sector Menu Cost Model," forthcoming in the *Quarterly Journal of Economics*.

[38] Ravina, E. (2007), "Habit Persistence and Keeping Up with the Joneses: Evidence from Micro Data," mimeo available at http://www0.gsb.columbia.edu/faculty/eravina/habit_microdata.pdf.

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

[40] Sheedy, K. (2007), "Intrinsic Inflation Persistence," mimeo available at http://personal.lse.ac.uk/sheedy/.

[41] Shimer, R. (2009), "Convergence in Macroeconomics: The Labor Wedge," *American Economic Journal: Macroeconomics* 1: 280–297.

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

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

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

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

[46] Wolman, A. (1999), "Sticky Prices, Marginal Cost, and the Behavior of Inflation," Federal Reserve Bank of Richmond Economic Quarterly 85 (Fall 1999), 29-48.

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

[48] Wulfsberg, F. (2009), "Price Adjustments and Inflation: Evidence from Consumer Price Data in Norway 1975-2004," Norges Bank WP 2009/11.

[49] Yao, F. (2009), "Real and Nominal Rigidities in Price Setting: A Bayesian Analysis Using Aggregate Data," mimeo available at http://ideas.repec.org/p/hum/wpaper/sfb649dp2009-057.html.

Appendix

A Identification

A.1 When $\zeta \neq 1$

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

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

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.[42] We conduct both a large- (1000 observations) and a small-sample exercise (100 observations, as in our actual sample). Table 10 reports the results.

Table 10: 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.[43] The

[42] 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).

[43] 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.

"Ini. guess" column reports the average value of the initial guesses supplied for the optimization algorithm across 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).

A.3 In a multi-sector Calvo (1983) model

Under the assumption of Calvo pricing, equations (6) and (5) are replaced by, respectively:

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

and

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

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

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

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

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

We focus on the case of strategic neutrality in price setting ($\zeta = 1$). Then, (15) simplifies to:

$$\begin{aligned} x_t(k) &= (1 - \beta\alpha_k) E_t \sum_{i=0}^{\infty} (\alpha_k\beta)^i \left(m_{t+i} - y_{t+i}^n\right) \\ &= F(\beta, \rho_1, \alpha_k) m_t - F(\beta, \delta_1, \alpha_k) y_t^n. \end{aligned}$$

In that case the aggregate price level evolves according to:

$$\begin{aligned} p_t &= \sum_{j=0}^{\infty}\sum_{k=1}^{K} \omega_k (1 - \beta\alpha_k) F(\beta, \rho_1, \alpha_k) \alpha_k^{-j} m_{t-j} \\ &\quad - \sum_{j=0}^{\infty}\sum_{k=1}^{K} \omega_k (1 - \beta\alpha_k) F(\beta, \delta_1, \alpha_k) \alpha_k^{-j} y_{t-j}^n. \end{aligned}$$

We illustrate how identification obtains in a model with $K = 2$. As in Subsection 2.4, the starting point is a set of consistent estimates of the coefficients on m_{t-j} $(\sum_{k=1}^{K}\omega_k\,(1-\beta\alpha_k)\,F(\beta,\rho_1,\alpha_k)\,\alpha_k^{-j})$, which we denote by a_j. With $K = 2$, this implies the following system of equations:

$$\omega_1\,(1-\beta\alpha_1)\,F(\beta,\rho_1,\alpha_1) + (1-\omega_1)\,(1-\beta\alpha_2)\,F(\beta,\rho_1,\alpha_2) = a_0$$
$$\omega_1\,(1-\beta\alpha_1)\,F(\beta,\rho_1,\alpha_1)\,\alpha_1^{-1} + (1-\omega_1)\,(1-\beta\alpha_2)\,F(\beta,\rho_1,\alpha_2)\,\alpha_2^{-1} = a_1$$
$$\omega_1\,(1-\beta\alpha_1)\,F(\beta,\rho_1,\alpha_1)\,\alpha_1^{-2} + (1-\omega_1)\,(1-\beta\alpha_2)\,F(\beta,\rho_1,\alpha_2)\,\alpha_2^{-2} = a_2,$$

which can be solved for ω_1, α_1, and α_2 as a function of a_0, a_1 and a_2 (and β, ρ_1).

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.[44] 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

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

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[45] (PSR) factors for each parameter and inspecting the 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:

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

[45] 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. 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.

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

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

with:

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

So:

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

To recover the v_k's from ω simply set:

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

C Estimates for other countries

Tables 11 and 12 present the full set of estimates for the remaining countries:

Table 11: Parameter estimates for remaining countries - K=8, flat prior

	Prior	*Switzerland*	*France*	*Italy*	*Norway*
ζ	4.440 (0.466;16.863)	0.403 (0.103;1.387)	0.193 (0.074;0.474)	1.272 (0.479;3.312)	8.068 (3.157;18.516)
ω_1	0.094 (0.007;0.348)	0.072 (0.016;0.233)	0.365 (0.143;0.615)	0.278 (0.105;0.519)	0.128 (0.028;0.335)
ω_2	0.094 (0.007;0.348)	0.043 (0.003;0.190)	0.058 (0.005;0.205)	0.027 (0.002;0.106)	0.163 (0.022;0.383)
ω_3	0.094 (0.007;0.348)	0.219 (0.088;0.359)	0.110 (0.016;0.234)	0.055 (0.005;0.169)	0.055 (0.004;0.206)
ω_4	0.094 (0.007;0.348)	0.080 (0.007;0.236)	0.045 (0.003;0.170)	0.037 (0.003;0.137)	0.089 (0.008;0.309)
ω_5	0.094 (0.007;0.348)	0.109 (0.010;0.305)	0.029 (0.002;0.119)	0.077 (0.006;0.255)	0.123 (0.010;0.398)
ω_6	0.094 (0.007;0.348)	0.050 (0.004;0.191)	0.034 (0.002;0.134)	0.070 (0.006;0.237)	0.070 (0.006;0.264)
ω_7	0.094 (0.007;0.348)	0.061 (0.005;0.203)	0.049 (0.004;0.187)	0.271 (0.051;0.523)	0.098 (0.008;0.341)
ω_8	0.094 (0.007;0.348)	0.255 (0.061;0.467)	0.221 (0.083;0.385)	0.067 (0.003;0.276)	0.087 (0.007;0.313)
$\bar{k}$	4.501 (3.245;5.760)	4.940 (3.711;5.979)	3.848 (2.718;5.028)	4.496 (3.203;5.633)	4.305 (3.198;5.410)
σ_k	2.139 (1.584;2.678)	2.284 (1.812;2.683)	2.785 (2.350;3.088)	2.562 (2.067;2.925)	2.235 (1.720;2.713)
ρ_0	0.000 (−8.224;8.224)	0.000 (−0.001;0.002)	−0.000 (−0.001;0.001)	−0.000 (−0.002;0.001)	0.002 (−0.002;0.006)
ρ_1	0.000 (−8.224;8.224)	1.528 (1.382;1.671)	1.610 (1.492;1.726)	1.523 (1.387;1.657)	0.920 (0.748;1.091)
ρ_2	0.000 (−8.224;8.224)	−0.547 (−0.690;−0.403)	−0.618 (−0.734;−0.500)	−0.534 (−0.670;−0.398)	−0.011 (−0.184;0.162)
σ_m	0.059 (0.009;0.195)	0.006 (0.006;0.007)	0.005 (0.004;0.006)	0.009 (0.008;0.010)	0.022 (0.020;0.025)
δ_0	0.000 (−8.224;8.224)	0.001 (−0.002;0.006)	0.000 (−0.002;0.002)	0.000 (−0.001;0.002)	−0.000 (−0.003;0.003)
δ_1	0.000 (−8.224;8.224)	0.223 (−0.081;0.771)	0.770 (0.442;1.037)	0.862 (0.589;1.110)	0.461 (0.248;0.642)
δ_2	0.000 (−8.224;8.224)	0.454 (0.093;0.676)	0.137 (−0.111;0.419)	0.062 (−0.169;0.311)	0.441 (0.261;0.639)
σ_n	0.059 (0.009;0.195)	0.030 (0.012;0.107)	0.012 (0.008;0.025)	0.009 (0.007;0.013)	0.016 (0.014;0.020)

Note: The first column reports the medians of the marginal prior distributions; the other columns provide the analogous statistics for the posterior marginal distributions in each country; numbers in parentheses correspond to the 5$^{\text{th}}$ and 95$^{\text{th}}$ percentiles.

Table 12: Parameter estimates for remaining countries - K=8, flat prior, continued

	Prior	*U.K.*	*Canada*	*Australia*	*Korea*	*Sweden*
ζ	4.440 (0.466;16.863)	3.681 (1.269;10.330)	1.040 (0.389;2.776)	0.599 (0.231;1.490)	0.556 (0.102;1.630)	3.035 (1.053;8.697)
ω_1	0.094 (0.007;0.348)	0.254 (0.100;0.484)	0.253 (0.090;0.490)	0.197 (0.069;0.417)	0.211 (0.064;0.452)	0.247 (0.086;0.488)
ω_2	0.094 (0.007;0.348)	0.040 (0.003;0.153)	0.055 (0.004;0.181)	0.055 (0.004;0.178)	0.119 (0.017;0.271)	0.035 (0.003;0.136)
ω_3	0.094 (0.007;0.348)	0.037 (0.003;0.147)	0.028 (0.002;0.108)	0.037 (0.003;0.130)	0.037 (0.003;0.128)	0.063 (0.005;0.222)
ω_4	0.094 (0.007;0.348)	0.086 (0.006;0.304)	0.070 (0.006;0.223)	0.035 (0.003;0.129)	0.062 (0.005;0.197)	0.078 (0.006;0.299)
ω_5	0.094 (0.007;0.348)	0.072 (0.005;0.264)	0.054 (0.004;0.193)	0.065 (0.006;0.209)	0.078 (0.007;0.243)	0.046 (0.003;0.192)
ω_6	0.094 (0.007;0.348)	0.093 (0.007;0.321)	0.060 (0.005;0.220)	0.114 (0.012;0.295)	0.078 (0.007;0.248)	0.072 (0.006;0.272)
ω_7	0.094 (0.007;0.348)	0.093 (0.007;0.315)	0.147 (0.015;0.398)	0.179 (0.019;0.430)	0.079 (0.006;0.271)	0.143 (0.012;0.413)
ω_8	0.094 (0.007;0.348)	0.158 (0.014;0.444)	0.209 (0.031;0.461)	0.210 (0.030;0.443)	0.219 (0.037;0.451)	0.146 (0.012;0.425)
$\bar{k}$	4.501 (3.245;5.760)	4.456 (3.220;5.710)	4.665 (3.338;5.883)	5.012 (3.650;6.162)	4.529 (3.241;5.741)	4.501 (3.231;5.732)
σ_k	2.139 (1.584;2.678)	2.540 (1.975;3.020)	2.685 (2.195;3.049)	2.573 (2.102;2.927)	2.615 (2.137;2.991)	2.558 (1.994;3.022)
ρ_0	0.000 (−8.224;8.224)	−0.000 (−0.002;0.001)	0.001 (−0.001;0.002)	0.000 (−0.001;0.002)	−0.001 (−0.004;0.002)	0.001 (−0.001;0.003)
ρ_1	0.000 (−8.224;8.224)	1.198 (1.034;1.363)	1.516 (1.381;1.650)	1.478 (1.334;1.622)	1.476 (1.321;1.676)	1.074 (0.907;1.241)
ρ_2	0.000 (−8.224;8.224)	−0.207 (−0.373;−0.042)	−0.553 (−0.687;−0.418)	−0.493 (−0.636;−0.350)	−0.485 (−0.697;−0.327)	−0.099 (−0.263;0.067)
σ_m	0.059 (0.009;0.195)	0.008 (0.007;0.009)	0.007 (0.007;0.008)	0.009 (0.008;0.010)	0.016 (0.015;0.019)	0.013 (0.012;0.015)
δ_0	0.000 (−8.224;8.224)	0.000 (−0.001;0.001)	0.001 (−0.001;0.003)	0.000 (−0.002;0.003)	0.001 (−0.004;0.009)	0.001 (−0.001;0.003)
δ_1	0.000 (−8.224;8.224)	1.100 (0.842;1.322)	0.867 (0.591;1.112)	0.584 (0.297;0.848)	0.677 (0.365;0.949)	0.865 (0.650;1.068)
δ_2	0.000 (−8.224;8.224)	−0.148 (−0.367;0.102)	0.041 (−0.186;0.287)	0.242 (0.013;0.475)	0.196 (−0.068;0.482)	0.094 (−0.106;0.304)
σ_n	0.059 (0.009;0.195)	0.006 (0.005;0.008)	0.012 (0.009;0.019)	0.021 (0.014;0.037)	0.037 (0.025;0.144)	0.012 (0.010;0.016)

Note: The first column reports the medians of the marginal prior distributions; the other columns provide the analogous statistics for the posterior marginal distributions in each country; numbers in parentheses correspond to the 5^{th} and 95^{th} percentiles.

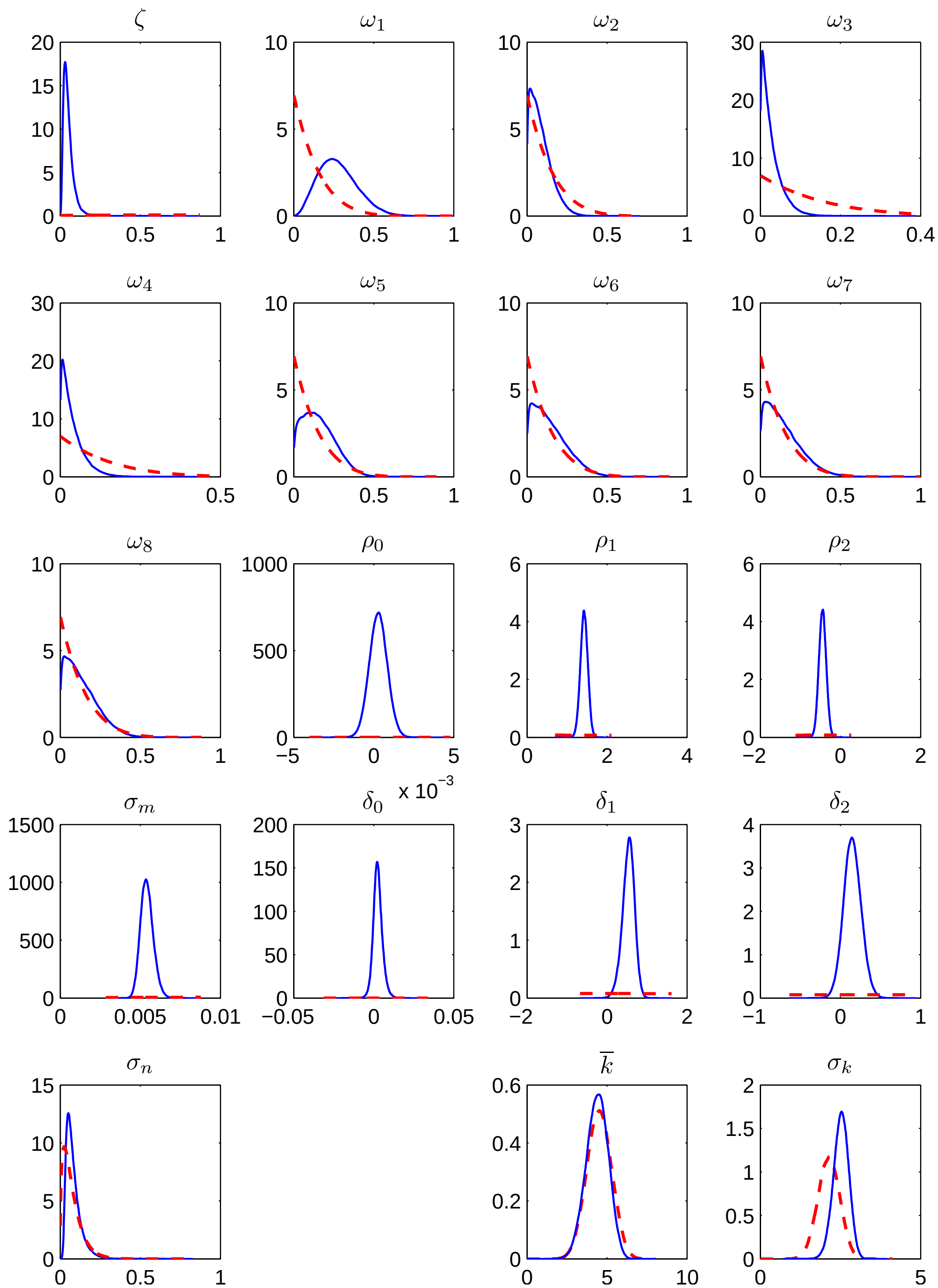

Figure 1: Marginal prior (dashed line) and posterior (solid line) distributions, K=8, flat prior - U.S.

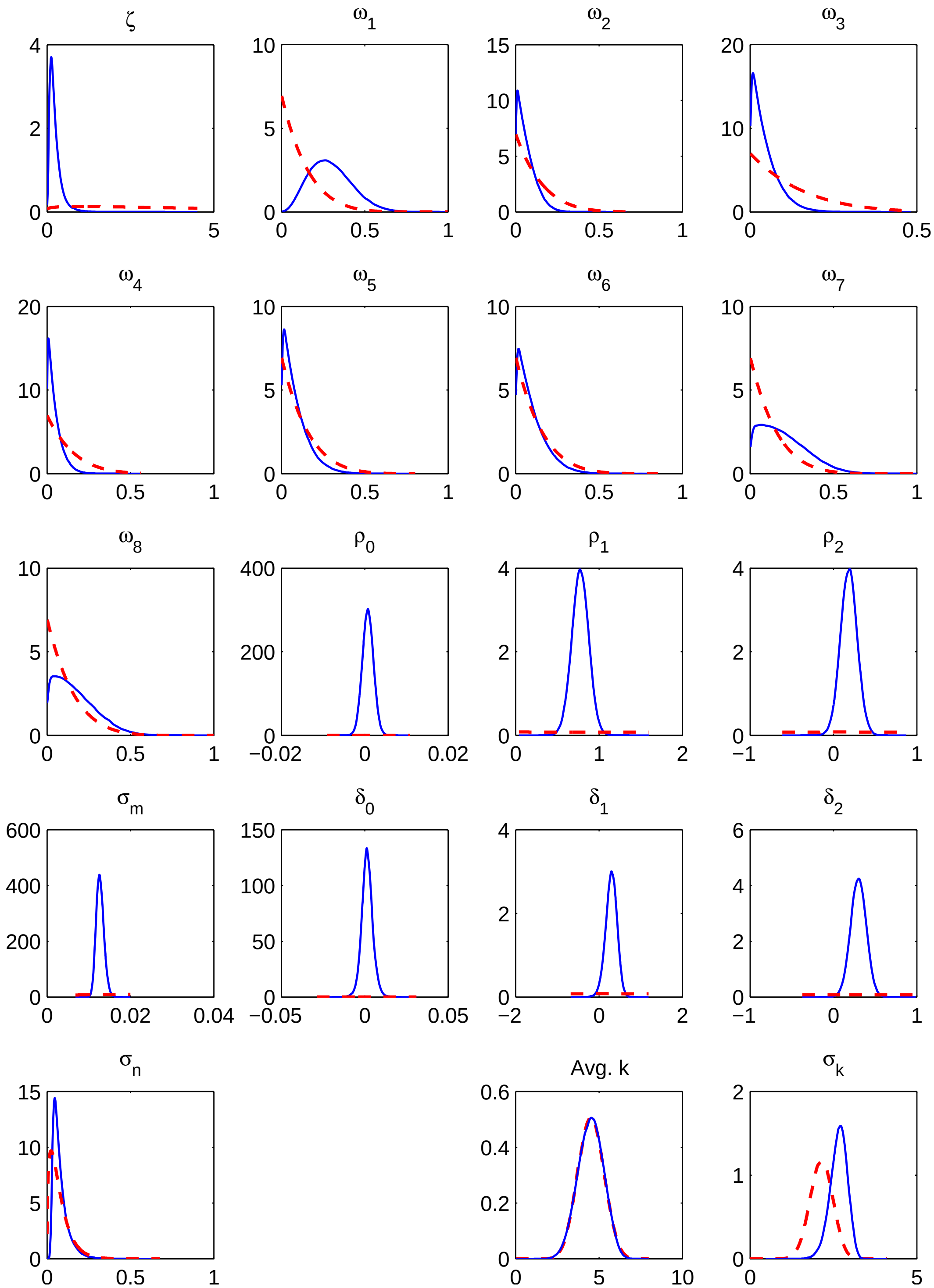

Figure 2: Marginal prior (dashed line) and posterior (solid line) distributions, K=8, flat prior - Denmark

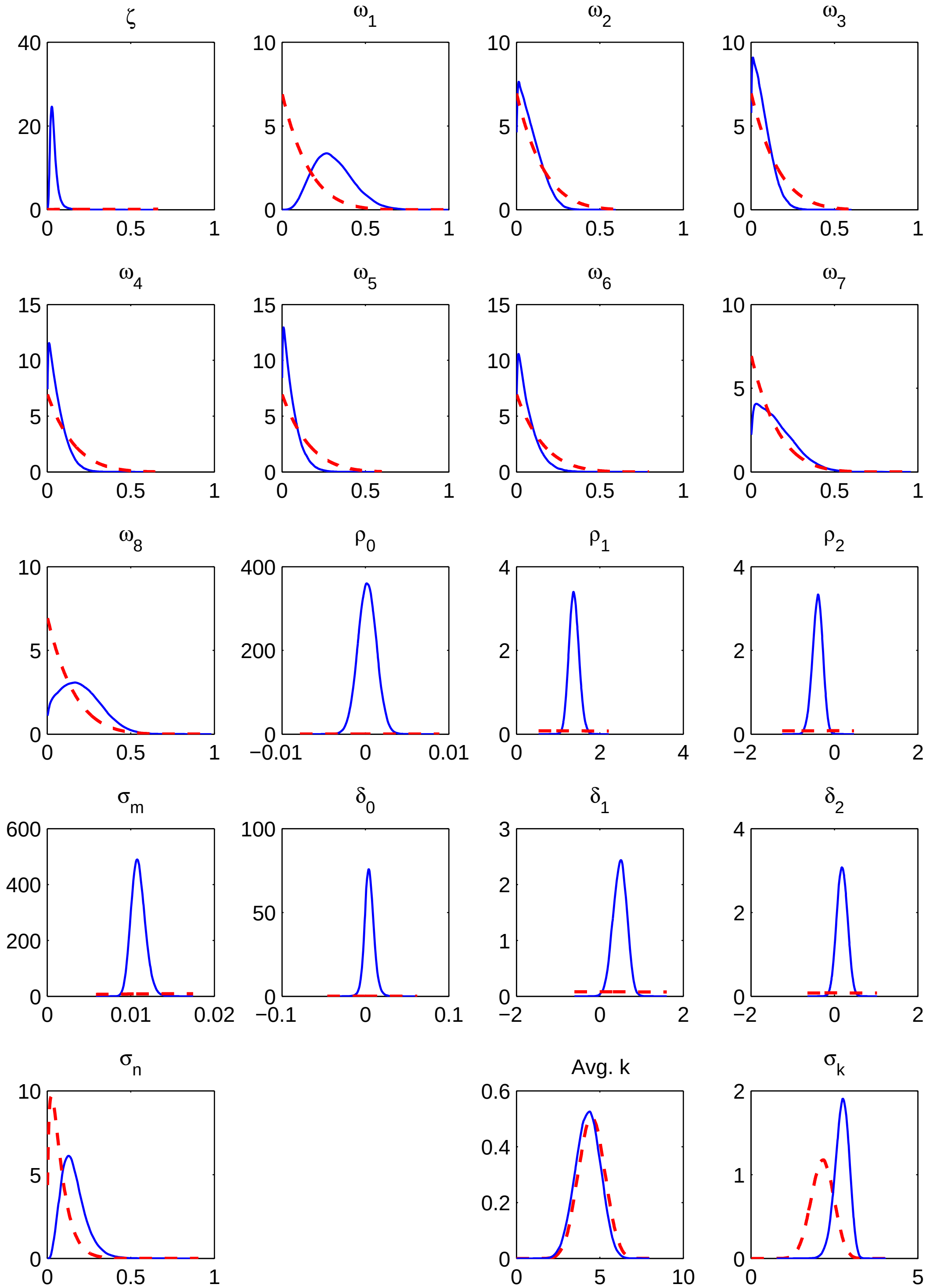

Figure 3: Marginal prior (dashed line) and posterior (solid line) distributions, K=8, flat prior - Japan

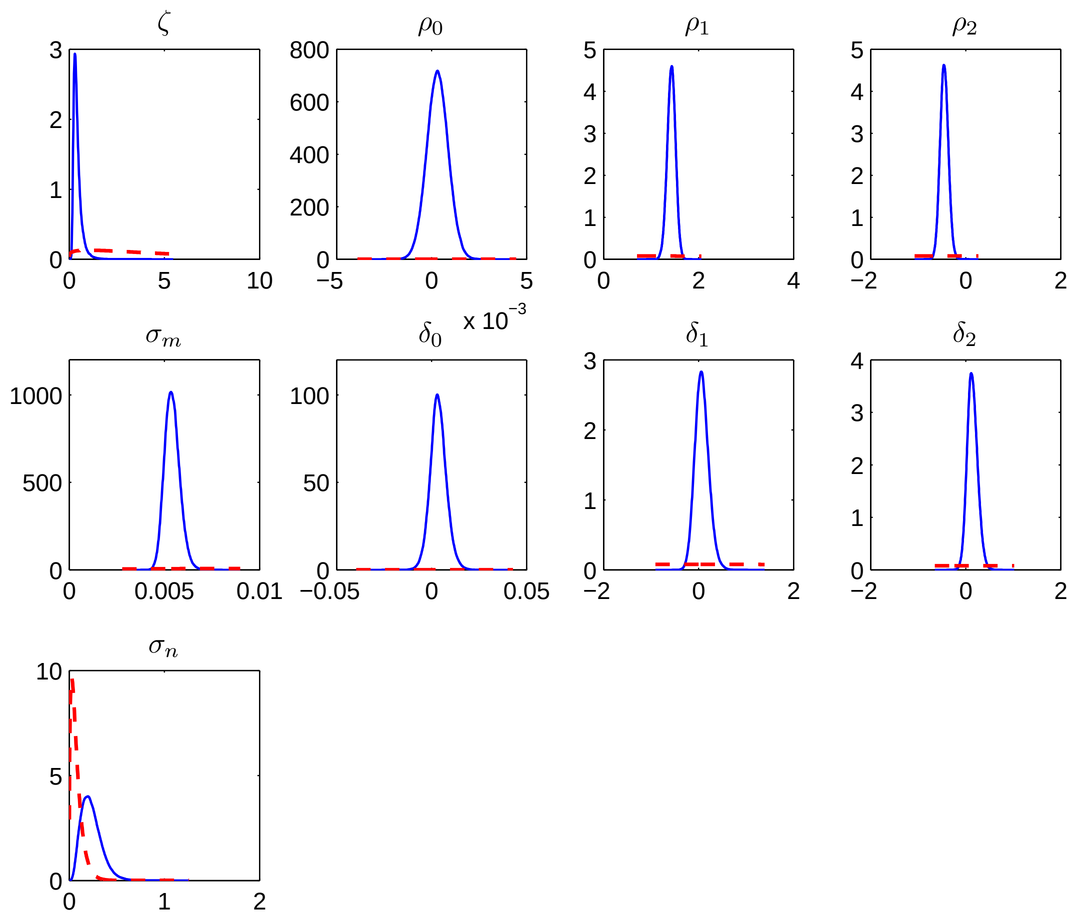

Figure 4: Marginal prior (dashed line) and posterior (solid line) distributions, one-sector model with 7-quarter price spells - U.S.

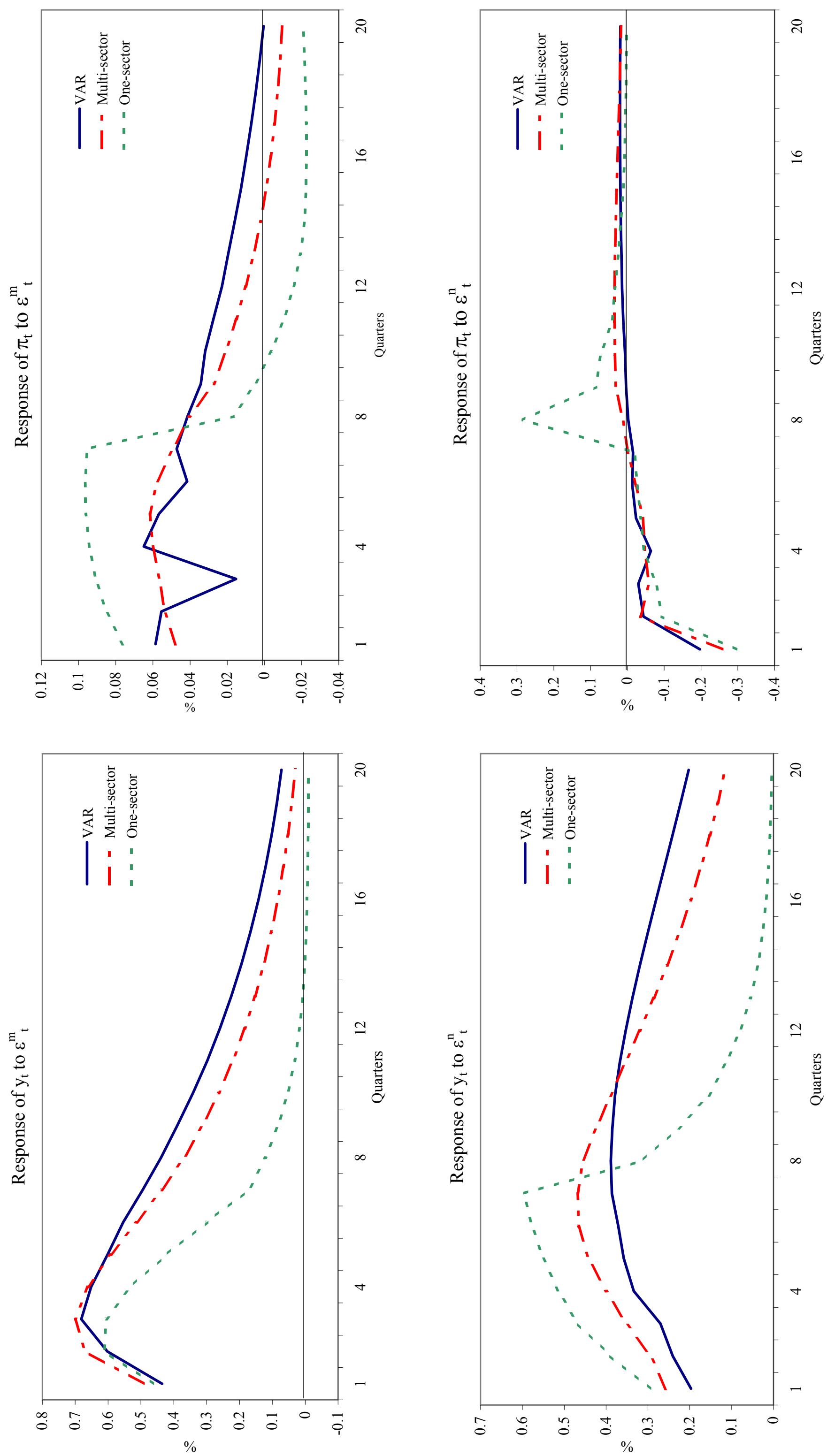

Figure 5: Impulse response functions of models and bivariate VAR - U.S.

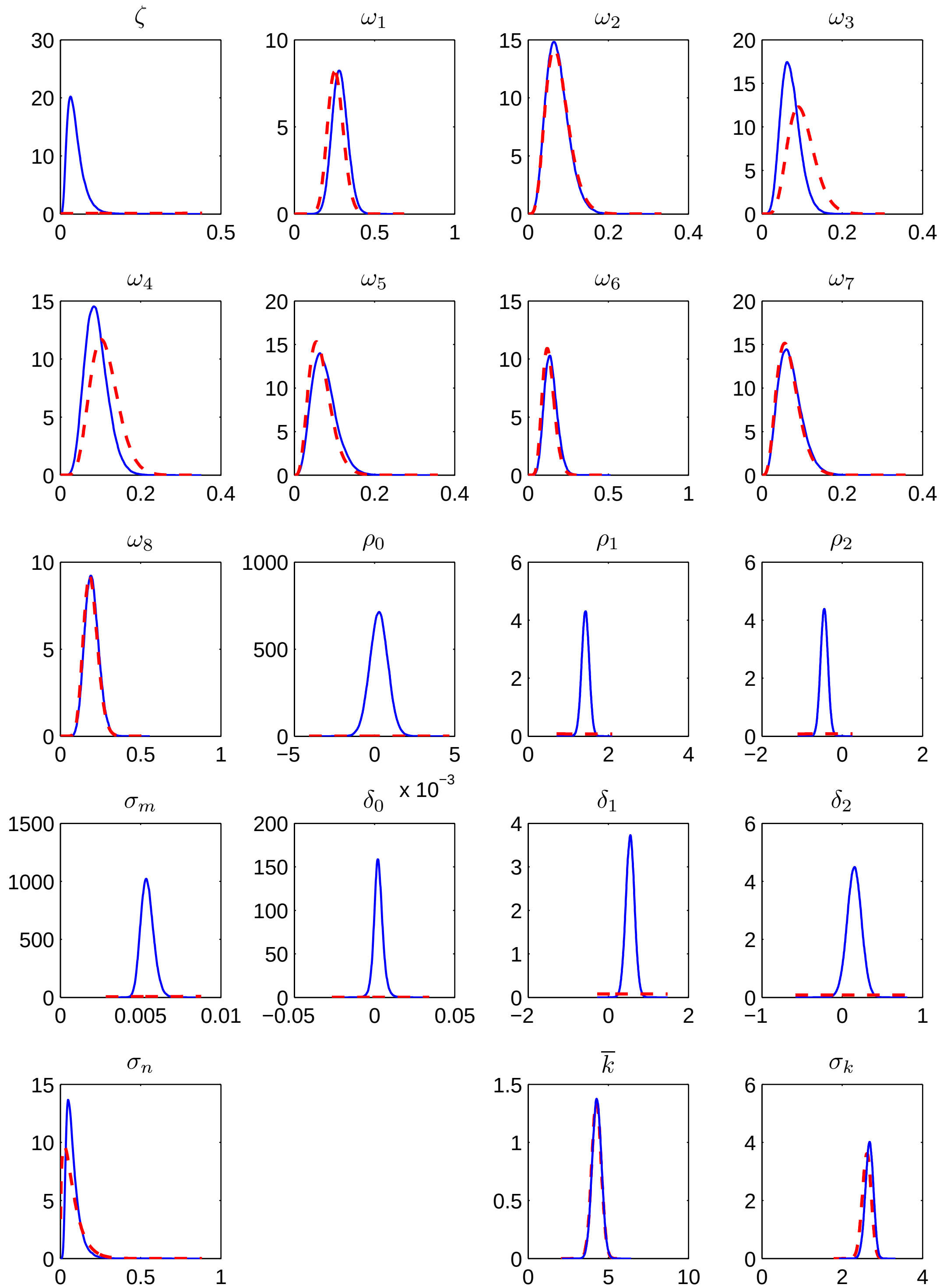

Figure 6: Marginal prior (dashed line) and posterior (solid line) distributions, K=8, informative prior - U.S.

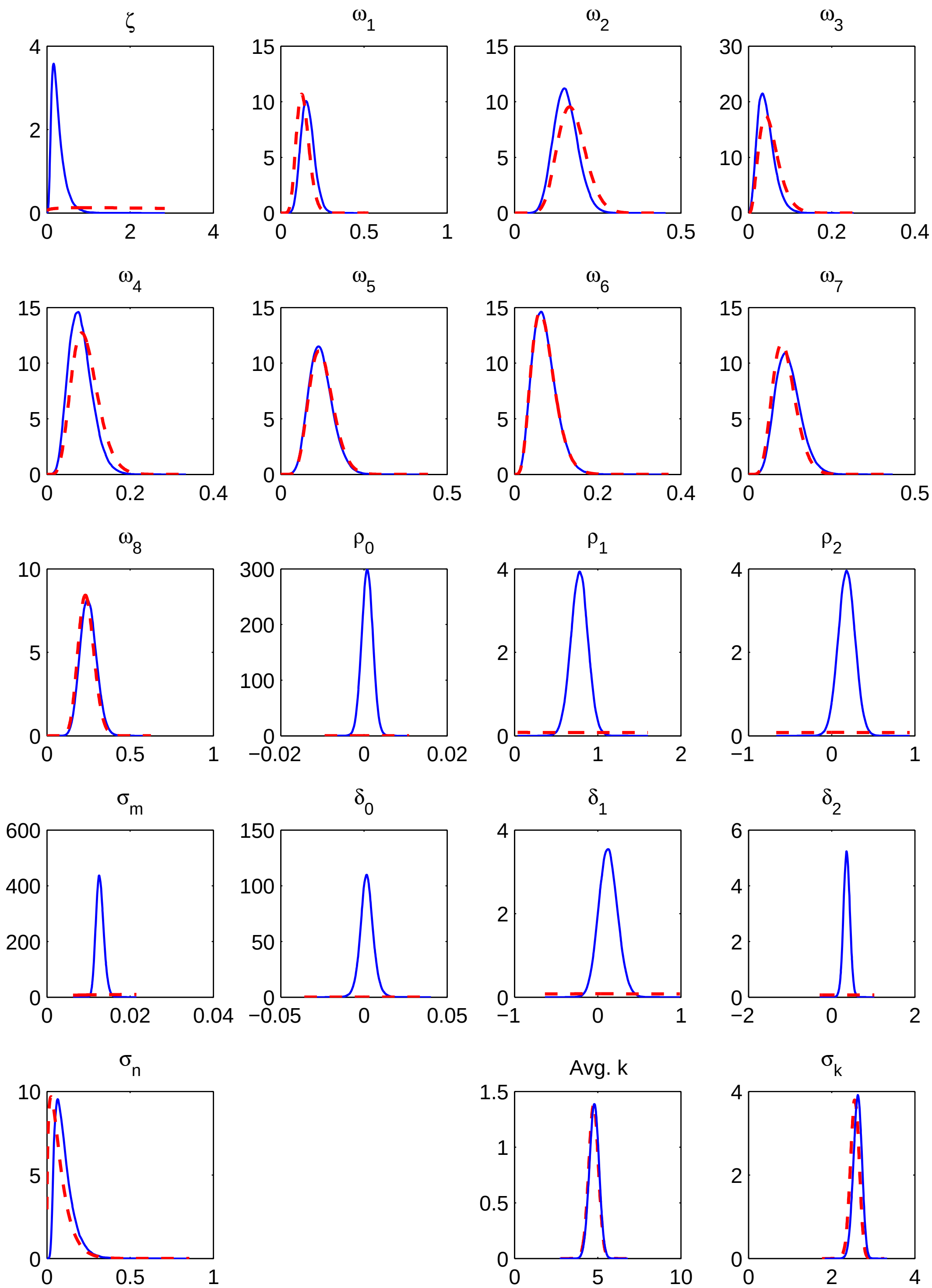

Figure 7: Marginal prior (dashed line) and posterior (solid line) distributions, K=8, informative prior - Denmark

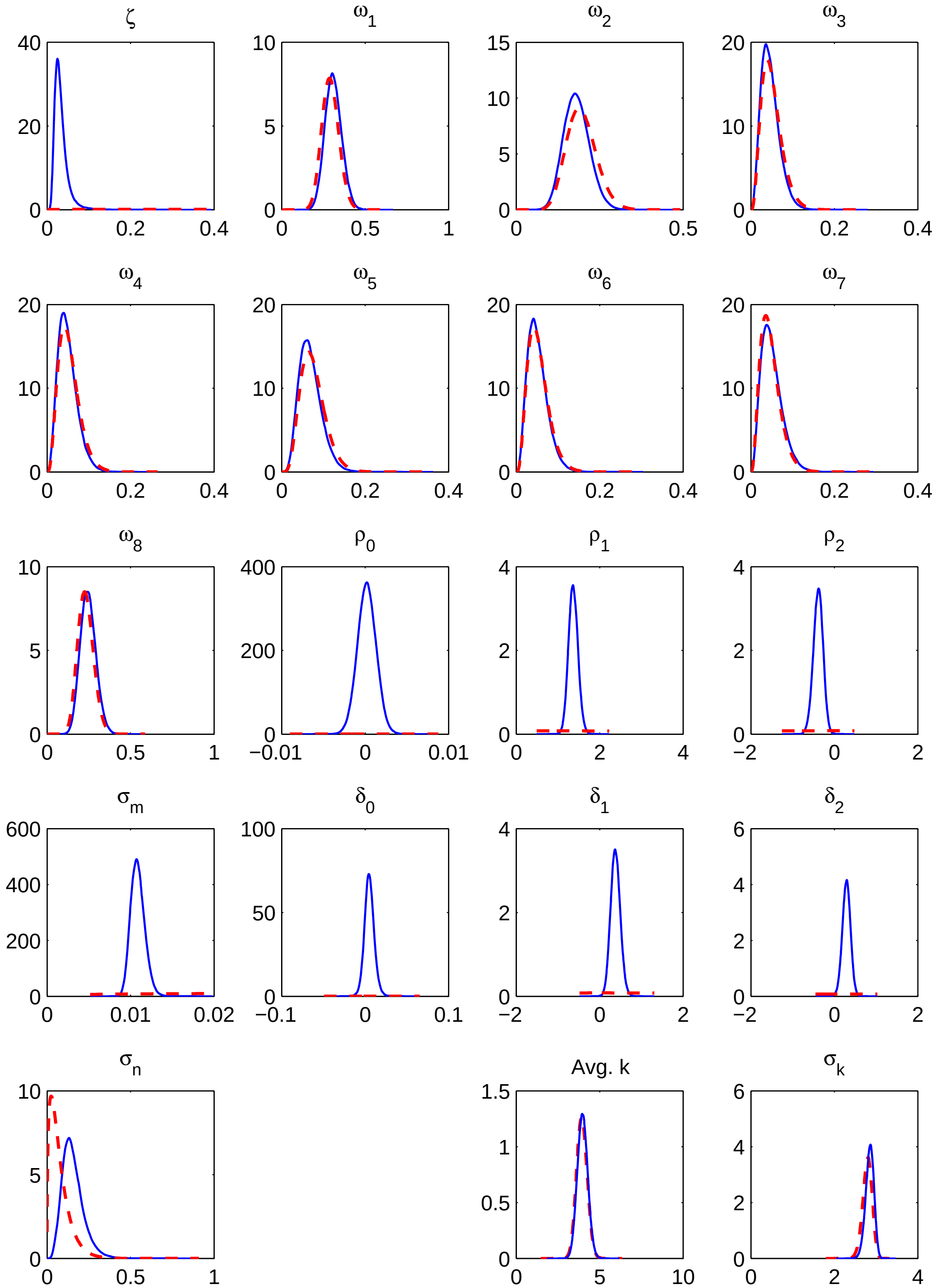

Figure 8: Marginal prior (dashed line) and posterior (solid line) distributions, K=8, informative prior - Japan