Estimating the Cross-sectional Distribution of Price Stickiness from Aggregate Data*

(COMMENTS WELCOME)

Carlos Carvalho
Federal Reserve Bank of New York

Niels Arne Dam
Danmarks Nationalbank

November, 2008

Abstract

We estimate sticky-price models for the U.S. economy in which the degree of price stickiness is allowed to vary across sectors. Perhaps surprisingly, we use only aggregate data on nominal and real output. In our models, identification of the cross-sectional distribution of price stickiness is made possible by the fact that different sectors are relatively more important in determining the response of aggregate variables to shocks *at different frequencies*. We find that the distribution of price stickiness inferred from aggregate data is strikingly similar to the distribution obtained from the recent empirical literature on microeconomic aspects of price setting in the U.S. economy. We also employ a Bayesian approach to combine time-series data on aggregate nominal and real output with such microeconomic information. Our results show that heterogeneity in price stickiness is of critical importance for understanding the joint dynamics of output and prices. Moreover, allowing for enough heterogeneity - in particular for prices in some sectors to last beyond one year - is crucial to avoid producing estimates that imply too little average nominal rigidity at the expense of too much real rigidity, relative to the specification of the model favored by the data.

JEL classification codes: E10, E30

Keywords: heterogeneity, price stickiness, micro data, macro data, Bayesian estimation

*For helpful comments and suggestions, we would like to thank Marco Del Negro, Stefano Eusepi, John Leahy, Frank Schorfheide, and seminar participants at LAMES, Danmarks Nationalbank, University of Copenhagen, Riksbank, and NY Fed. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve System, or Danmarks Nationalbank. E-mails: carlos.carvalho@ny.frb.org, nad@nationalbanken.dk.

1 Introduction

The recent empirical literature on price setting that analyzes the datasets underlying the construction of price indices documents a large amount of heterogeneity in the frequency of price changes across different economic sectors. Starting with Bils and Klenow (2004), who use data from the Bureau of Labor Statistics for the U.S. economy, shortly thereafter a large number of papers helped establish the fact that such type of heterogeneity is pervasive in modern industrial economies (e.g. Dhyne et al., 2006, and references cited therein document similar facts for the Euro area; Gagnon, 2008 details the evidence for Mexico). Inasmuch as heterogeneity in price stickiness is concerned, this recent literature corroborates and provides better measures supporting the findings of previous work on price-setting behavior (e.g. Blinder et al., 1998).

The evidence of substantial heterogeneity in the degree of price stickiness across sectors stands in sharp contrast with the assumption, common to the vast majority of papers on sticky prices, that all firms in the economy change prices with the same frequency. Apart from tractability, the only reason to resort to this assumption and not take heterogeneity explicitly into account in macroeconomic models would be if it did not matter qualitatively in aggregate terms, or at least not quantitatively. However, recent work undermines these arguments. Heterogeneity in price stickiness has been shown to affect optimal monetary policy (Aoki, 2001, and Benigno, 2004). Moreover, the extent of heterogeneity revealed by the recent microeconomic evidence has been found to have substantive implications for the dynamic response of the economy to shocks in a wide range of models with nominal rigidities.[1] In particular, in the presence of heterogeneity the response of the output gap to various shocks tends to be larger and more persistent than in otherwise identical "one-sector" economies in which all firms face the same degree of price stickiness.

The microeconomic evidence on heterogeneity in price stickiness and the theoretical results about its importance for aggregate dynamics jointly underscore the goal of this paper. We estimate a series of small multi-sector models for the U.S. economy in which the degree of price stickiness is allowed to vary across sectors. Perhaps surprisingly, we use only *aggregate* data on nominal and real output. With our models, we show that identification of the cross-sectional distribution of price stickiness is made possible by the fact that different sectors are relatively more important in determining the response of aggregate variables to shocks *at different frequencies*. This result is clear from inspection of estimates of the spectral density of inflation and output based on artificial data generated by versions of our multi-sector model with different cross-sectional distributions of price stickiness. This analysis also suggests

[1] Carvalho (2006) and Carvalho and Schwartzman (2008) cover a large class of sticky-price and sticky-information models. Nakamura and Steinsson (2007) and Enomoto (2007) study menu-cost models.

that our estimation results should allow sharp discrimination between models with heterogeneity in price stickiness and their one-sector, homogeneous-firms counterparts. We confirm the insights obtained with the spectral analysis through a small Monte-Carlo exercise, in which we estimate the parameters of a model with heterogeneity in price stickiness by maximum likelihood, using model-generated data on aggregate nominal and real output. As expected, we find that aggregate data do contain information about the cross-sectional distribution of price stickiness. For large samples (one thousand observations), the estimates are very close to the true parameter values, and the standard errors are small. However, for samples of the same size as our data, we find the estimates to be moderately biased and somewhat less precise.

Drawing upon our Monte Carlo analysis, to estimate our models we use a full-information Bayesian approach that allows us to combine recently available information about the cross-sectional distribution of price stickiness derived from microeconomic data with time series of aggregate nominal and real output. We focus on two complementary questions. First, we assess how estimates of the cross-sectional distribution of the frequency of price changes obtained *without* use of the microeconomic evidence compares with such evidence. To that end, we estimate the model under an uninformative ("flat") prior on the cross-sectional distribution of price stickiness, and compare the results with empirical cross-sectional distributions of the frequency of price changes derived from Bils and Klenow (2004) and Nakamura and Steinsson (2008). Second, we perform formal statistical model comparisons to assess whether the aggregate data favor specifications with heterogeneity in price stickiness over alternatives with homogenous price-setting behavior. To answer that question we estimate, in addition, versions of the model that incorporate information from the microeconomic data through different sets of priors on the cross-sectional distribution of price stickiness.[2]

Under uninformative priors on the cross-sectional distribution of price stickiness, we find that the distribution inferred from aggregate data, conditional on our model, is strikingly similar to the distribution obtained from the recent empirical literature that uses only microeconomic data. When we restrict the maximum duration of price stickiness to one year, the distribution implied by various point estimates closely tracks the empirical distribution constructed from Bils and Klenow (2004). At the same time, the results point to an incredibly large degree of real rigidity (in the sense of Ball and Romer, 1990), which renders the adjustment process of the output gap to shocks sluggish despite the relatively low levels of nominal price rigidity. In contrast, when we allow prices in some sectors to last beyond one year (up to a maximum of either one and a half, or two years), the resulting cross-

[2] Another promising approach, which we do not explore in this paper, involves augmenting the model with sectoral shocks, and using sectoral data as well. See Lee (2007) and Boukaez et al (2007).

sectional distribution becomes more similar to the empirical distribution derived from Nakamura and Steinsson (2008). Moreover, in this case the degree of real rigidity falls somewhat, and approaches levels that are arguably consistent with those implied by plausible calibrations of structural parameters in fully-specified macroeconomic models.

Despite very different empirical methodologies, our results under the restriction of a maximum of one year of price rigidity are in line with those obtained by Coenen et al. (2007), who estimate a model with Taylor-staggered price setting and heterogeneous contract lengths of up to four quarters, using indirect-inference methods. They find a distribution of price spells that resembles the evidence documented by Bils and Klenow (2004), combined with an incredibly large degree of real rigidities. However, we find that the restriction of at most one year of price rigidity comes at significant costs for the empirical validity of the model. This is true irrespective of whether the prior on the cross-sectional distribution of price stickiness is informative or not. Specifically, in the comparison between the specifications with at most one year price spells, and the models with up to 6 or 8 quarters of price rigidity, the posterior odds ratios favor the latter by an order of magnitude that ranges from $10^3 : 1$ to $10^5 : 1$, depending on the specification.

After learning that the evidence allows such a clear distinction between models with more of less heterogeneity in price stickiness, it comes as no surprise that it supports an even sharper discrimination between models with heterogeneity in price stickiness and their homogeneous-firms counterparts. In the comparisons between the various specifications of these models, the results strongly support the versions with heterogeneity in price stickiness. As an example, picking the *worst* heterogeneous model and the *best* homogeneous model, the posterior odds in favor of the heterogeneous model is of the order of $10^5 : 1$. 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.[3]

To summarize, our results provide strong evidence in favor of models with heterogeneity in price stickiness. In terms of formal statistical model comparisons, these receive overwhelming support from the data, relative to their homogeneous-firms counterparts. Moreover, the estimates that allow for heterogeneity appear more reasonable in economic terms, since they match not only the average degree of price rigidity observed in the micro-data, but also the whole cross-sectional distribution of price stickiness. Finally, allowing for *enough* heterogeneity - namely, for prices in some sectors to last beyond one year - is also important: neglecting to do so leads to estimates that imply too little nominal rigidity on the one hand, compensated by too much real rigidity on the other, relative to the specifications

[3]Note that this result is in line with estimates of price stickiness obtained with one-sector DSGE models, such as Smets and Wouters (2003).

favored by the data.

In constructing our models, we derive the structure of the supply side from 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 (the demand side, if you will), we assume exogenous stochastic processes for nominal output and for an unobservable natural rate of output; hence, we refer to our models as "semi-structural".[4] Given that our focus is 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 some flexibility relative to a fully-structural model. This approach allows us to avoid having to find a demand-side specification that performs well empirically, which typically requires the introduction of various "frictions" that would move us away from the focus of our paper.

The recent literature that focuses on the implications of heterogeneity in the extent of nominal rigidities continues to grow. Dixon and Kara (2005) study sectoral heterogeneity in wage setting. Carlstrom et al. (2006a,b) analyze the effects of heterogeneity in price stickiness on equilibrium determinacy and aggregate dynamics in two-sector models. Nakamura and Steinsson (2007) and Enomoto (2007) study multi-sector menu-cost models. Sheedy (2007) focuses on how heterogeneity in price stickiness affects inflation persistence. 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. Barsky et al. (2007) study a two-sector model with durable consumption goods and heterogeneity in the frequency of price changes. Carvalho and Schwartzman (2008) show how heterogeneity in price-setting behavior affects aggregate dynamics in a large class of sticky-price and sticky-information models. Carvalho and Nechio (2008) show that an open-economy model with heterogeneity in price stickiness can account for the sluggish dynamics of real exchange rates observed in the data, whereas a one-sector version of the model with the same average degree of price stickiness fails to do so. Lee (2007) and Boukaez et al. (2008) estimate multi-sector DSGE models with heterogeneity in price rigidity using sectoral data. Jadresic (1999) and Coenen et al. (2007) estimate models with heterogeneity in price stickness using only aggregate data.

We proceed as follows. Section 2 presents the semi-structural model. Section 3 studies the extent to which aggregate data contains information about the cross-sectional distribution of price stickiness. Section 4 describes our empirical methodology. We detail the aggregate data used in the estimation,

[4] 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.

as well as our method for incorporating the information from the micro price data in our priors. This section also details our prior assumptions, and our particular Bayesian estimation algorithm. We follow with the analysis of our results in Section 5. We analyze the findings under informative and uninformative, study multi-sector models with different limits on the extent of price rigidity, and compare these with homonegeous-firms specifications. We conclude in the last section.

2 The semi-structural model

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

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

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

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:

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

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.[5]

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

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

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

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

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. Firms meet all demand for their products at the posted prices. 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:

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

where $Q_{t,t+i}$ is a (possibly stochastic) nominal discount factor. In this context, 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 $X_t(k)$. 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:

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

We close the model by specifying a stochastic process for the shock ξ_t, and by positing that nominal output $M_t \equiv P_t Y_t$ also evolves in an exogenous fashion. This is a standard assumption in theoretical work on price setting (e.g. Mankiw and Reis, 2002, or Woodford, 2003, chapter 3). It allows us to focus on the implications of our particular model of price setting for the decomposition of changes in nominal output into purely real and purely nominal effects, without having to specify a full model of the economy.

2.1 A loglinear approximation

We assume that the economy has a deterministic zero inflation steady state characterized by $\xi_{t+i} = 0, Y_{t+i} = \overline{Y}, Q_{t,t+i} = \beta$, and for all (k,j), $X_t(k,j) = P_t = \overline{P}$, and loglinearize (1) around it to obtain:[6]

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

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 the level of output that would prevail in a flexible-price economy. It is referred to as the natural level of output, and 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, the natural output level moves *pari passu* with $\log(\xi_t)$: $y_t^n \propto \log(\xi_t)$.

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:[7]

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

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

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

[6] We write all such expressions as equalities, ignoring higher order terms.

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

whether there is an explicit input-output structure etc.[8]

In the context of our model, the index itself can be regarded as a primitive parameter. 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:

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

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, *ceteris paribus*. 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 output m_t and natural output level y_t^n

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

$$m_t = \rho_0 + \rho_1 m_{t-1} + ... + \rho_{p1} m_{t-p_1} + \varepsilon_t^m, \tag{6}$$

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

$$y_t^n = \delta_0 + \delta_1 y_{t-1}^n + ... + \delta_{p2} y_{t-p_2}^n + \varepsilon_t^n, \tag{7}$$

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

2.3 State-space representation and likelihood function

The semi-structural model then consists of equations (2) through (7). We write it in state-space form in the notation of Sims (2002):

$$\Gamma_0 Z_t = \Gamma_1 Z_{t-1} + \mathcal{C} + \Psi \varepsilon_t + \Xi \eta_t,$$

where Z_t is a vector collecting all variables and additional "dummy" variables created to account for leads and lags, $\mathcal{C}$ is a vector of constants, ε_t is as defined before, and η_t is a vector of one period

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

ahead expectational errors. Γ_0, Γ_1, Ψ, and Ξ are the appropriate matrices, which are functions of the primitive parameters of the model, collected for notational simplicity in a vector:

$$\theta = \begin{pmatrix} K & p_1 & p_2 & \beta & \zeta & \sigma_m^2 & \sigma_n^2 & \omega_1 & \cdots & \omega_K & \rho_0 & \cdots & \rho_{p_1} & \delta_0 & \cdots & \delta_{p_2} \end{pmatrix}.$$

We solve the model with Gensys (Sims, 2002), to obtain:

$$Z_t = C(\theta) + G_1(\theta) Z_{t-1} + B(\theta) \varepsilon_t. \tag{8}$$

In all estimations that follow, we make use of a 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 log-linear model (8). 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:

$$\begin{aligned} Z_t|Z_{t-1} &\sim N\left(C(\theta) + G_1(\theta) Z_{t-1}, B(\theta)\Omega B(\theta)'\right), \\ Z_t^*|\{Z_\tau^*\}_{\tau=1}^{t-1} &\sim N(M(\theta), V(\theta)), \end{aligned} \tag{9}$$

where $M(\theta) \equiv HC(\theta) + HG_1(\theta)\hat{Z}_{t|t-1}$, $V(\theta) \equiv HB(\theta)\hat{\Sigma}_{t|t-1}B(\theta)'H'$, $\hat{Z}_{t|t-1}$ 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.

3 Informational content of the aggregate data

In estimating our multi-sector models we only use data on aggregate nominal and real output. Although we use a Bayesian approach in our empirical implementation (detailed subsequently), it is natural to ask whether the aggregate data do contain usefull information about the cross-sectional distribution of price stickiness. Our goal in this section is to provide some insight into this question through a small Monte-Carlo exercise.

As a first step, we illustrate how different cross-sectional distributions of price stickiness affect aggregate dynamics in a very stylized setting. We shut down shocks to the natural rate of output ($\sigma_n = 0$), and assume that nominal income follows a random-walk $\left(\rho_1 = 1, \rho_2 = ... = \rho_{p_1} = 0\right)$. We also fix $\zeta = 0.1$, and $\beta = 0.99$. Then, we present estimates of the (log) spectral density of inflation based on artificial data generated under different cross-sectional distributions of price stickiness.[9] The results are presented in Figure 1.[10] The top row presents results for simple heterogeneous economies in which half of the prices are flexible ($\omega_1 = 0.5$), and the other half of firms change prices every k quarters. The bottom row shows results for one-sector, homogeneous-firms models with duration of price spells ranging from one year ($\omega_4 = 1$) to two years ($\omega_8 = 1$). For each specification, the dashed red lines indicate particular frequencies. From left to right, for a model with price spells of up to k quarters, they correspond to $k/2$, $k/4$, $k/6$ quarters, and so on. For example, for $k = 8$, the lines correspond to frequencies of (from left to right) 4, 2, 1.333, and 1 quarter(s).s

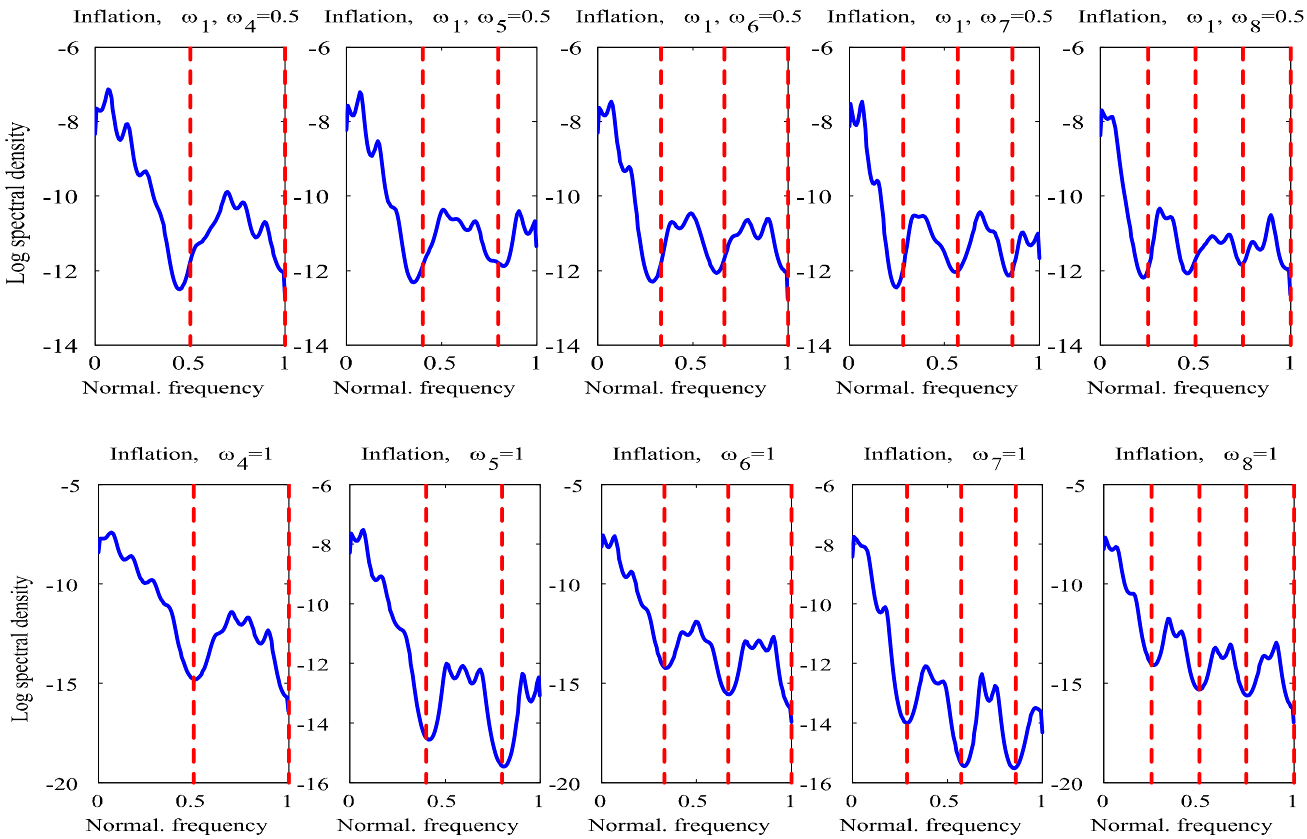

Figure 1: Spectral densities of inflation in models with different distributions of price stickiness.

The first noteworthy fact is that the spectral densities of inflation present very strong patterns that differ across specifications. This is good news when it comes to discriminating between the various alternatives. The second fact is that specifications with homogeneous and heterogeneous firms

[9] Although we estimate the models on nominal and real output, we present the spectral densities for inflation rates for convenience. Since our two observables and the price level are linearly related, they contain the same information that can be explored by our full-information methods.

[10] The estimates are obtained by fitting an AR(20) to the series, and computing the implied spectral density. The sample size is one thousand observations, which corresponds to the large sample in our subsequent Monte-Carlo analysis. Non-parametric methods deliver similar results.

that have a common average duration of price spells also present different patterns. For instance, the specifications with $\omega_4 = 1$ (bottom row, first from left to right) and $\omega_1, \omega_7 = 0.5$ (top row, second from right to left): both have an average duration of price spells of one year. Yet, the latter appears closer to the specification with $\omega_7 = 1$ than to the one with $\omega_4 = 1$. This is a manifestation of what Carvalho (2006) refers to as the *strategic interaction effect*: in the presence of pricing complementarities, lower frequencies of price adjustment have a disproportionate aggregate effect. However, the pattern of the spectral density remains different from that of the homogeneous-firms economy with $k = 7$, in that the pronounced drops in the spectral density do not happen at the exact same frequencies, and the magnitudes of the fluctuations in the spectral densities are different.

These differences in the patterns of the spectral densities are a manifestation of what Carvalho (2006) termed the *frequency composition effect*. It allows inference about the distribution of sectoral weights at the different frequencies that correspond to different durations of price spells from the behavior of aggregates.[11] This inference can, of course, be made harder or easier, depending on several factors. For instance, stronger pricing complementarities make it harder to to discriminate between a model with $\omega_1 = \omega_k = 0.5$ and one with $\omega_k = 1$. Smaller samples, of course, also make the differences on which this inference relies less stark.

Having provided some insight into the effects of heterogeneity in price stickiness that allow for cross-sectional inference from aggregates, we now try to shed light on the reliability of such inference in an empirically more realistic setting. We generate artificial data on aggregate nominal and real output using a model with $K = 4$, and parameter values that resemble what we find when we estimate the model on actual data. Then, we estimate the parameters of the model by maximum likelihood.[12] We conduct both a large sample (one thousand observations) and a small sample exercise (samples of the same size as our data, i.e. 100 observations).

Table 1 reports the true parameter values used to generate the data in the first column. 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.[13] The "Ini. guess" column reports the average value of the initial guesses supplied for the optimization algorithm across the corresponding samples. Figures 2 and 3 present histograms of the point estimates

[11] In a related, albeit different context, Zaffaroni (2004) shows that under some semi-parametric assumptions on the cross-sectional distribution of micro-parameters, as the number of sectors increases one should be able to obtain information about such distribution by observing only the aggregate (average) of the microeconomic units.

[12] 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 4.4).

[13] 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.

for each parameter obtained with the artificial samples, for the "Large sample" and "Small sample" Monte-Carlo exercises, respectivelly.[14] They also indicate the true values, the sample means, and the (average) values of the initial guess.

Table 1: Monte-Carlo - Maximum Likelihood estimation

		Large sample				*Small sample*			
	True	*Mean*	*5th perc.*	*95th perc.*	*Ini. guess*	*Mean*	*5th perc.*	*95th 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

Note: The columns under "Large sample" report statistics across 75 artificial samples of 1000 observations. The "Small sample" columns report statistics across 240 artificial samples of 100 observations. The "Ini. guess" column reports the average value of the initial guesses across the corresponding samples. Following the procedure that we use in the actual estimation algorithm, the values for ζ and $\omega_1 - \omega_4$ are fixed across replications; the values for the remaining parameters in each replication are set equal to the ordinary least squares estimates based on nominal income (for the ρ's) and *actual* output (for the δ's).

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

[14] We omit ρ_0 and δ_0, for which the distributions are essentially a point mass at zero.

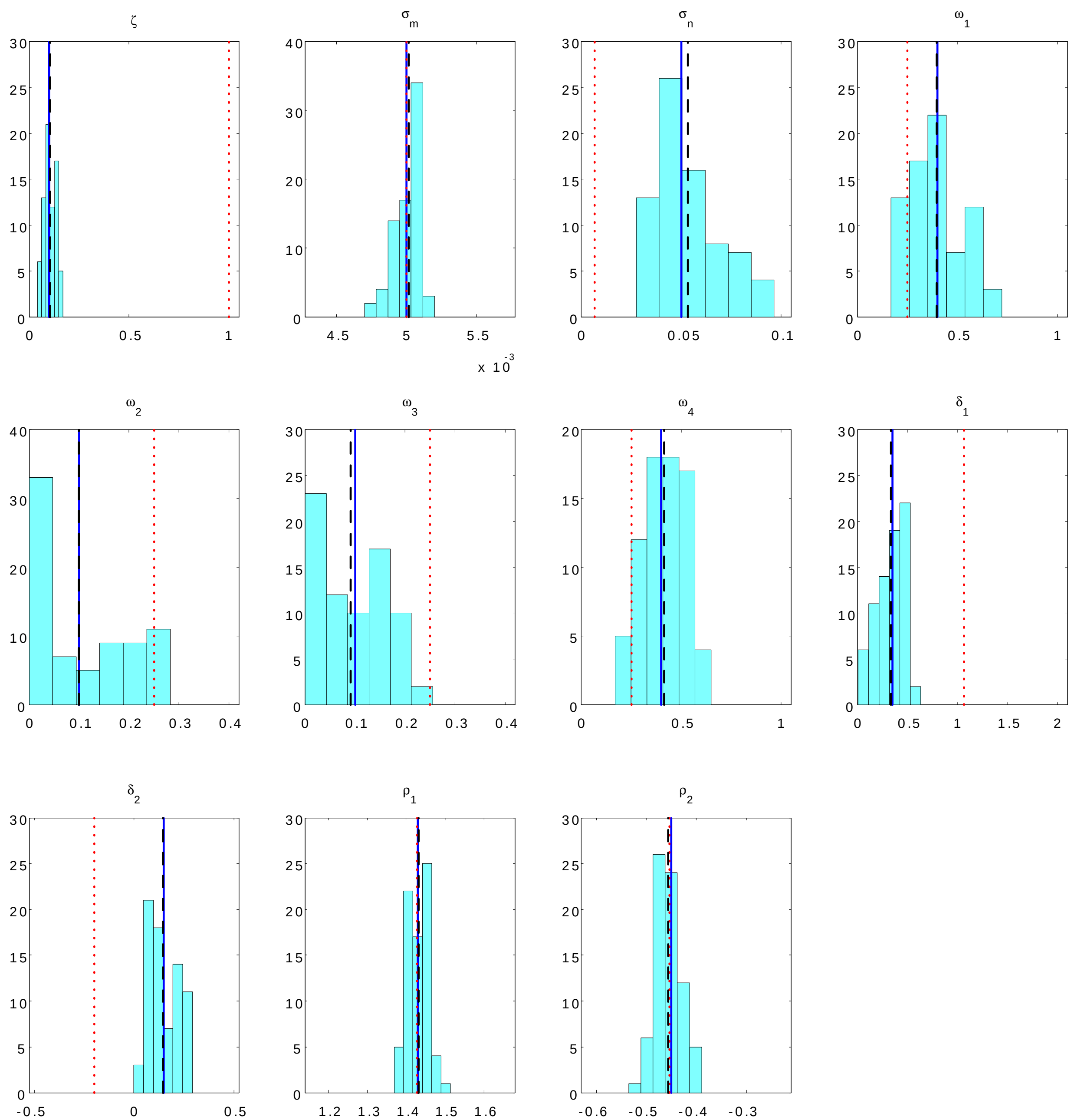

Figure 2: Monte-Carlo - maximum likelihood estimation, large samples. The solid (blue) line indicates the true value, the dashed (black) line indicates the average across replications, and the dotted (red) line indicates the initial value used in the maximum likelihood estimation.

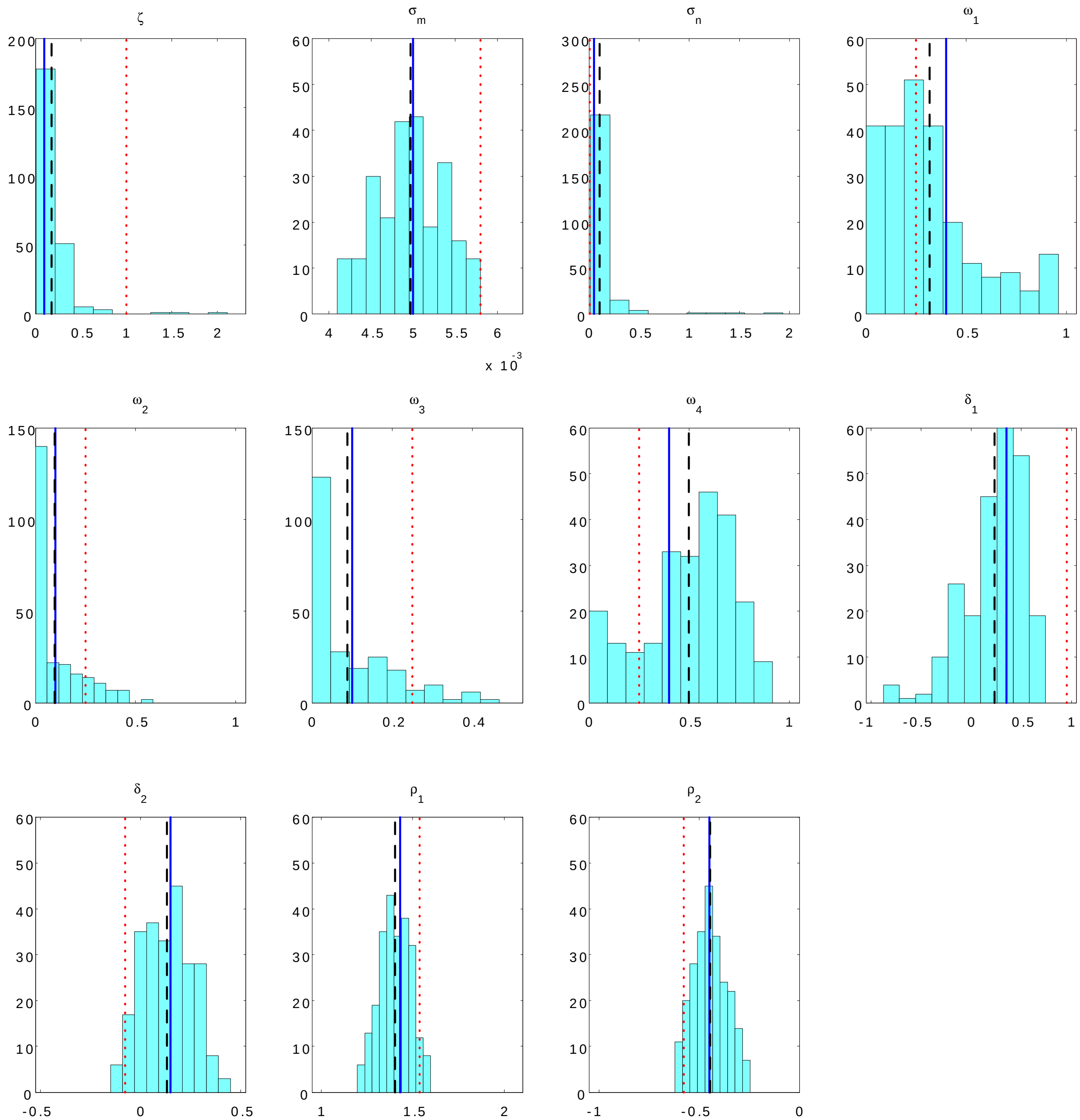

Figure 3: Monte-Carlo - maximum likelihood estimation, small samples. The solid (blue) line indicates the true value, the dashed (black) line indicates the average across replications, and the dotted (red) line indicates the initial value used in the maximum likelihood estimation.

4 Empirical Methodology and Data

With the challenges involved in bridging the gap between microeconomic information on firms' pricing behavior and the National Account time-series on U.S. real and nominal output, the choice of empirical methodology is of critical importance. We employ a Bayesian approach as this allows us to integrate the microeconomic information in the Bureau of Labor Statistics (BLS) price data with macroeconomic time-series. 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 data in two complementary ways. In one direction, we estimate the cross-sectional distribution of price stickiness using time-series data on aggregate nominal and real output, under a prior distribution that is flat (uninformative) over the space of sectoral weights. We then compare the results with the recent microeconomic evidence documented by Bils and Klenow (2004) and Nakamura and Steinsson (2008). Alternatively, we incorporate the information from the microeconomic data analyzed in these papers through our prior on the cross-sectional distribution of the frequency of price changes, and estimate the model using, again, aggregate nominal and real output as observables. The use of the microeconomic findings from these papers is discussed in Subsection 4.2 below. While we could have opted to integrate the results of the two papers into one set of priors, we chose to utilize them in separate estimations. The reason is that it allows us to investigate how the profile of the cross-sectional distribution affects the estimated aggregate dynamics, and thus gain information on which of the two distributions seems more in line with the observed behavior of aggregate output and prices, and whether they improve on the estimation obtained without the use of microeconomic evidence through the priors.

When estimating the model, whether using priors based on the microeconomic evidence or not, we estimate several specifications of the model for different values of K, treating it as fixed in each estimation. We choose this approach over the alternative of specifying a prior for this parameter as well and estimating its distribution due to computational constraints.

In the next subsection we provide details on the macroeconomic time-series that constitute the observables in our Bayesian estimation. We then explain how we use the statistics reported by Bils and Klenow (2004) and Nakamura and Steinsson (2008) to construct empirical cross-sectional distributions of price stickiness. This is followed by a specification of our prior assumptions, with particular emphasis on how we use the cross-sectional information from those two papers. Finally, we detail the algorithm that we use to simulate the joint posterior distribution of the parameters.

4.1 Macroeconomic time-series

The model laid out is tested against the development of quarterly nominal and real output. These are measured as seasonally adjusted gross domestic product at, respectively, current and constant prices. We take natural logarithms and remove a linear trend from the data. Whereas the assumptions underlying the model include one of an unchanged economic environment, the U.S. economy has undergone profound changes in the recent decades, including phenomena such as the so-called "Great Moderation" and the Volcker Disinflation. As a consequence, we do not attempt to confront the model with the full sample of post-war data. We use the period from 1979 to 1982 only as a pre-sample, and thus exclude the monetary regimes up to and including the Federal Reserve money-supply targeting policy of 1979-1982 from the developments assessed by the likelihood criterion. Ultimately we evaluate the model according to its ability to match business cycle developments in nominal and real output post 1982. 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. In running the Kalman filter, we use the parameters values in each draw.[15] Our effective estimation sample is thus 1983-2007. The time-series are depicted in Figures 4 and 5.

4.2 Empirical cross-sectional distributions of price stickiness

We extract our information about the cross-sectional distribution of price stickiness from two recent papers that 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 the Consumer Price Index. The first study is Bils and Klenow (2004, henceforth BK), who pioneered the use of such data to analyze the frequency of price changes and other aspects of price setting. The other is Nakamura and Steinsson (2008, henceforth NS).

The main difference between these two papers is that BK analyze the BLS data at an intermediate level of disaggregation,[16] whereas NS use the data at its most disaggregated level.[17] Moreover, NS provide statistics with different treatments of sales price observations and product substitutions. When

[15] For the initial condition for the pre-sample, we use the unconditional mean and a large variance-covariance matrix.

[16] They used data on so called "entry level items," from the Commodities and Services Substitution Rate Tables for the period 1995-1997.

[17] Their data comes from the CPI Research Database, and covers the period 1988-2005.

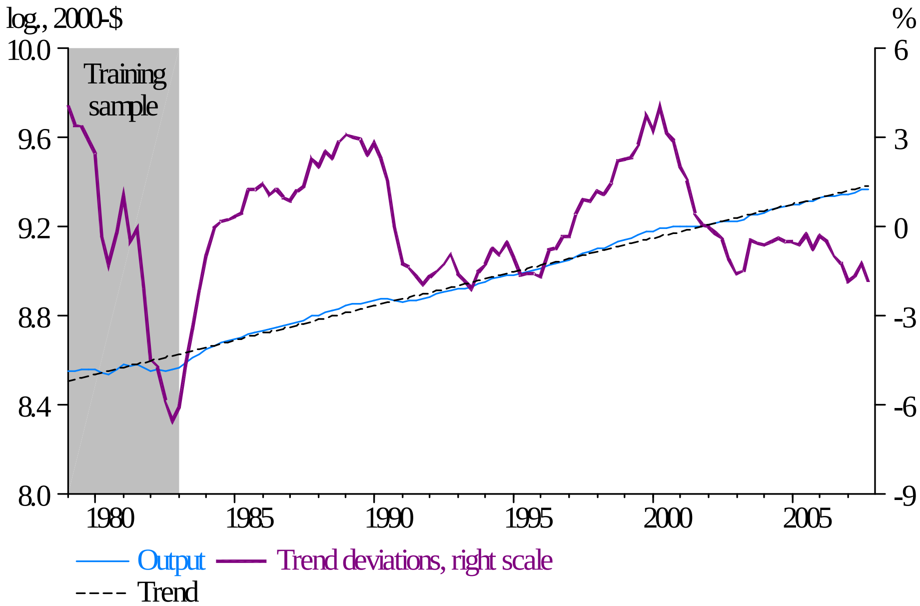

Figure 4: Real GDP ("output")

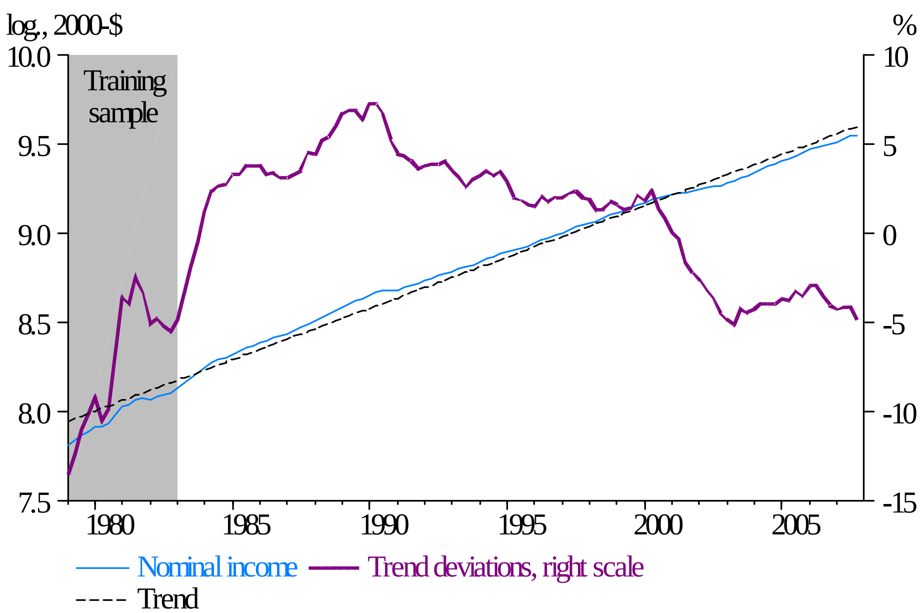

Figure 5: Nominal GDP ("nominal income")

they apply treatments for these events that make their sample close to being a more disaggregated version of the data analyzed by BK, the two papers produce very similar results on the cross-sectional distribution of the frequency of price changes. In contrast, when NS use the observations on what they refer to as *regular* price changes, they find more nominal price rigidity and more dispersion in the frequency of price changes across different goods.[18]

We work with the statistics on the frequency of price changes for the 350 categories of goods and services ("entry level items") analyzed by BK, and with the 272 entry level items covered by NS (using the statistics for regular prices). Our goal is to map those statistics into an empirical distribution of sectoral weights over spells of price rigidity with different durations. We work at a quarterly frequency, and consider economies with at most 8 quarters of price stickiness. Thus, we consider price spells which are multiples of one quarter, and for each of the BK and NS data, we aggregate the goods and services categories so that the ones which have an average duration of price spells between zero and one quarter (inclusive) are assigned to the first sector; the ones with an expected duration of price spells 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 CPI expenditure weights. We proceed in this fashion until the sector with 7-quarter price spells. Finally, we aggregate all the remaining categories, which have mean durations of price rigidity of 8 quarters and beyond, into a sector with 2-year price spells. This gives rise to the empirical cross-sectional distributions of price stickiness that we use in our estimation, which are summarized in Table 2. We denote the sectoral weight for sector k obtained from this procedure by $\widehat{\omega}_k$. For each of the BK and NS distributions, we also compute the average duration of price spells, $\widehat{\overline{k}} = \sum_{k=1}^{K} \widehat{\omega}_k k$, and the standard deviation of the underlying distribution, $\widehat{\sigma}_k = \sqrt{\sum_{k=1}^{K} \widehat{\omega}_k \left(k - \widehat{\overline{k}}\right)^2}$.

4.3 Assumptions on the prior distributions

The model parameters of interest essentially fall into four categories that are dealt with in turn. 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 2 for both processes, i.e. $p_1 = p_2 = 2$. The choice of lag lengths was based on likelihood ratio tests and the Schwarz and Akaike information criteria from univariate estimations on the pre-sample time-series (for this purpose, we use actual real output as a proxy for the unobserved natural output

[18] We refer the reader to the two papers for a detailed description of their methodologies.

Table 2: Cross-sectional distributions of price stickiness

Parameter	BK	NS
$\widehat{\omega}_1$	0.40	0.27
$\widehat{\omega}_2$	0.24	0.07
$\widehat{\omega}_3$	0.12	0.10
$\widehat{\omega}_4$	0.12	0.11
$\widehat{\omega}_5$	0.04	0.06
$\widehat{\omega}_6$	0.03	0.13
$\widehat{\omega}_7$	0.03	0.06
$\widehat{\omega}_8$	0.03	0.20
$\widehat{\overline{k}}^{(*)}$	2.54	4.23
$\widehat{\sigma}_k^{(*)}$	1.86	2.66

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

process).[19] The second set of parameters consists of the standard deviations of the shocks to nominal output (σ_m) and natural output (σ_n). These are strictly positive parameters to which we assign loose gamma priors. The third block of parameters includes only the output-gap elasticity of price setting, also known as 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 significant degree of strategic complementarity or substitutability in price setting should be a feature of the estimation rather than of our prior assumptions. The priors for this first set of parameters are summarized in Table 3.[20]

Table 3: Prior distributions

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

Note: The hyper-parameters for the gamma distribtution specify shape and inverse scale, respectively, cf. Gelman et al. (2003).

The fourth and final set of parameters are the sectoral weights $\omega = (\omega_1, ..., \omega_K)$. The combined

[19] In principle we could have specified priors over p_1, p_2 and estimated their posterior distribution as well. However, the computational challenge of the set of model configurations analyzed in this paper is already tremendous given our implementation of the estimation algorithm. Therefore, we restrict ourselves to this specification.

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

restrictions of non-negativity and summation to unity of the ω's are captured through a Dirichlet distribution, which is a multivariate generalization of the beta distribution. Notationally, $\omega \sim D(\alpha)$ with density function:

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

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] It follows that for a given K, $\alpha_0 \equiv \sum_k \alpha_k$ captures the overall level of information in the prior distribution. Thus, we control the tightness of our sectoral mass priors through α_0/K, where the normalization by K makes this measure more comparable across models with different number of sectors. For a given level of information in the prior distribution of sectoral weights, 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).

For a given number of sectors K and a given value for α_0, we use the empirical cross-sectional distributions derived from BK and NS described in the previous subsection to assign values for the hyperparameters $\alpha_1, ..., \alpha_K$. We do this in a way that sets the mode of the distribution of sectoral weights equal to the empirical distribution implied by the $\widehat{\omega}_k$'s. For specifications with $K = 8$, we simply set $\alpha_k = 1 + \widehat{\omega}_k(\alpha_0 - K)$. For any specification with $K < 8$, we set $\alpha_k = 1 + \widehat{\omega}_k(\alpha_0 - K)$ for all $k \leq K - 1$, and $\alpha_K = 1 + \left(\sum_{j=K}^{8} \widehat{\omega}_j\right)(\alpha_0 - K)$. The $\widehat{\omega}_k$'s being taken from the empirical distribution derived from BK gives rise to what we refer to as a *Bils-Klenow prior*, whereas taking the $\widehat{\omega}_k$'s from the corresponding distribution based on NS yields a *Nakamura-Steinsson prior*.

In addition to using the BK and NS priors, we also estimate versions of the model with 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 assess the information that the aggregate data contains about the cross-sectional distribution of price stickiness.

Generally, we shall refer to cases with $\alpha_0 = 2K$ as *loose priors*. We also consider a tighter set of priors. For each configuration of the model (i.e. the number of sectors K, and whether the prior is

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

BK or NS) the *tight priors* are defined by setting $\alpha_0 = 10K$. Figure 6 illustrates how the marginal distribution varies with α_0 when all α_k's are equal. For $K = 4$, it shows the flat ($\alpha_0 = 4$), loose ($\alpha_0 = 8$), and tight ($\alpha_0 = 40$) priors. Figure 7 compares the loose and tight priors of the BK and NS variants for the case of four sectors.

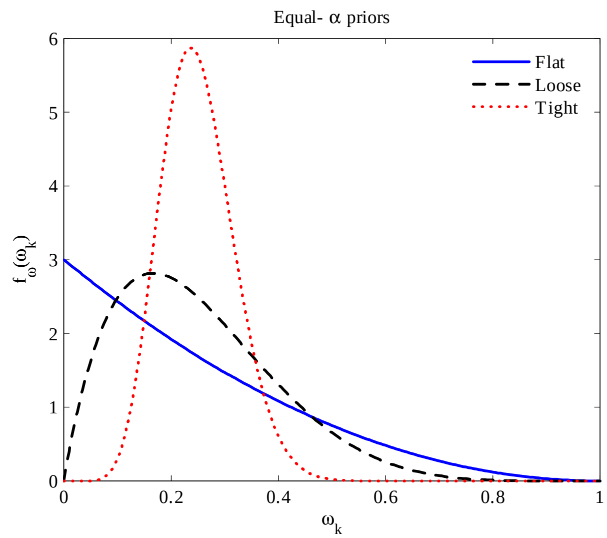

Figure 6: Equal-α sectoral mass marginal priors with different tightness, $K = 4$.

4.4 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 parameter set that converges to the joint posterior distribution under certain conditions.[22] The numerous restrictions and sizeable dimension of the parameter set, especially in relation to our data set consisting of only two aggregate time-series, place high demands on the exact implementation of the algorithm in order to obtain efficiency and convergence of the Markov chain within a manageable number of iterations.

Our specific estimation strategy for each configuration of the model is as follows. We run two numerical optimization routines sequentially in order to determine the starting point of the Markov chain and gain a first crude estimate of the covariance matrix for our Random-Walk Metropolis Gaussian jumping distribution.[23] Before running the Markov chains we transform all parameters to

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

[23] The first optimization routine is `csminwel` by Chris Sims, while the second is `fsminsearch` 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 income (for the ρ's) and *actual* output (for the δ's). Following

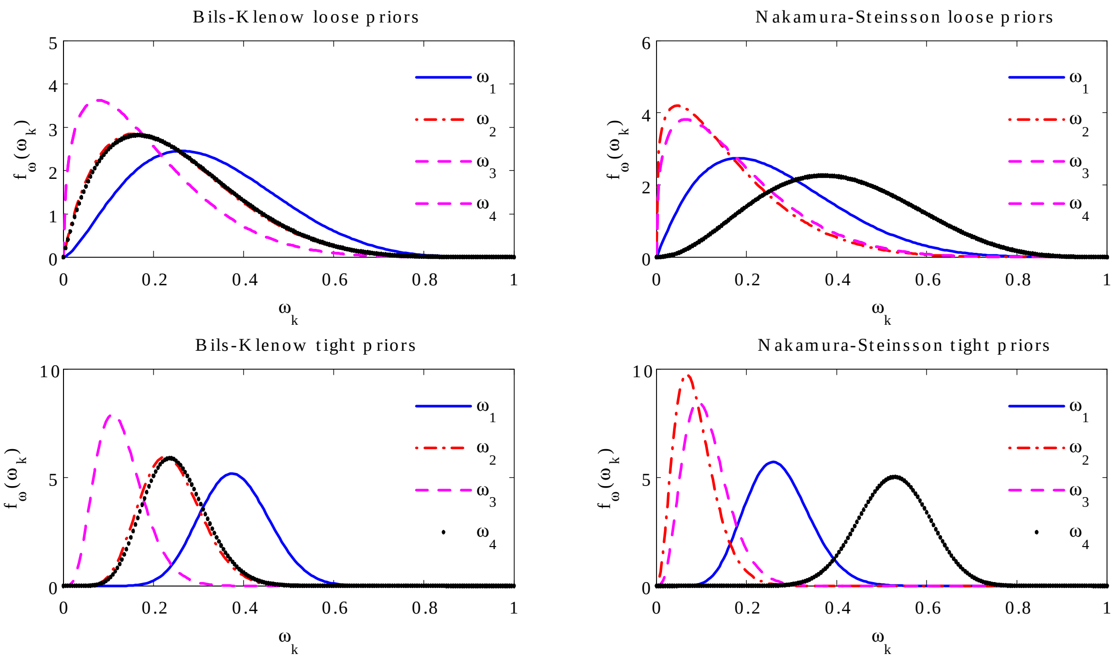

Figure 7: Sectoral mass marginal priors based on BK and NS

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 Appendix A). 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.[24] Next we run the so-called *fixed phase* of the MCMC. We take the estimate of the posterior mode and sample covariance matrix from the adaptive phase, and run 5 parallel chains of 300,000 iterations each. Again, before making the draws that will form the sample we rescale such covariance matrix in order to get rejection rates as close as possible to 0.77. To initialize each chain we draw from a candidate normal distribution centered on the posterior mode estimate, with covariance matrix given by 9 times the fine-tuned covariance matrix. We check for convergence for the latter 2/3s of the draws of all 5 chains by calculating the potential scale reduction[25] (PSR) factors for each

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.

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

[25] 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

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.

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

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

where $\mathcal{M}_j$ 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 MD_j for each model using Geweke's (1999) modified harmonic mean. We use this to evaluate the empirical fit of the models relative to one another; the MD ratio of two models 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^*.

5 Results

5.1 Flat prior on ω

We start with the results from specifications in which the prior ditribution for the sectoral weights ω is uninformative, which are summarized in Table 4 and in Figures 8-10 in terms of marginal distributions for the parameters.[26]. The first noteworhty result is that when we restrict the maximum degree of price rigidity to one year ($K = 4$), the resulting distribution of sectoral weights is somewhat similar to the empirical distribution with up to 4 quarters of price rigidity constructed from Bils and Klenow (2004).[27] In particular, using the means as the point estimates for the sectoral weights (Table 4, third column under $K = 4$), the fraction of firms changing prices every quarter is 58.3% (versus 39.5% in BK), every two quarters is 15.6% (24% in BK), every three quarters is 3.6% (11.6% in BK), and once a year is 22.5% (25% in BK). The correlation coefficient between the two is around 97%.[28] In contrast, the correlation with the corresponding distribution constructed from NS is much lower at

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.

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

[27] In constructing this truncated empirical distribution from the microeconomic data, we add the weights of all sectors that reprice on average once a year or less often into the last sector (see Subsection 4.3).

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

31%. This reflects the fact that, despite still having significant mass in the flexible price sector, the NS distribution has disproportionally more mass for price changes that occur with an average frequency of once-a-year or less.

As we allow the degree of price rigidity in some sectors to exceed one year ($K = 6, 8$), the estimated distribution of sectoral weights shifts in the direction of more average price rigidity (higher $\bar{k}$), and more heterogeneity in price stickiness (higher σ_k). As a result, the distribution of sectoral weights becomes more similar to the empirical distributions constructed from NS. This pattern accords well with the fact that the cross-sectional distribution of price stickiness derived from BK is much more concentrated in the sectors with very frequent price changes, when compared with the distribution derived from NS based on the frequency of regular price changes (see Subsection 4.2). When $K = 6$, the correlation between the mean sectoral weights and the NS empirical distribution increases to 73%. In turn, the correlation with the BK empirical distribution drops to 56%. For the model with up to two years of price rigidity ($K = 8$), the correlation with NS is also higher, at 63%, while the correlation with BK is 43%.

In addition, when we increase the number of sectors beyond $K = 4$, the degree of real rigidity falls (higher ζ). Note that when $K = 4$, the extent of real rigidity is incredibly high when assessed against the range of values that can be obtained in fully-specified models under reasonable calibrations for preference and technology parameters - the 95^{th} percentile of ζ is 0.02. In contrast, for $K = 8$ the degree of real rigidity is more in line with what can be obtained through calibration, although it is still on the high side of what is usually deemed plausible in the literature.[29]

Finally, the distribution of remaining parameters is remarkably stable, perhaps with the exception of the variance of shocks to the natural rate of output. These tend to become less volatile as we allow for more sectors and both the average degree of price rigidity and the extent of heterogeneity and price stickiness increase. In contrast, and perhaps somewhat surprisingly, the dynamic properties of this latent stochastic process remain relatively unchanged across the various specifications. As for the parameters describing the nominal income process, the stability is less surprising given that this variable is included as one of the observable time-series.

Despite the very different empirical methodology, our results with $K = 4$ are quite similar to those obtained by Coenen et al. (2007). They also estimate a model with Taylor staggered price setting and heterogeneous price spells with durations of up to four quarters, but rely on indirect inference methods. They find an average contract length that is in line with the evidence in Bils and

[29]The posterior mean of ζ is 0.05 and the 95^{th} percentile equals 0.11, which falls within the 0.10-0.15 range that Woodford (2003) argues can be made consistent with fully-specified models.

Klenow (2004), but at the same time find an incredible amount of real rigidity. However, besides the economically meaningful differences between the estimated models with $K = 4$ and with $K = 6, 8$, our results point to an important discrepancy in terms of their statistical performance. In particular, the data strongly favor models with 6 or 8 sectors over the specification with at most one year of price rigidity. While the difference between the models with $K = 6$ and $K = 8$ is small (the posterior odds ratio in favor of the latter is around $5 : 1$), the evidence against the specification with $K = 4$ is quite strong: the posterior odds ratios favor the models with $K = 6, 8$ by an order of $10^5 : 1$.

Our findings suggest that specifications allowing for more heterogeneity can produce significantly different results, in particular in terms of the amount of nominal and real rigidities. Results based on these specifications may be more informative for the purpose of identifying the roles of such rigidities in shaping the dynamic response of the economy to aggregate shocks.

5.2 Informative priors on ω

Given our Monte-Carlo evidence on the small sample properties of cross-sectional estimates obtained from aggregate data only, in this subsection we present results of estimations that make use of informed priors for the sectoral weights. They are summarized in Table 5, and presented in more detail in Tables 6-9, and in Figures 11-22 in terms of marginal distributions for the parameters.[30] A general trend from the case of the flat priors is confirmed; as we allow for sectors with price spells that last beyond four quarters, these sectors are assigned significant mass, and usually more than in the prior. In case of the Bils-Klenow priors, this implies that the posterior estimate of the average duration of price rigidity, $\bar{k}$, is pulled up relative to its prior level, increasing the average degree of nominal rigidity in the model. For the Nakamura-Steinsson priors, however, the increased posterior weight on long spells is associated with increased weight on the flexible price sector, leaving the posterior $\bar{k}$ close to the prior. Still, the posterior estimate of $\bar{k}$ is higher for NS priors than BK priors.

Across the different specifications of the pirors, higher estimates of $\bar{k}$ are associated with higher estimates of ζ, i.e., a lower degree of strategic complementarity in price setting. Furthermore, as ζ increases, the variance of the innovations to the natural output process, σ_n^2, decreases. The connection between these key parameters is summarized in Table 5. Another key feature, that proves to be common across specifications with heterogeneous firms, is that the estimation pushes towards an increased degree of heterogeneity as captured by the standard deviation of the durations of price rigidity, σ_k. Also, as prices respond more to movements in the output gap, the estimation implies less

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

volatility in the natural rate of output.

While the different priors appear to have a relatively small influence on the overall fit, allowing for more than four quarters of nominal rigidity proves to be important. For all specifications with $K = 4$, the posterior distribution for $\bar{k}$ indicates *less* nominal rigidity and only slightly more heterogeneity than in the priors. In particular, in all cases the posterior median for $\bar{k}$ is in the range of 2 to 2.6 quarters, and therefore in line with the estimates for average nominal price rigidity obtained by Bils and Klenow (2004). However, these results are at odds with specifications with $K > 4$, which clearly indicate the presence of both more nominal rigidity *and* heterogeneity.

As already discussed in the previous subsection, the estimation with flat priors resemble the BK data quite closely for $K = 4$, but once we allow for more price rigidity in some sectors, the estimated sectoral weights under flat priors look more like the NS priors. This pattern is confirmed in the estimations; for $K = 4$, the posterior distributions from the flat-prior estimation and the BK priors look very similar, while the posteriors based on the NS tight prior stands out with more nominal rigidity (higher $\bar{k}$) and less real rigidity (higher ζ). For $K = 4$, the evidence favors the BK distribution, yielding the highest marginal density for the BK tight priors, although the differences are not very large (the posterior odds over the tight NS priors is $8 : 1$).

As soon as we increase K to 6 or 8, however, the ranking of the configuation changes. As we tighten the BK priors, the estimation produces less nominal rigidity and compensates with more real rigidity, compared with the flat or NS priors. This worsens the empirical fit. For $K = 6, 8$, the worst fit with tight NS priors has a posterior odds over the best fit with tight BK priors of $30 : 1$. For these configurations, the NS priors produce posteriors that are quite similar to those based on flat priors, and there are only minor differences in the empirical fit as measured by the marginal densities. Hence, the finding from the flat priors remain: the marginal density is markedly reduced if one does not allow for $K > 4$.

From these results we conclude that the evidence favors specifications with more average nominal rigidity and more heterogeneity in price stickiness than can be accommodated in a model with $K = 4$. Our results also confirm the conclusions from the studies based on BLS microeconomic data, that four-quarter price spells seem to be of no particular empirical relevance. This contrasts with the view based on the earlier empirical literature on price setting, that the typical firm in the U.S. tends to change prices once a year, which was often used as a justification to calibrate one-sector models with one year of price rigidity. Finally, we find that estimations restricting price spells to last no more than one year produce "too much" real rigidity in an attempt to generate sufficiently sluggish dynamics. In specifications with $K > 4$ real rigidity is "traded-off" against a more spread-out distribution of price spells to yield a better fit.

5.3 Models with homogeneous firms

In this subsection we ask how sharply the data allows us to discriminate between our multi-sector models with heterogeneity in price stickiness, and one-sector models with homogeneous firms. We estimate models with price spells ranging from two to eight quarters. We keep the same prior distributions for all parameters besides the sectoral weights, and thus each such model with price spells of length k, say, can be seen as a restriction of the more general heterogeneous model, with a prior over the distribution of sectoral weights that puts probability one on $\omega_k = 1$. The results are presented in Tables 10-11 and in Figures 23-29.

The homogenous models produce estimates for some parameter values that are incredible. In particular, they can generate extreme values of ζ coupled with also extreme variances of the natural output process, σ_n^2. In our view, these results arise because of the attempt of such models with Taylor pricing and homogeneous price stickiness to generate relatively smooth dynamics. With $\zeta = 1$, such models generate somewhat unrealistic dynamics, with impulse response functions that tend to display sharp kinks at the time when all prices have been reset after the shock hit. Such kinks can be smoothed by moving ζ significantly away from unity, but that has implications for how the economy responds to different types of shocks. In particular, low values of ζ require high values of σ_n to fit the data, and vice-versa. In contrast, models with heterogeneity naturally tend to display smoother dynamics, and thus preclude the need for such large shifts in ζ and σ_n.

All specifications with homogeneous firms produce a poor statistical fit relative to the models with heterogeneity discussed in the previous subsections. As an example, picking the *worst* heterogeneous model and the *best* homogeneous model, the posterior odds in favor of the heterogeneous model is of the order of $10^6 : 1$. 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. Overall, we conclude that in the context of our semi-structural specifications, homogeneous models produce a much worse statistical fit, and economically unappealing results.

6 Conclusion

We estimate a series of small semi-structural models for the U.S. economy in which the extent of price rigidity varies across firms. We provide estimates based only on aggregate data on real and nominal output, and estimates that incorporate prior information on the extent of price stickiness at the microeconomic level from the recent empirical literature. Perhaps surprisingly, we find that the former accords extremely well with the latter evidence. More generally, we find overwhelming evidence in favor of specifications with heterogeneity in price stickiness, over the ones with homonegeous-firms.

Within the set of models with heterogeneity, it is important to allow for prices in some sector of the economy to last beyond one year. Neglecting to do so produces "too little" nominal rigidity, and increases the estimated degree of real rigidity beyond what is found in the specifications favored by the data. The latter produce estimates of real rigidity that are close to what is deemed plausible in the literature (e.g. Woodford, 2003), and deliver cross-sectional distributions of the frequency of price changes that are in line with the recent evidence based on microeconomic data for the U.S. economy.

We find the results sufficiently compelling to warrant further work. In particular, it would be interesting to evaluate the consequences of allowing for heterogenous pricing behavior when estimating models of fully structured DSGE economies. The experience with our semi-structural model suggests that combining microeconomic and macroeconomic data within a Bayesian framework can help us integrate our views on price setting at the microeconomic level, and its implications for aggregate dynamics. In addition, making use of sectoral data as well, along the lines of Lee (2007) and Boukaez et al. (2008), can further enhance our understanding of how actual monetary economies work.

Finally, our approach can be applied to other countries and time periods for which no direct microeconomic evidence on the frequency of price changes is available. This can be used to address questions such as how the distribution of the frequency of price changes is affected by the level of inflation or the monetary policy regime, and what are the implications of the estimated distribution of price stickiness for the conduct of policy.

Appendix

A 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 application of exponentials in this way guarantees the non-negativity and summation to unity constraints. However, without additional restrictions on h^{-1}, 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 adopt the restriction $v(K) = 0$ and in effect draw vectors $\widetilde{v} = (v_1, ..., v_{K-1})$ in $\mathbb{R}^{K-1}$. Thus, the transformation becomes $\widetilde{\omega} = \widetilde{h}^{-1}(\widetilde{v})$, with $\widetilde{\omega} = (\omega_1, ..., \omega_{K-1})$ and:

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

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

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

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

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

with:

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

So:

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

To recover the v_k's from ω simply set:

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

References

[1] Aoki, K. (2001), "Optimal Monetary Policy Responses to Relative-Price Changes," *Journal of Monetary Economics* 48: 55-80.

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

[3] Benigno, P. (2004), "Optimal Monetary Policy in a Currency Area," *Journal of International Economics* 63: 293-320.

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

[5] Blinder, A., E. Canetti, D. Lebow and J. Rudd (1998), *Asking about Prices: A New Approach to Understanding Price Stickiness*, Russell Sage Foundation.

[6] Boukaez, H., E. Cardia and F. Ruge-Murcia (2008), "The Transmission of Monetary Policy in a Multi-Sector Economy," forthcoming in the *International Economic Review.*

[7] Carlstrom, C., T. Fuerst and F. Ghironi (2006a), "Does It Matter (for Equilibrium Determinacy) What Price Index the Central Bank Targets?," *Journal Economic Theory* 128: 214-231.

[8] Carlstrom, C., T. Fuerst, F. Ghironi and K. Hernandez (2006b), "Relative Price Dynamics and the Aggregate Economy," mimeo available at http://copland.udel.edu/~kolver/research/index.html.

[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://cvianac.googlepages.com/papers.

[11] Carvalho, C. and F. Nechio (2008), "Aggregation and the PPP Puzzle in a Sticky Price Model," mimeo, available at http://cvianac.googlepages.com/papers.

[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] Dixon, H. and E. Kara (2005), "Persistence and Nominal Inertia in a Generalized Taylor Economy: How Longer Contracts Dominate Shorter Contracts," ECB Working Paper Series no. 489.

[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] Gagnon, E. (2008), "Price Setting During Low and High Inflation: Evidence from Mexico," forthcoming in the *Quarterly Journal of Economics*.

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

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

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

[20] Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.

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

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

[23] Lee, Jae Won (2007), "Heterogeneous Households, Real Rigidity, and Estimated Duration of Price Contract in a Sticky Price DSGE Model," mimeo, Princeton University.

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

[25] Nakamura, E. and J. Steinsson (2007), "Monetary Non-Neutrality in a Multi-Sector Menu Cost Model," mimeo available at http://www.columbia.edu/~js3204/cu-papers.html.

[26] ____________ (2008), "Five Facts About Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics* 123: 1415-1464.

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

[28] Sheedy, K. (2007), "Inflation Persistence When Price Stickiness Differs Between Industries," mimeo available at http://personal.lse.ac.uk/sheedy/.

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

[30] Smets, F. and R. Wouters (2003), "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association* 1: 1123-1175.

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

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

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

[34] Zaffaroni, P. (2004), "Contemporaneous Aggregation of Linear Dynamic Models in Large Economies," *Journal of Econometrics* 120: 75-102.

Table 4: Parameter estimates - Flat prior

	$K=4$			$K=6$			$K=8$		
ζ	4.440 (0.47;16.86)	0.006 (0.00;0.02)	0.008	4.440 (0.47;16.86)	0.021 (0.01;0.06)	0.025	4.440 (0.47;16.86)	0.042 (0.01;0.11)	0.050
ω_1	0.206 (0.02;0.63)	0.593 (0.31;0.82)	0.583	0.129 (0.01;0.45)	0.374 (0.16;0.63)	0.382	0.094 (0.01;0.35)	0.264 (0.10;0.49)	0.276
ω_2	0.206 (0.02;0.63)	0.134 (0.02;0.37)	0.156	0.129 (0.01;0.45)	0.090 (0.01;0.25)	0.104	0.094 (0.01;0.35)	0.072 (0.01;0.21)	0.086
ω_3	0.206 (0.02;0.63)	0.024 (0.00;0.11)	0.036	0.129 (0.01;0.45)	0.024 (0.00;0.09)	0.033	0.094 (0.01;0.35)	0.020 (0.00;0.08)	0.027
ω_4	0.206 (0.02;0.63)	0.215 (0.08;0.40)	0.225	0.129 (0.01;0.45)	0.033 (0.00;0.13)	0.045	0.094 (0.01;0.35)	0.027 (0.00;0.11)	0.037
ω_5	–	–	–	0.129 (0.01;0.45)	0.157 (0.02;0.39)	0.173	0.094 (0.01;0.35)	0.144 (0.02;0.34)	0.156
ω_6	–	–	–	0.129 (0.01;0.45)	0.258 (0.05;0.50)	0.264	0.094 (0.01;0.35)	0.123 (0.01;0.35)	0.144
ω_7	–	–	–	–	–	–	0.094 (0.01;0.35)	0.120 (0.01;0.35)	0.143
ω_8	–	–	–	–	–	–	0.094 (0.01;0.35)	0.112 (0.01;0.32)	0.132
$\bar{k}$	2.499 (1.67;3.32)	1.885 (1.39;2.48)	1.904	3.504 (2.44;4.56)	3.322 (2.34;4.25)	3.315	4.503 (3.25;5.76)	4.394 (3.21;5.46)	4.374
σ_k	0.983 (0.62;1.32)	1.193 (0.88;1.39)	1.229	1.559 (1.09;2.01)	2.087 (1.77;2.30)	2.155	2.137 (1.58;2.68)	2.523 (2.11;2.89)	2.616
ρ_0	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000
ρ_1	0.000 (-8.22;8.22)	1.422 (1.27;1.57)	1.422	0.000 (-8.22;8.22)	1.410 (1.26;1.56)	1.410	0.000 (-8.22;8.22)	1.426 (1.27;1.58)	1.426
ρ_2	0.000 (-8.22;8.22)	-0.442 (-0.59;-0.29)	-0.441	0.000 (-8.22;8.22)	-0.430 (-0.57;-0.28)	-0.429	0.000 (-8.22;8.22)	-0.446 (-0.59;-0.30)	-0.446
σ_m	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005
δ_0	0.000 (-8.22;8.22)	0.003 (-0.00;0.01)	0.003	0.000 (-8.22;8.22)	0.003 (-0.00;0.01)	0.003	0.000 (-8.22;8.22)	0.002 (-0.00;0.01)	0.003
δ_1	0.000 (-8.22;8.22)	0.417 (0.15;0.63)	0.408	0.000 (-8.22;8.22)	0.519 (0.25;0.74)	0.510	0.000 (-8.22;8.22)	0.541 (0.27;0.76)	0.532
δ_2	0.000 (-8.22;8.22)	0.110 (-0.03;0.26)	0.111	0.000 (-8.22;8.22)	0.128 (-0.04;0.31)	0.130	0.000 (-8.22;8.22)	0.146 (-0.03;0.33)	0.149
σ_n	0.059 (0.01;0.20)	0.122 (0.05;0.26)	0.135	0.059 (0.01;0.20)	0.084 (0.04;0.19)	0.096	0.059 (0.01;0.20)	0.069 (0.03;0.17)	0.081
log MD	–	794.993	–	–	806.418	–	–	808.031	–

Note: For each configuration, the first two columns report the median - with the 5^{th} and 95^{th} percentiles in parentheses - for, respectively, the marginal prior and the marginal posterior; the third column gives the mean of the marginal distribution. The logarithm of the Marginal Posterior (log MD, bottom row) is approximated with Geweke's (1999) modified harmonic mean.

Table 5: Selected parameter estimates - models with heterogeneous firms

Prior	$K=4$			$K=6$			$K=8$		
Flat									
ζ	4.440 (0.47;16.86)	0.006 (0.00;0.02)	0.008	4.440 (0.47;16.86)	0.021 (0.01;0.06)	0.025	4.440 (0.47;16.86)	0.042 (0.01;0.11)	0.050
$\bar{k}$	2.499 (1.67;3.32)	1.885 (1.39;2.48)	1.904	3.504 (2.44;4.56)	3.322 (2.34;4.25)	3.315	4.503 (3.25;5.76)	4.394 (3.21;5.46)	4.374
σ_k	0.983 (0.62;1.32)	1.193 (0.88;1.39)	1.229	1.559 (1.09;2.01)	2.087 (1.77;2.30)	2.155	2.137 (1.58;2.68)	2.523 (2.11;2.89)	2.616
σ_n	0.059 (0.01;0.20)	0.122 (0.05;0.26)	0.135	0.059 (0.01;0.20)	0.084 (0.04;0.19)	0.096	0.059 (0.01;0.20)	0.069 (0.03;0.17)	0.081
log MD	–	794.993	–	–	806.418	–	–	808.031	–
Loose NS									
ζ	4.440 (0.47;16.86)	0.008 (0.00;0.02)	0.010	4.440 (0.47;16.86)	0.024 (0.01;0.06)	0.028	4.440 (0.47;16.86)	0.042 (0.02;0.11)	0.049
$\bar{k}$	2.736 (2.02;3.38)	2.114 (1.64;2.61)	2.117	3.637 (2.75;4.51)	3.534 (2.72;4.29)	3.523	4.364 (3.38;5.36)	4.367 (3.45;5.25)	4.360
σ_k	1.172 (0.88;1.38)	1.281 (1.07;1.42)	1.306	1.843 (1.49;2.15)	2.136 (1.91;2.30)	2.180	2.405 (1.98;2.79)	2.612 (2.28;2.91)	2.670
σ_n	0.059 (0.01;0.20)	0.128 (0.05;0.27)	0.140	0.059 (0.01;0.20)	0.087 (0.04;0.20)	0.099	0.059 (0.01;0.20)	0.066 (0.03;0.16)	0.077
log MD	–	795.011	–	–	806.944	–	–	808.270	–
Tight NS									
ζ	4.440 (0.47;16.86)	0.012 (0.01;0.03)	0.014	4.440 (0.47;16.86)	0.027 (0.01;0.06)	0.031	4.440 (0.47;16.86)	0.041 (0.02;0.10)	0.047
$\bar{k}$	2.901 (2.55;3.22)	2.616 (2.31;2.91)	2.613	3.753 (3.31;4.18)	3.727 (3.30;4.14)	3.724	4.258 (3.79;4.74)	4.280 (3.81;4.75)	4.281
σ_k	1.290 (1.14;1.39)	1.360 (1.27;1.42)	1.368	2.049 (1.90;2.18)	2.132 (2.01;2.24)	2.145	2.610 (2.42;2.78)	2.676 (2.51;2.83)	2.691
σ_n	0.059 (0.01;0.20)	0.150 (0.06;0.30)	0.162	0.059 (0.01;0.20)	0.094 (0.04;0.20)	0.104	0.059 (0.01;0.20)	0.066 (0.03;0.16)	0.077
log MD	–	793.446	–	–	806.990	–	–	807.530	–
Loose BK									
ζ	4.440 (0.47;16.86)	0.007 (0.00;0.02)	0.008	4.440 (0.47;16.86)	0.020 (0.01;0.05)	0.023	4.440 (0.47;16.86)	0.032 (0.01;0.08)	0.037
$\bar{k}$	2.350 (1.73;3.02)	1.992 (1.54;2.49)	2.000	2.951 (2.21;3.80)	3.140 (2.37;3.92)	3.143	3.495 (2.64;4.49)	3.776 (2.91;4.69)	3.786
σ_k	1.104 (0.82;1.33)	1.208 (0.97;1.37)	1.235	1.668 (1.29;2.02)	2.050 (1.77;2.25)	2.093	2.222 (1.75;2.67)	2.515 (2.17;2.85)	2.579
σ_n	0.059 (0.01;0.20)	0.130 (0.06;0.27)	0.142	0.059 (0.01;0.20)	0.086 (0.04;0.20)	0.098	0.059 (0.01;0.20)	0.068 (0.03;0.16)	0.078
log MD	–	795.350	–	–	806.024	–	–	807.563	–
Tight BK									
ζ	4.440 (0.47;16.86)	0.008 (0.00;0.02)	0.010	4.440 (0.47;16.86)	0.015 (0.01;0.04)	0.018	4.440 (0.47;16.86)	0.019 (0.01;0.05)	0.022
$\bar{k}$	2.245 (1.94;2.56)	2.155 (1.88;2.44)	2.158	2.543 (2.21;2.91)	2.683 (2.33;3.07)	2.690	2.724 (2.39;3.12)	2.878 (2.52;3.29)	2.887
σ_k	1.190 (1.06;1.30)	1.214 (1.09;1.31)	1.223	1.643 (1.43;1.85)	1.836 (1.63;2.01)	1.849	1.969 (1.70;2.26)	2.165 (1.90;2.43)	2.184
σ_n	0.059 (0.01;0.20)	0.146 (0.06;0.29)	0.158	0.059 (0.01;0.20)	0.093 (0.04;0.21)	0.105	0.059 (0.01;0.20)	0.077 (0.03;0.18)	0.088
log MD	–	795.516	–	–	803.345	–	–	803.558	–

Note: For each configuration, the first two columns report the median - with the 5^{th} and 95^{th} percentiles in parentheses - for, respectively, the marginal prior and the marginal posterior; the third column gives the mean of the marginal distribution. The logarithm of the Marginal Posterior (log MD) is approximated with Geweke's (1999) modified harmonic mean.

Table 6: Parameter estimates - Nakamura-Steinsson loose prior

	$K=4$			$K=6$			$K=8$		
ζ	4.440 (0.47;16.86)	0.008 (0.00;0.02)	0.010	4.440 (0.47;16.86)	0.024 (0.01;0.06)	0.028	4.440 (0.47;16.86)	0.042 (0.02;0.11)	0.049
ω_1	0.241 (0.06;0.53)	0.513 (0.29;0.72)	0.510	0.204 (0.06;0.43)	0.337 (0.17;0.54)	0.344	0.186 (0.06;0.38)	0.277 (0.14;0.45)	0.283
ω_2	0.132 (0.02;0.40)	0.138 (0.02;0.33)	0.153	0.098 (0.01;0.30)	0.086 (0.01;0.22)	0.097	0.081 (0.01;0.24)	0.069 (0.01;0.18)	0.080
ω_3	0.147 (0.02;0.42)	0.037 (0.00;0.12)	0.047	0.112 (0.02;0.32)	0.037 (0.01;0.11)	0.044	0.095 (0.02;0.26)	0.033 (0.01;0.09)	0.038
ω_4	0.396 (0.15;0.69)	0.285 (0.15;0.45)	0.290	0.118 (0.02;0.32)	0.050 (0.01;0.14)	0.059	0.102 (0.02;0.27)	0.047 (0.01;0.12)	0.054
ω_5	–	–	–	0.092 (0.01;0.29)	0.108 (0.02;0.28)	0.123	0.076 (0.01;0.23)	0.109 (0.02;0.26)	0.120
ω_6	–	–	–	0.265 (0.10;0.50)	0.331 (0.16;0.52)	0.334	0.112 (0.03;0.28)	0.142 (0.04;0.31)	0.154
ω_7	–	–	–	–	–	–	0.077 (0.01;0.23)	0.086 (0.01;0.24)	0.101
ω_8	–	–	–	–	–	–	0.147 (0.04;0.33)	0.160 (0.05;0.32)	0.169
$\bar{k}$	2.736 (2.02;3.38)	2.114 (1.64;2.61)	2.117	3.637 (2.75;4.51)	3.534 (2.72;4.29)	3.523	4.364 (3.38;5.36)	4.367 (3.45;5.25)	4.360
σ_k	1.172 (0.88;1.38)	1.281 (1.07;1.42)	1.306	1.843 (1.49;2.15)	2.136 (1.91;2.30)	2.180	2.405 (1.98;2.79)	2.612 (2.28;2.91)	2.670
ρ_0	0.000 (−8.22;8.22)	0.000 (−0.00;0.00)	0.000	0.000 (−8.22;8.22)	0.000 (−0.00;0.00)	0.000	0.000 (−8.22;8.22)	0.000 (−0.00;0.00)	0.000
ρ_1	0.000 (−8.22;8.22)	1.423 (1.27;1.57)	1.423	0.000 (−8.22;8.22)	1.409 (1.26;1.55)	1.409	0.000 (−8.22;8.22)	1.427 (1.27;1.58)	1.426
ρ_2	0.000 (−8.22;8.22)	−0.443 (−0.59;−0.29)	−0.442	0.000 (−8.22;8.22)	−0.429 (−0.57;−0.28)	−0.428	0.000 (−8.22;8.22)	−0.447 (−0.59;−0.30)	−0.446
σ_m	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005
δ_0	0.000 (−8.22;8.22)	0.003 (−0.00;0.01)	0.003	0.000 (−8.22;8.22)	0.003 (−0.00;0.01)	0.003	0.000 (−8.22;8.22)	0.002 (−0.00;0.01)	0.002
δ_1	0.000 (−8.22;8.22)	0.416 (0.18;0.62)	0.410	0.000 (−8.22;8.22)	0.506 (0.27;0.71)	0.501	0.000 (−8.22;8.22)	0.545 (0.32;0.75)	0.541
δ_2	0.000 (−8.22;8.22)	0.098 (−0.04;0.24)	0.099	0.000 (−8.22;8.22)	0.133 (−0.02;0.30)	0.135	0.000 (−8.22;8.22)	0.151 (−0.01;0.32)	0.152
σ_n	0.059 (0.01;0.20)	0.128 (0.05;0.27)	0.140	0.059 (0.01;0.20)	0.087 (0.04;0.20)	0.099	0.059 (0.01;0.20)	0.066 (0.03;0.16)	0.077
log MD	–	795.011	–	–	806.944	–	–	808.270	–

Note: For each configuration, the first two columns report the median - with the 5^{th} and 95^{th} percentiles in parentheses - for, respectively, the marginal prior and the marginal posterior; the third column gives the mean of the marginal distribution. The logarithm of the Marginal Posterior (log MD, bottom row) is approximated with Geweke's (1999) modified harmonic mean.

Table 7: Parameter estimates - Nakamura-Steinsson tight prior

	$K=4$			$K=6$			$K=8$		
ζ	4.440 (0.47;16.86)	0.012 (0.01;0.03)	0.014	4.440 (0.47;16.86)	0.027 (0.01;0.06)	0.031	4.440 (0.47;16.86)	0.041 (0.02;0.10)	0.047
ω_1	0.267 (0.16;0.39)	0.365 (0.26;0.48)	0.366	0.260 (0.17;0.36)	0.294 (0.21;0.39)	0.296	0.256 (0.18;0.34)	0.282 (0.21;0.37)	0.283
ω_2	0.082 (0.03;0.17)	0.100 (0.04;0.20)	0.106	0.076 (0.03;0.14)	0.075 (0.03;0.14)	0.079	0.073 (0.03;0.13)	0.071 (0.03;0.12)	0.074
ω_3	0.107 (0.04;0.20)	0.072 (0.03;0.14)	0.077	0.100 (0.05;0.18)	0.071 (0.03;0.12)	0.074	0.097 (0.05;0.16)	0.069 (0.04;0.11)	0.071
ω_4	0.528 (0.40;0.65)	0.451 (0.34;0.56)	0.451	0.111 (0.06;0.19)	0.088 (0.04;0.15)	0.091	0.108 (0.06;0.17)	0.087 (0.05;0.14)	0.090
ω_5	–	–	–	0.066 (0.03;0.13)	0.069 (0.03;0.13)	0.074	0.063 (0.03;0.12)	0.071 (0.03;0.13)	0.074
ω_6	–	–	–	0.365 (0.27;0.47)	0.385 (0.29;0.48)	0.386	0.126 (0.07;0.19)	0.138 (0.08;0.21)	0.141
ω_7	–	–	–	–	–	–	0.064 (0.03;0.12)	0.067 (0.03;0.12)	0.071
ω_8	–	–	–	–	–	–	0.188 (0.12;0.27)	0.193 (0.13;0.27)	0.195
$\bar{k}$	2.901 (2.55;3.22)	2.616 (2.31;2.91)	2.613	3.753 (3.31;4.18)	3.727 (3.30;4.14)	3.724	4.258 (3.79;4.74)	4.280 (3.81;4.75)	4.281
σ_k	1.290 (1.14;1.39)	1.360 (1.27;1.42)	1.368	2.049 (1.90;2.18)	2.132 (2.01;2.24)	2.145	2.610 (2.42;2.78)	2.676 (2.51;2.83)	2.691
ρ_0	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000
ρ_1	0.000 (-8.22;8.22)	1.429 (1.28;1.57)	1.428	0.000 (-8.22;8.22)	1.411 (1.26;1.55)	1.410	0.000 (-8.22;8.22)	1.429 (1.28;1.58)	1.428
ρ_2	0.000 (-8.22;8.22)	-0.448 (-0.59;-0.30)	-0.447	0.000 (-8.22;8.22)	-0.430 (-0.57;-0.28)	-0.429	0.000 (-8.22;8.22)	-0.449 (-0.60;-0.30)	-0.448
σ_m	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005
δ_0	0.000 (-8.22;8.22)	0.003 (-0.00;0.01)	0.003	0.000 (-8.22;8.22)	0.003 (-0.00;0.01)	0.003	0.000 (-8.22;8.22)	0.002 (-0.00;0.01)	0.002
δ_1	0.000 (-8.22;8.22)	0.414 (0.24;0.58)	0.412	0.000 (-8.22;8.22)	0.490 (0.31;0.67)	0.490	0.000 (-8.22;8.22)	0.544 (0.37;0.72)	0.543
δ_2	0.000 (-8.22;8.22)	0.059 (-0.06;0.18)	0.061	0.000 (-8.22;8.22)	0.131 (-0.00;0.27)	0.132	0.000 (-8.22;8.22)	0.150 (0.00;0.30)	0.150
σ_n	0.059 (0.01;0.20)	0.150 (0.06;0.30)	0.162	0.059 (0.01;0.20)	0.094 (0.04;0.20)	0.104	0.059 (0.01;0.20)	0.066 (0.03;0.16)	0.077
log MD	–	793.446	–	–	806.990	–	–	807.530	–

Note: For each configuration, the first two columns report the median - with the 5^{th} and 95^{th} percentiles in parentheses - for, respectively, the marginal prior and the marginal posterior; the third column gives the mean of the marginal distribution. The logarithm of the Marginal Posterior (log MD, bottom row) is approximated with Geweke's (1999) modified harmonic mean.

Table 8: Parameter estimates - Bils-Klenow loose prior

	$K=4$			$K=6$			$K=8$		
ζ	4.440 (0.47;16.86)	0.007 (0.00;0.02)	0.008	4.440 (0.47;16.86)	0.020 (0.01;0.05)	0.023	4.440 (0.47;16.86)	0.032 (0.01;0.08)	0.037
ω_1	0.307 (0.09;0.60)	0.528 (0.29;0.74)	0.525	0.268 (0.10;0.51)	0.376 (0.20;0.58)	0.381	0.250 (0.10;0.45)	0.324 (0.17;0.51)	0.329
ω_2	0.223 (0.05;0.51)	0.175 (0.05;0.37)	0.189	0.186 (0.05;0.41)	0.132 (0.04;0.27)	0.140	0.169 (0.05;0.36)	0.123 (0.04;0.24)	0.130
ω_3	0.156 (0.02;0.43)	0.038 (0.01;0.13)	0.048	0.121 (0.02;0.33)	0.037 (0.01;0.10)	0.044	0.105 (0.02;0.27)	0.035 (0.01;0.09)	0.041
ω_4	0.228 (0.05;0.52)	0.231 (0.11;0.39)	0.238	0.122 (0.02;0.33)	0.050 (0.01;0.14)	0.059	0.106 (0.02;0.27)	0.049 (0.01;0.13)	0.057
ω_5	–	–	–	0.080 (0.01;0.27)	0.127 (0.02;0.32)	0.141	0.064 (0.01;0.21)	0.106 (0.02;0.26)	0.119
ω_6	–	–	–	0.110 (0.02;0.31)	0.230 (0.06;0.42)	0.235	0.062 (0.01;0.21)	0.100 (0.01;0.27)	0.116
ω_7	–	–	–	–	–	–	0.060 (0.01;0.21)	0.090 (0.01;0.25)	0.105
ω_8	–	–	–	–	–	–	0.061 (0.01;0.21)	0.088 (0.01;0.24)	0.102
$\bar{k}$	2.350 (1.73;3.02)	1.992 (1.54;2.49)	2.000	2.951 (2.21;3.80)	3.140 (2.37;3.92)	3.143	3.495 (2.64;4.49)	3.776 (2.91;4.69)	3.786
σ_k	1.104 (0.82;1.33)	1.208 (0.97;1.37)	1.235	1.668 (1.29;2.02)	2.050 (1.77;2.25)	2.093	2.222 (1.75;2.67)	2.515 (2.17;2.85)	2.579
ρ_0	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000
ρ_1	0.000 (-8.22;8.22)	1.424 (1.27;1.57)	1.423	0.000 (-8.22;8.22)	1.415 (1.26;1.56)	1.413	0.000 (-8.22;8.22)	1.425 (1.27;1.57)	1.424
ρ_2	0.000 (-8.22;8.22)	-0.444 (-0.59;-0.30)	-0.443	0.000 (-8.22;8.22)	-0.434 (-0.58;-0.28)	-0.433	0.000 (-8.22;8.22)	-0.445 (-0.59;-0.30)	-0.444
σ_m	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005
δ_0	0.000 (-8.22;8.22)	0.003 (-0.00;0.01)	0.003	0.000 (-8.22;8.22)	0.003 (-0.00;0.01)	0.003	0.000 (-8.22;8.22)	0.002 (-0.00;0.01)	0.002
δ_1	0.000 (-8.22;8.22)	0.384 (0.14;0.59)	0.378	0.000 (-8.22;8.22)	0.477 (0.25;0.69)	0.475	0.000 (-8.22;8.22)	0.514 (0.30;0.72)	0.512
δ_2	0.000 (-8.22;8.22)	0.118 (-0.02;0.26)	0.119	0.000 (-8.22;8.22)	0.158 (-0.00;0.32)	0.159	0.000 (-8.22;8.22)	0.176 (0.01;0.34)	0.176
σ_n	0.059 (0.01;0.20)	0.130 (0.06;0.27)	0.142	0.059 (0.01;0.20)	0.086 (0.04;0.20)	0.098	0.059 (0.01;0.20)	0.068 (0.03;0.16)	0.078
log MD	–	795.350	–	–	806.024	–	–	807.563	–

Note: For each configuration, the first two columns report the median - with the 5^{th} and 95^{th} percentiles in parentheses - for, respectively, the marginal prior and the marginal posterior; the third column gives the mean of the marginal distribution. The logarithm of the Marginal Posterior (log MD, bottom row) is approximated with Geweke's (1999) modified harmonic mean.

Table 9: Parameter estimates - Bils-Klenow tight prior

	$K = 4$			$K = 6$			$K = 8$		
ζ	4.440 (0.47;16.86)	0.008 (0.00;0.02)	0.010	4.440 (0.47;16.86)	0.015 (0.01;0.04)	0.018	4.440 (0.47;16.86)	0.019 (0.01;0.05)	0.022
ω_1	0.379 (0.26;0.51)	0.430 (0.31;0.55)	0.431	0.371 (0.27;0.48)	0.395 (0.30;0.50)	0.397	0.367 (0.28;0.46)	0.387 (0.30;0.48)	0.388
ω_2	0.236 (0.14;0.36)	0.228 (0.14;0.34)	0.232	0.229 (0.15;0.33)	0.202 (0.13;0.29)	0.205	0.226 (0.16;0.31)	0.204 (0.14;0.28)	0.206
ω_3	0.123 (0.05;0.22)	0.081 (0.04;0.15)	0.086	0.117 (0.06;0.20)	0.077 (0.04;0.13)	0.080	0.114 (0.06;0.18)	0.079 (0.04;0.13)	0.082
ω_4	0.245 (0.15;0.37)	0.248 (0.16;0.35)	0.251	0.118 (0.06;0.20)	0.097 (0.05;0.16)	0.100	0.115 (0.07;0.18)	0.098 (0.06;0.15)	0.100
ω_5	–	–	–	0.045 (0.01;0.10)	0.064 (0.02;0.14)	0.069	0.042 (0.01;0.09)	0.057 (0.02;0.11)	0.061
ω_6	–	–	–	0.097 (0.05;0.17)	0.146 (0.08;0.23)	0.150	0.039 (0.01;0.08)	0.054 (0.02;0.11)	0.058
ω_7	–	–	–	–	–	–	0.035 (0.01;0.08)	0.047 (0.02;0.10)	0.052
ω_8	–	–	–	–	–	–	0.037 (0.01;0.08)	0.049 (0.02;0.10)	0.053
$\bar{k}$	2.245 (1.94;2.56)	2.155 (1.88;2.44)	2.158	2.543 (2.21;2.91)	2.683 (2.33;3.07)	2.690	2.724 (2.39;3.12)	2.878 (2.52;3.29)	2.887
σ_k	1.190 (1.06;1.30)	1.214 (1.09;1.31)	1.223	1.643 (1.43;1.85)	1.836 (1.63;2.01)	1.849	1.969 (1.70;2.26)	2.165 (1.90;2.43)	2.184
ρ_0	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000	0.000 (-8.22;8.22)	0.000 (-0.00;0.00)	0.000
ρ_1	0.000 (-8.22;8.22)	1.426 (1.28;1.57)	1.426	0.000 (-8.22;8.22)	1.421 (1.27;1.57)	1.420	0.000 (-8.22;8.22)	1.423 (1.27;1.57)	1.423
ρ_2	0.000 (-8.22;8.22)	-0.446 (-0.59;-0.30)	-0.445	0.000 (-8.22;8.22)	-0.440 (-0.59;-0.29)	-0.440	0.000 (-8.22;8.22)	-0.444 (-0.59;-0.29)	-0.443
σ_m	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005
δ_0	0.000 (-8.22;8.22)	0.003 (-0.00;0.01)	0.003	0.000 (-8.22;8.22)	0.003 (-0.00;0.01)	0.003	0.000 (-8.22;8.22)	0.002 (-0.00;0.01)	0.003
δ_1	0.000 (-8.22;8.22)	0.344 (0.17;0.52)	0.345	0.000 (-8.22;8.22)	0.415 (0.24;0.59)	0.415	0.000 (-8.22;8.22)	0.449 (0.28;0.63)	0.450
δ_2	0.000 (-8.22;8.22)	0.117 (-0.00;0.24)	0.117	0.000 (-8.22;8.22)	0.187 (0.05;0.32)	0.186	0.000 (-8.22;8.22)	0.199 (0.06;0.34)	0.199
σ_n	0.059 (0.01;0.20)	0.146 (0.06;0.29)	0.158	0.059 (0.01;0.20)	0.093 (0.04;0.21)	0.105	0.059 (0.01;0.20)	0.077 (0.03;0.18)	0.088
log MD	–	795.516	–	–	803.345	–	–	803.558	–

Note: For each configuration, the first two columns report the median - with the 5^{th} and 95^{th} percentiles in parentheses - for, respectively, the marginal prior and the marginal posterior; the third column gives the mean of the marginal distribution. The logarithm of the Marginal Posterior (log MD, bottom row) is approximated with Geweke's (1999) modified harmonic mean.

Table 10: Parameter estimates - models with homogeneous firms: K=2-5

	Prior	$K=2$		$K=3$		$K=4$		$K=5$	
ζ	4.440 (0.47;16.86)	2.060 (1.32;3.69)	2.230	0.074 (0.05;0.13)	0.080	40.409 (23.59;63.74)	41.743	0.208 (0.11;0.52)	0.248
ρ_0	0.000 $(-8.22;8.22)$	0.000 $(-0.00;0.00)$	0.000	0.000 $(-0.00;0.00)$	0.000	0.000 $(-0.00;0.00)$	0.000	0.000 $(-0.00;0.00)$	0.000
ρ_1	0.000 $(-8.22;8.22)$	1.263 (1.09;1.42)	1.260	1.394 (1.25;1.54)	1.393	1.335 (1.18;1.49)	1.333	1.402 (1.25;1.54)	1.401
ρ_2	0.000 $(-8.22;8.22)$	-0.293 $(-0.45;-0.12)$	-0.290	-0.426 $(-0.57;-0.28)$	-0.424	-0.362 $(-0.52;-0.21)$	-0.361	-0.430 $(-0.57;-0.28)$	-0.429
σ_m	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.006	0.005 (0.00;0.01)	0.005	0.005 (0.00;0.01)	0.005	0.005 (0.00;0.01)	0.005
δ_0	0.000 $(-8.22;8.22)$	0.001 $(-0.00;0.00)$	0.001	0.004 $(-0.00;0.01)$	0.004	0.001 $(-0.00;0.00)$	0.001	0.003 $(-0.00;0.01)$	0.003
δ_1	0.000 $(-8.22;8.22)$	-0.133 $(-0.20;-0.07)$	-0.134	-0.224 $(-0.41;0.03)$	-0.213	-0.174 $(-0.26;-0.10)$	-0.175	0.113 $(-0.13;0.42)$	0.122
δ_2	0.000 $(-8.22;8.22)$	0.851 (0.79;0.91)	0.850	-0.079 $(-0.14;-0.00)$	-0.077	0.816 (0.73;0.89)	0.815	0.117 $(-0.10;0.33)$	0.117
σ_n	0.059 (0.01;0.20)	0.015 (0.01;0.02)	0.015	0.522 (0.29;0.82)	0.535	0.019 (0.02;0.02)	0.020	0.188 (0.07;0.36)	0.199
log MD	–	730.387	–	699.299	–	687.018	–	768.368	–

Note: The first column gives the medians for the marginal prior distributions. For each specification, the first and second columns give, respectively, the medians and means of the marginal posterior distributions. Numbers in parentheses correspond to the 5th and 95th percentiles. The logarithm of the Marginal Posterior (log MD, bottom row) is approximated with Geweke's (1999) modified harmonic mean.

Table 11: Parameter estimates - models with homogeneous firms: K=6-8

	Prior	$K=6$		$K=7$		$K=8$	
ζ	4.440 (0.47;16.86)	0.929 (0.52;1.84)	1.022	0.362 (0.19;0.83)	0.419	3.228 (1.67;7.23)	3.692
ρ_0	0.000 $(-8.22;8.22)$	0.000 $(-0.00;0.00)$	0.000	0.000 $(-0.00;0.00)$	0.000	0.000 $(-0.00;0.00)$	0.000
ρ_1	0.000 $(-8.22;8.22)$	1.289 (1.14;1.43)	1.289	1.429 (1.28;1.57)	1.428	1.338 (1.19;1.48)	1.337
ρ_2	0.000 $(-8.22;8.22)$	-0.320 $(-0.46;-0.17)$	-0.319	-0.453 $(-0.59;-0.31)$	-0.452	-0.372 $(-0.51;-0.23)$	-0.371
σ_m	0.059 (0.01;0.20)	0.005 (0.00;0.01)	0.005	0.005 (0.00;0.01)	0.005	0.005 (0.00;0.01)	0.005
δ_0	0.000 $(-8.22;8.22)$	0.004 $(-0.00;0.01)$	0.004	0.003 $(-0.00;0.01)$	0.004	0.002 $(-0.01;0.01)$	0.002
δ_1	0.000 $(-8.22;8.22)$	-0.491 $(-0.63;-0.34)$	-0.488	0.065 $(-0.15;0.32)$	0.071	-0.460 $(-0.60;-0.29)$	-0.455
δ_2	0.000 $(-8.22;8.22)$	0.469 (0.33;0.62)	0.472	0.135 $(-0.03;0.33)$	0.141	0.528 (0.39;0.69)	0.533
σ_n	0.059 (0.01;0.20)	0.202 (0.09;0.40)	0.218	0.216 (0.09;0.42)	0.230	0.155 (0.07;0.32)	0.169
log MD	–	737.782	–	781.326	–	717.135	–

Note: The first column gives the medians for the marginal prior distributions. For each specification, the first and second columns give, respectively, the medians and means of the marginal posterior distributions. Numbers in parentheses correspond to the 5th and 95th percentiles. The logarithm of the Marginal Posterior (log MD, bottom row) is approximated with Geweke's (1999) modified harmonic mean.

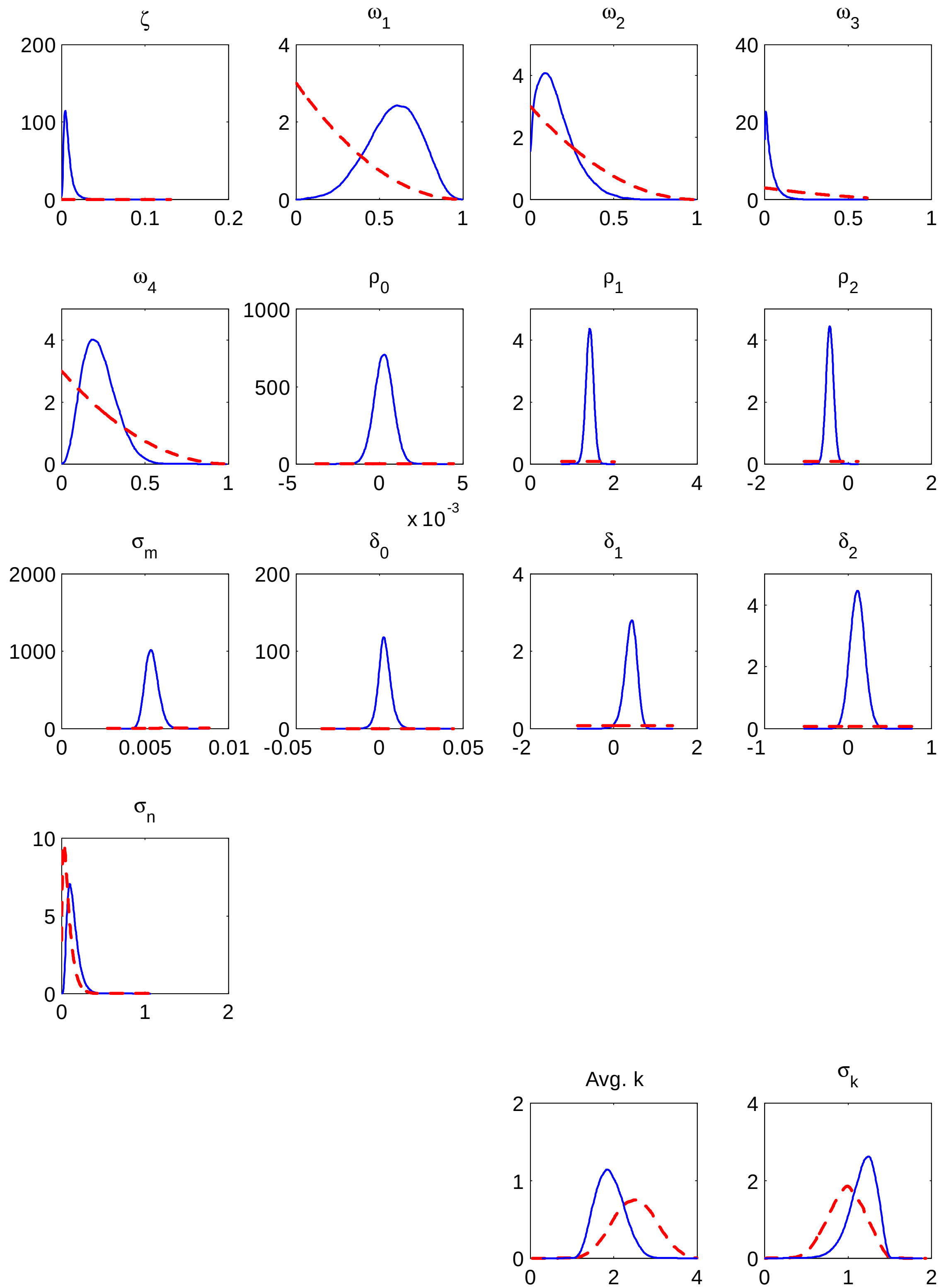

Figure 8: Flat prior (dashed line) and posterior (solid line) marginal distributions, $K = 4$

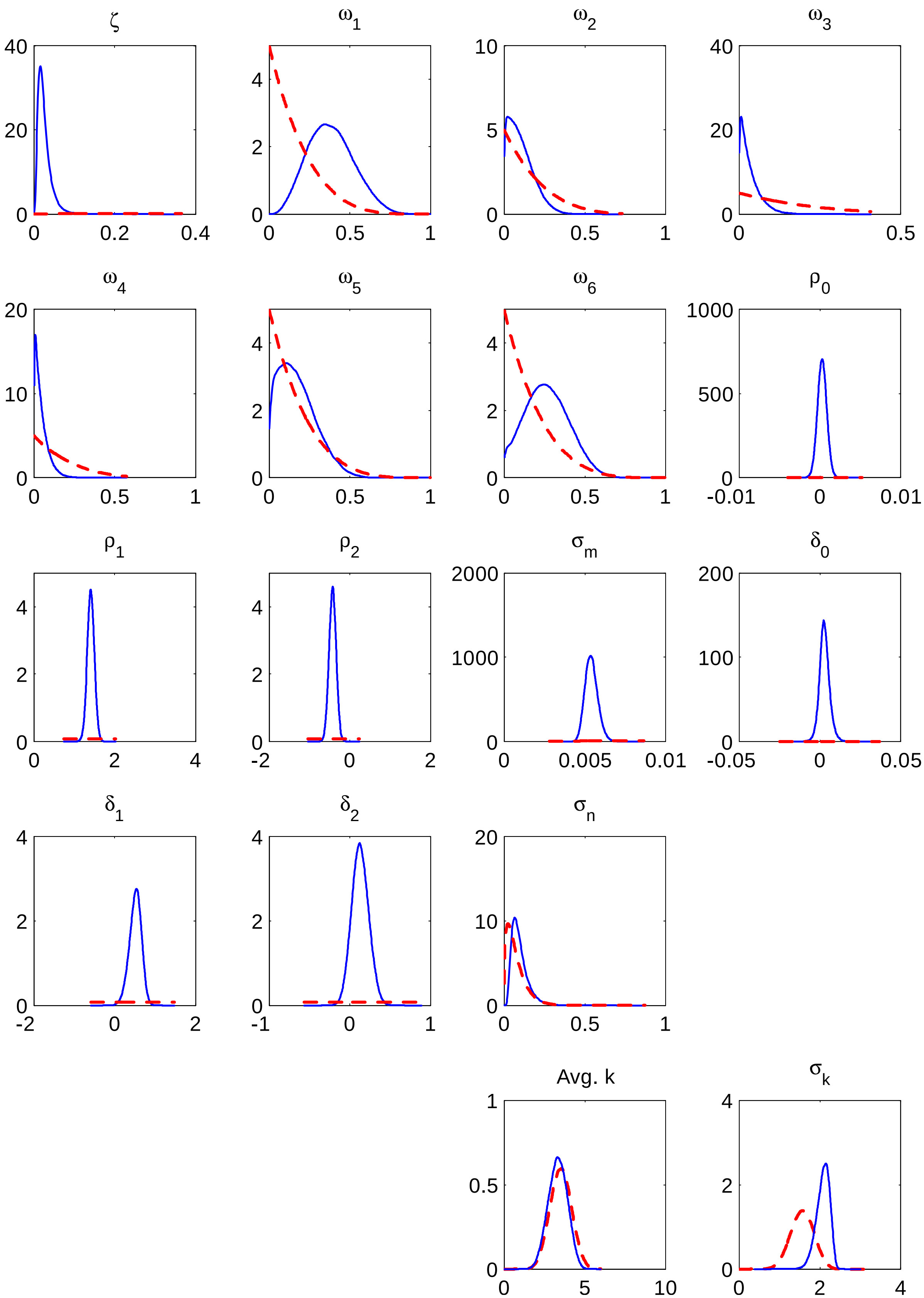

Figure 9: Flat prior (dashed line) and posterior (solid line) marginal distributions, $K = 6$

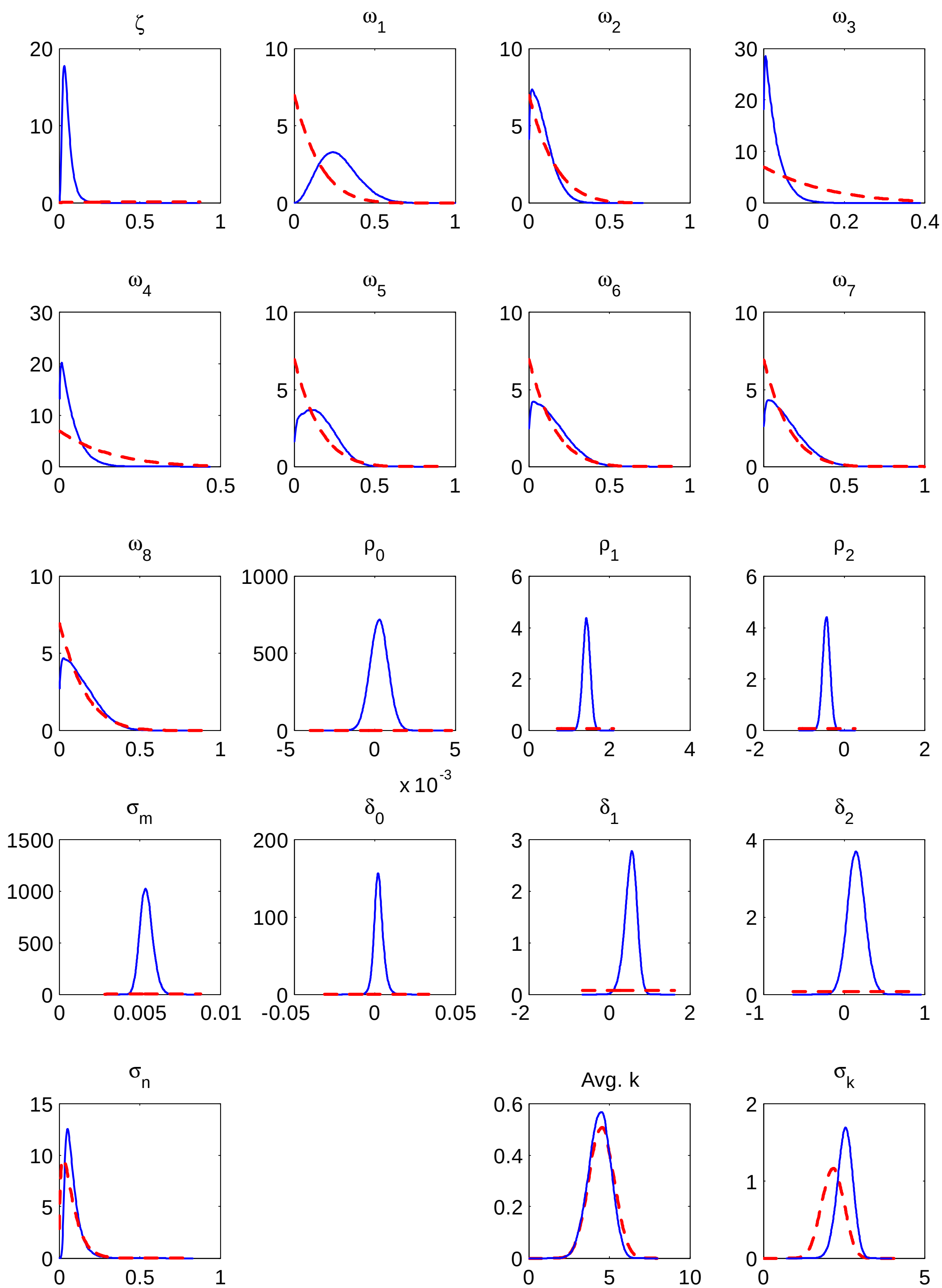

Figure 10: Flat prior (dashed line) and posterior (solid line) marginal distributions, $K = 8$

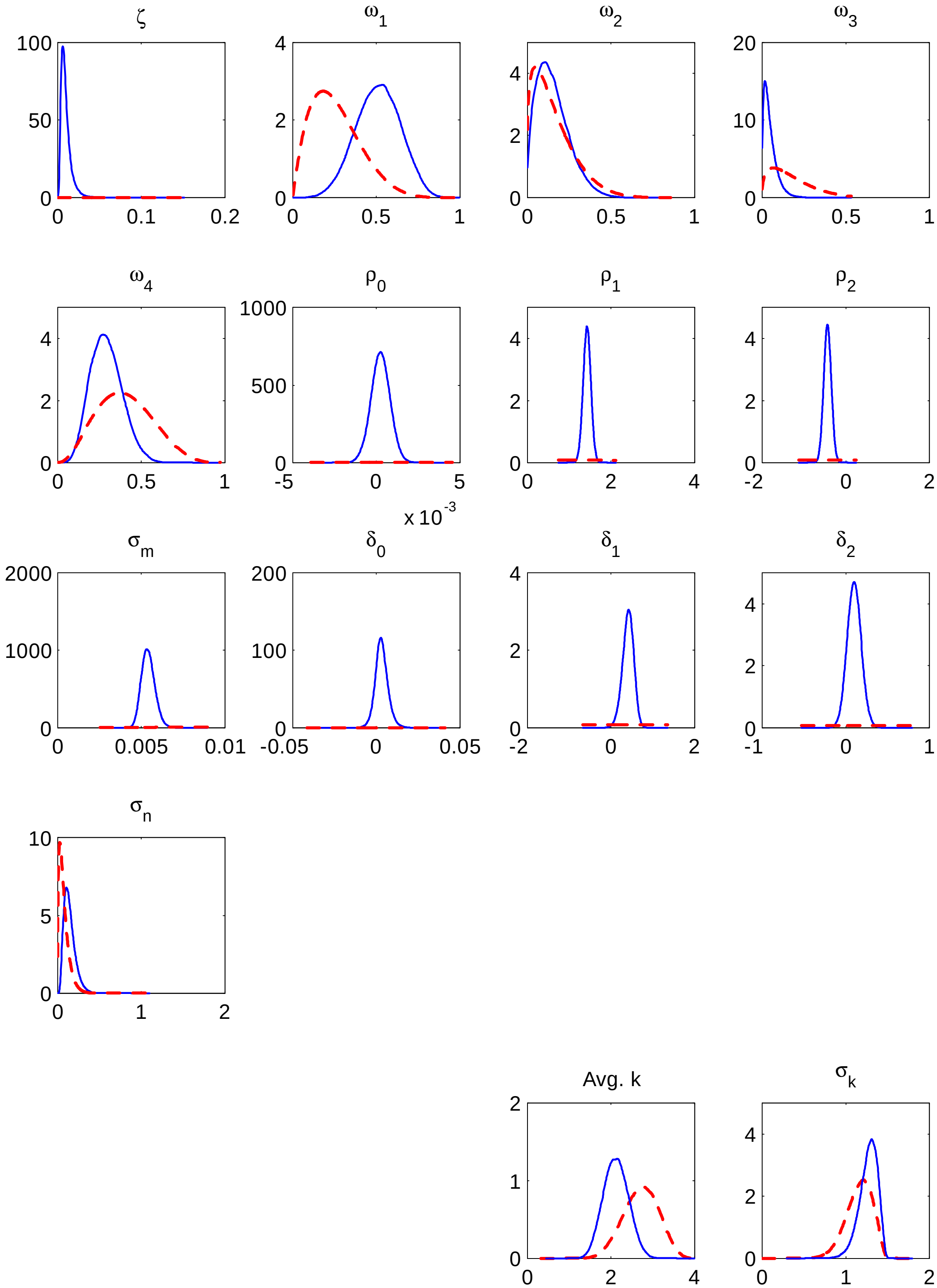

Figure 11: Nakamura-Steinsson loose prior (dashed line) and posterior (solid line) marginal distributions, $K = 4$

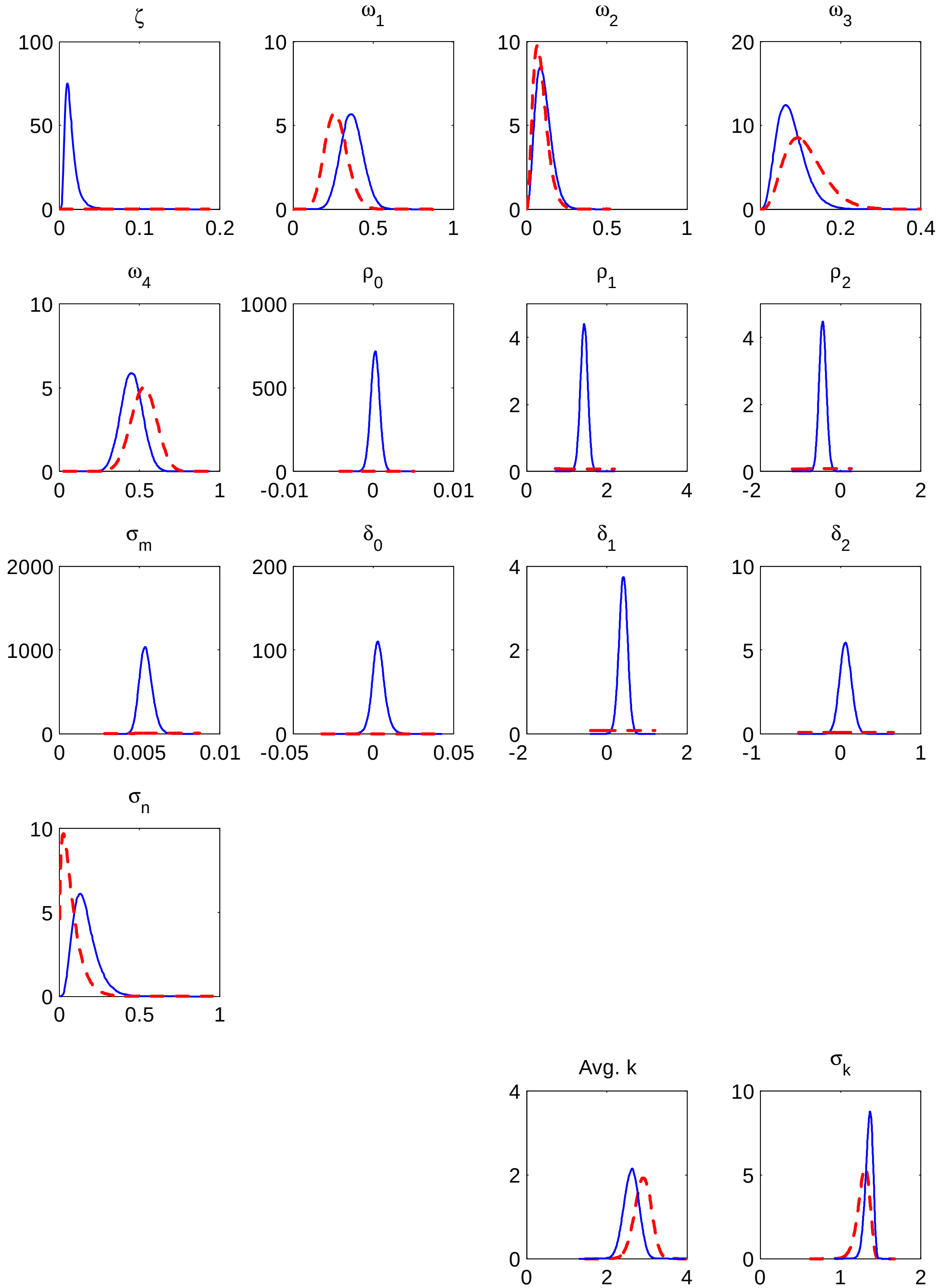

Figure 12: Nakamura-Steinsson tight prior (dashed line) and posterior (solid line) marginal distributions, $K = 4$

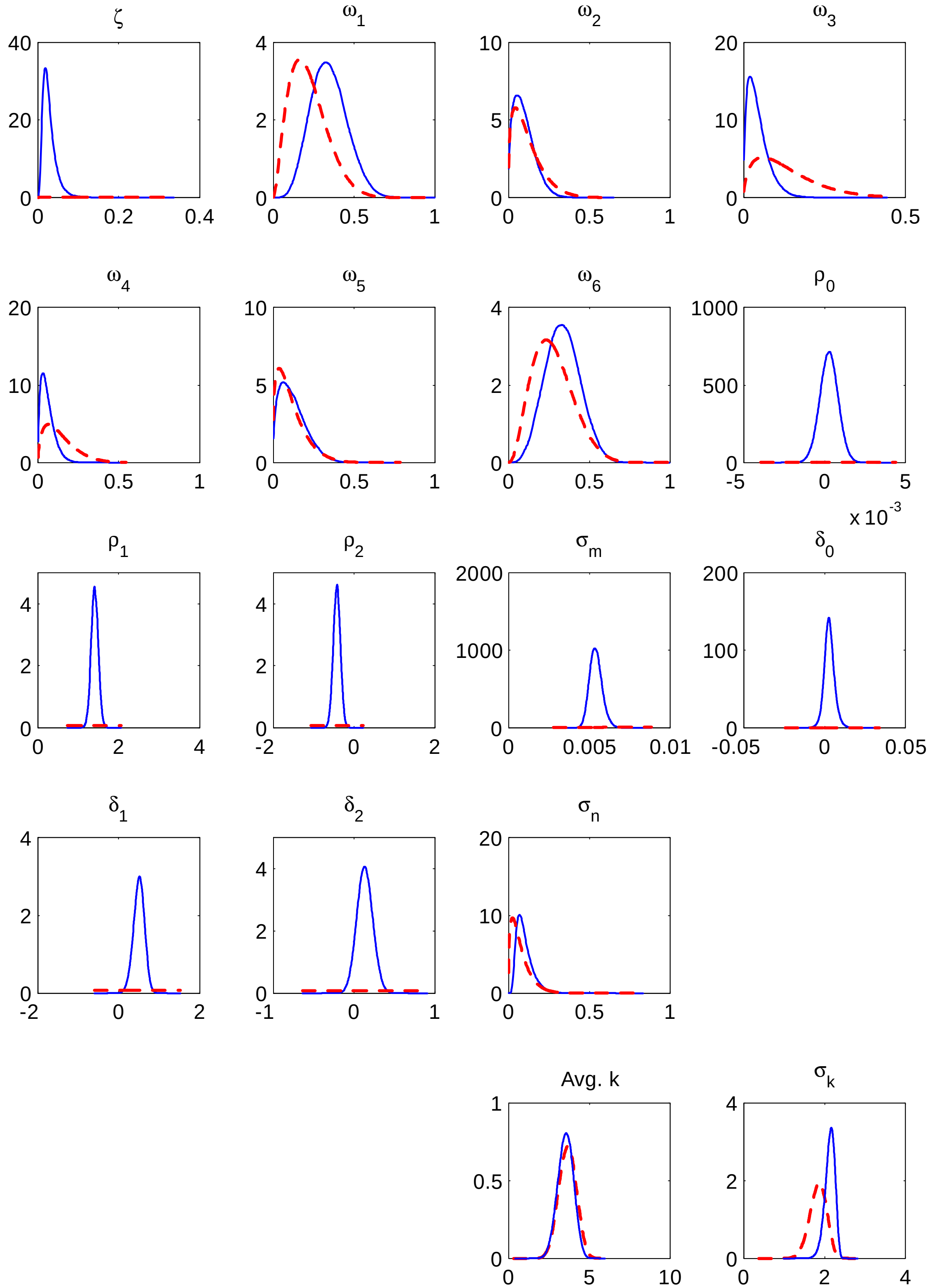

Figure 13: Nakamura-Steinsson loose prior (dashed line) and posterior (solid line) marginal distributions, $K = 6$

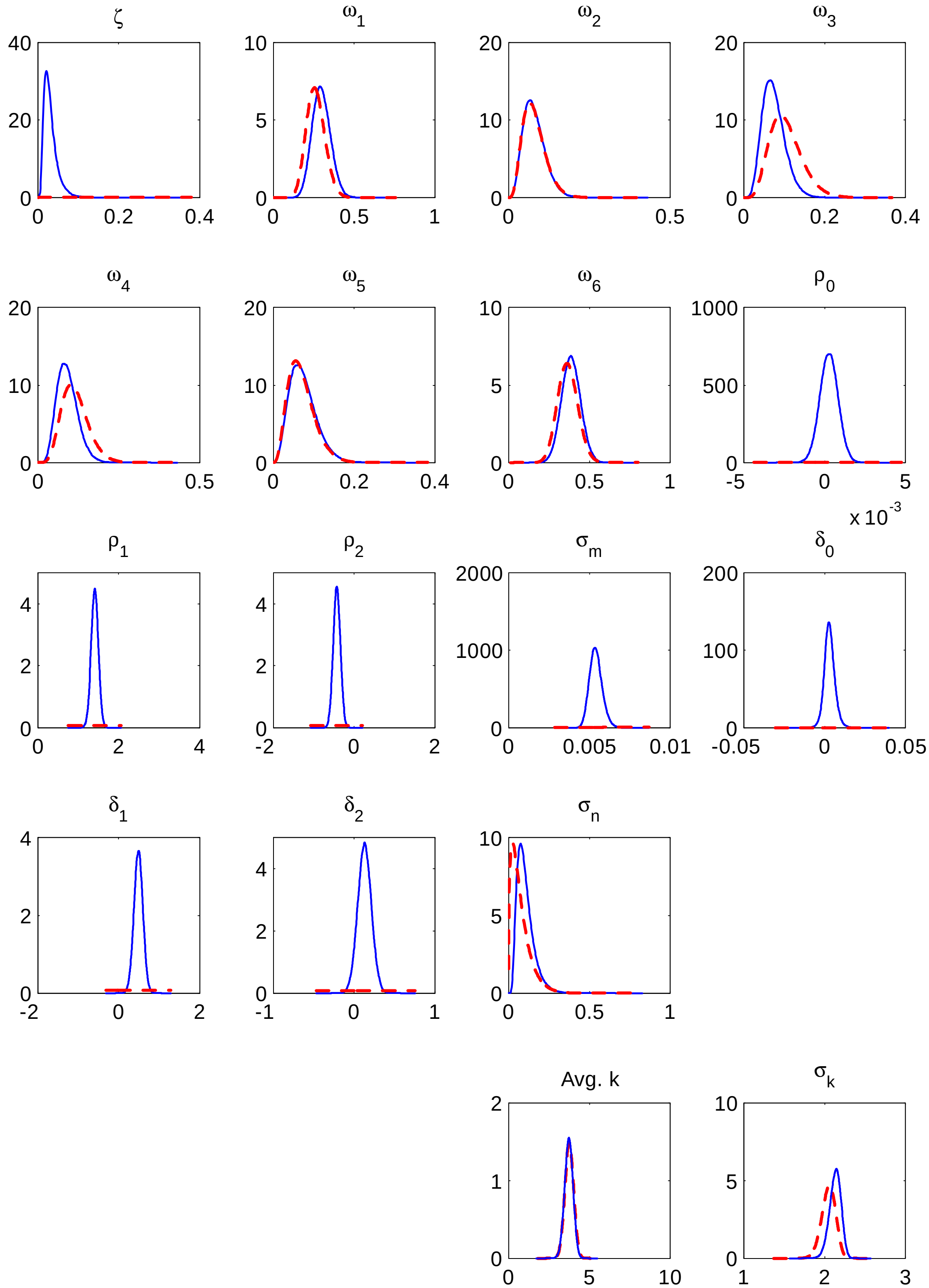

Figure 14: Nakamura-Steinsson tight prior (dashed line) and posterior (solid line) marginal distributions, $K = 6$

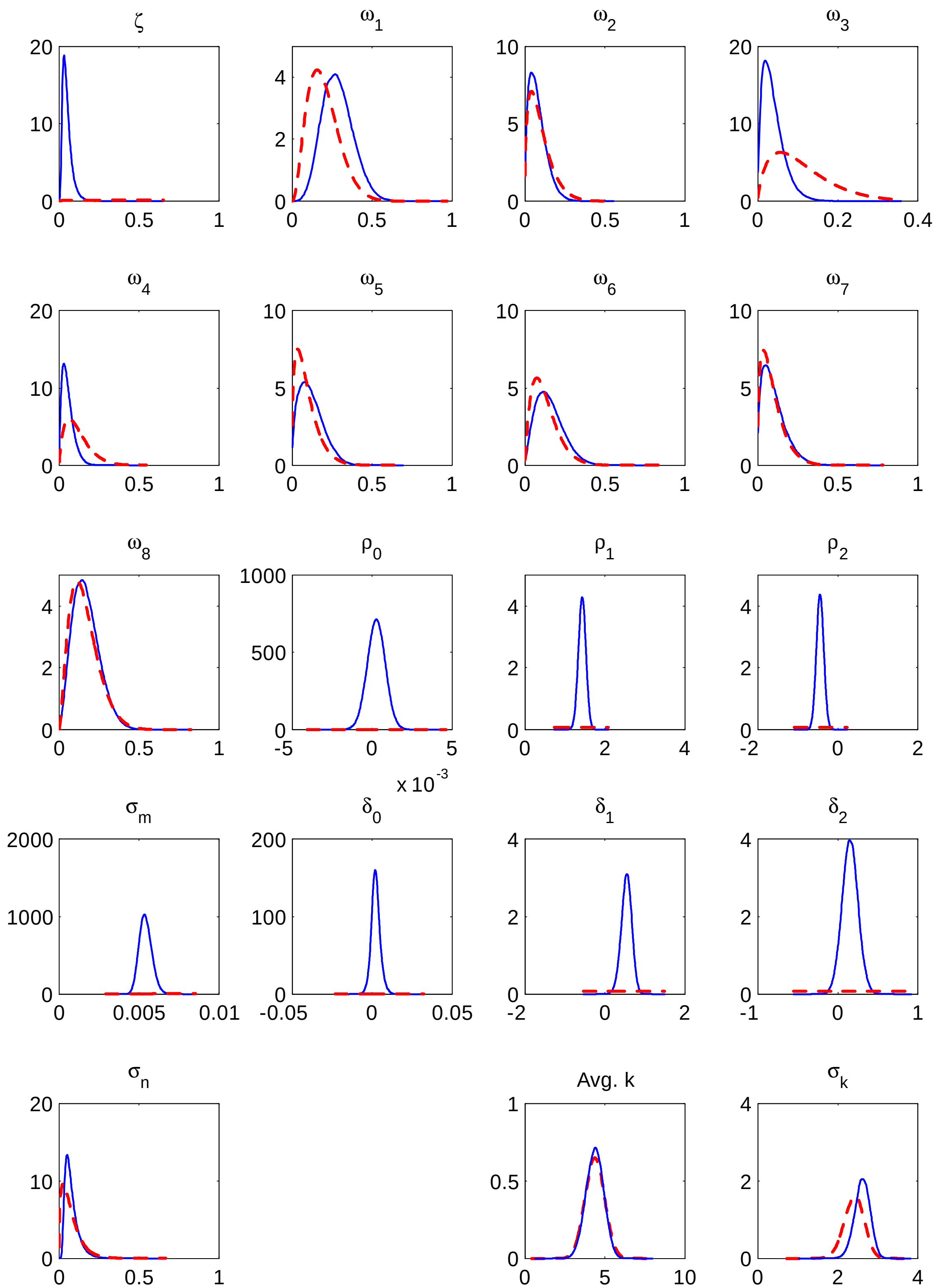

Figure 15: Nakamura-Steinsson loose prior (dashed line) and posterior (solid line) marginal distributions, $K = 8$

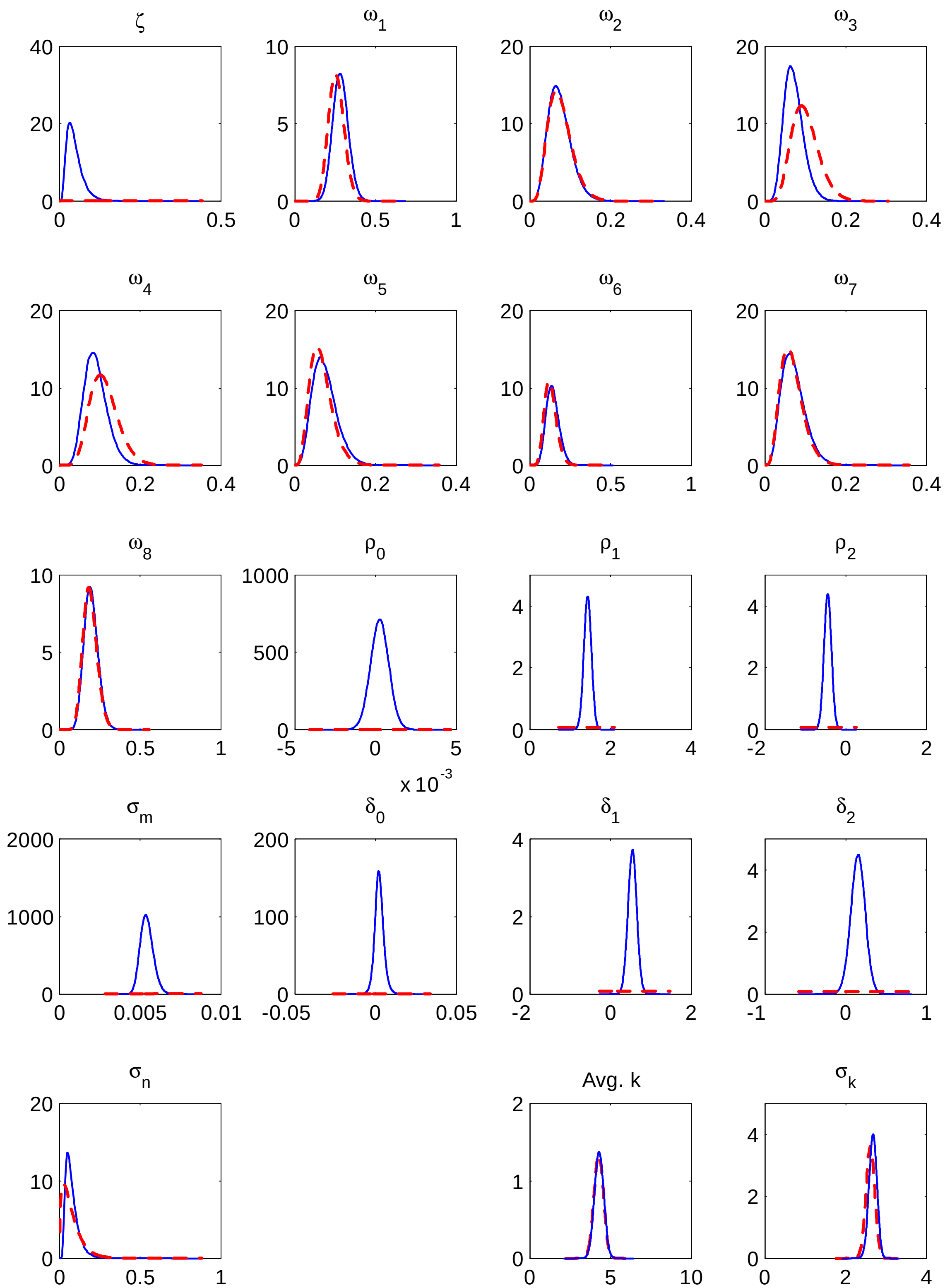

Figure 16: Nakamura-Steinsson tight prior (dashed line) and posterior (solid line) marginal distributions, $K = 8$

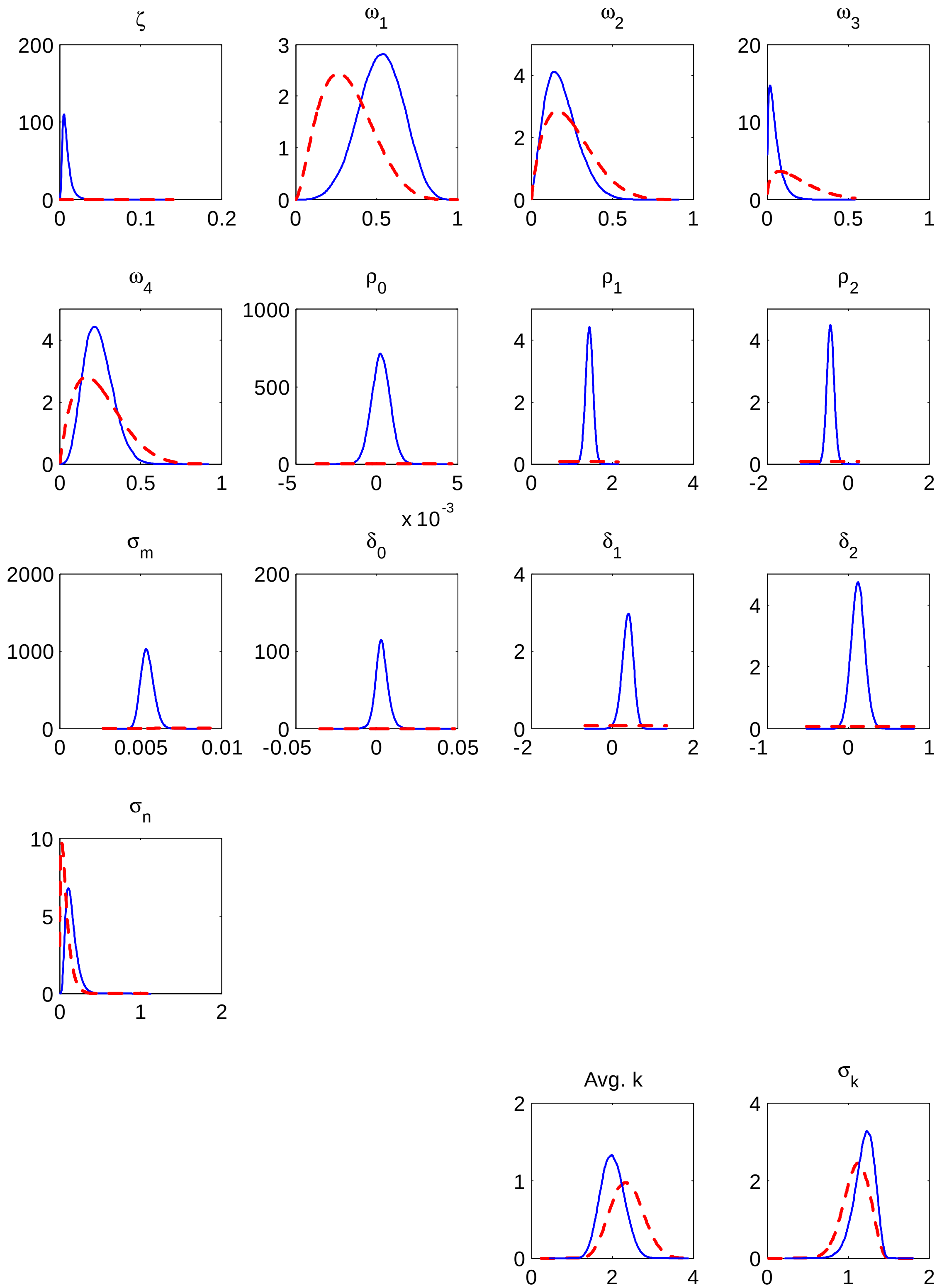

Figure 17: Bils-Klenow loose prior (dashed line) and posterior (solid line) marginal distributions, $K = 4$

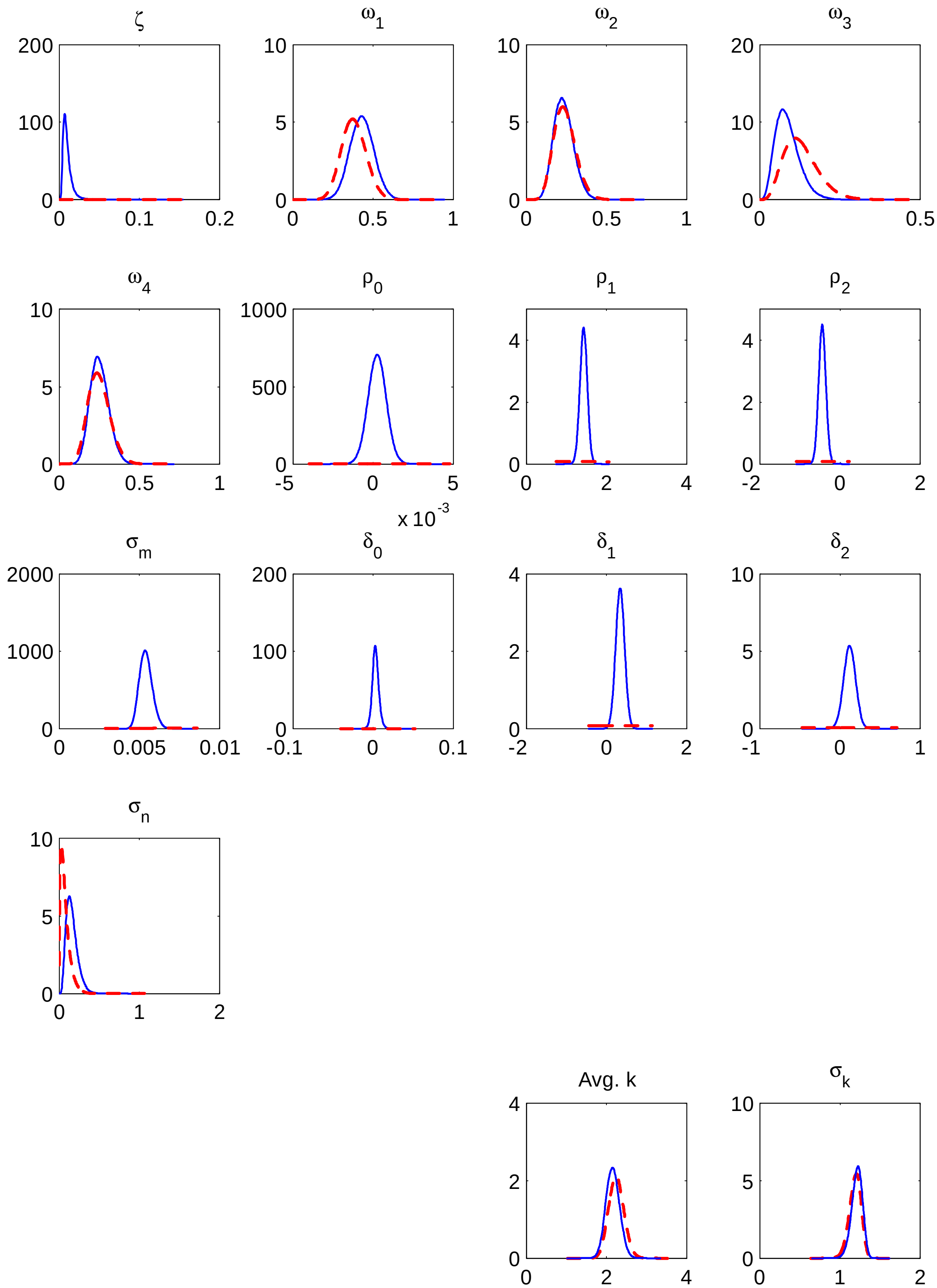

Figure 18: Bils-Klenow tight prior (dashed line) and posterior (solid line) marginal distributions, $K = 4$

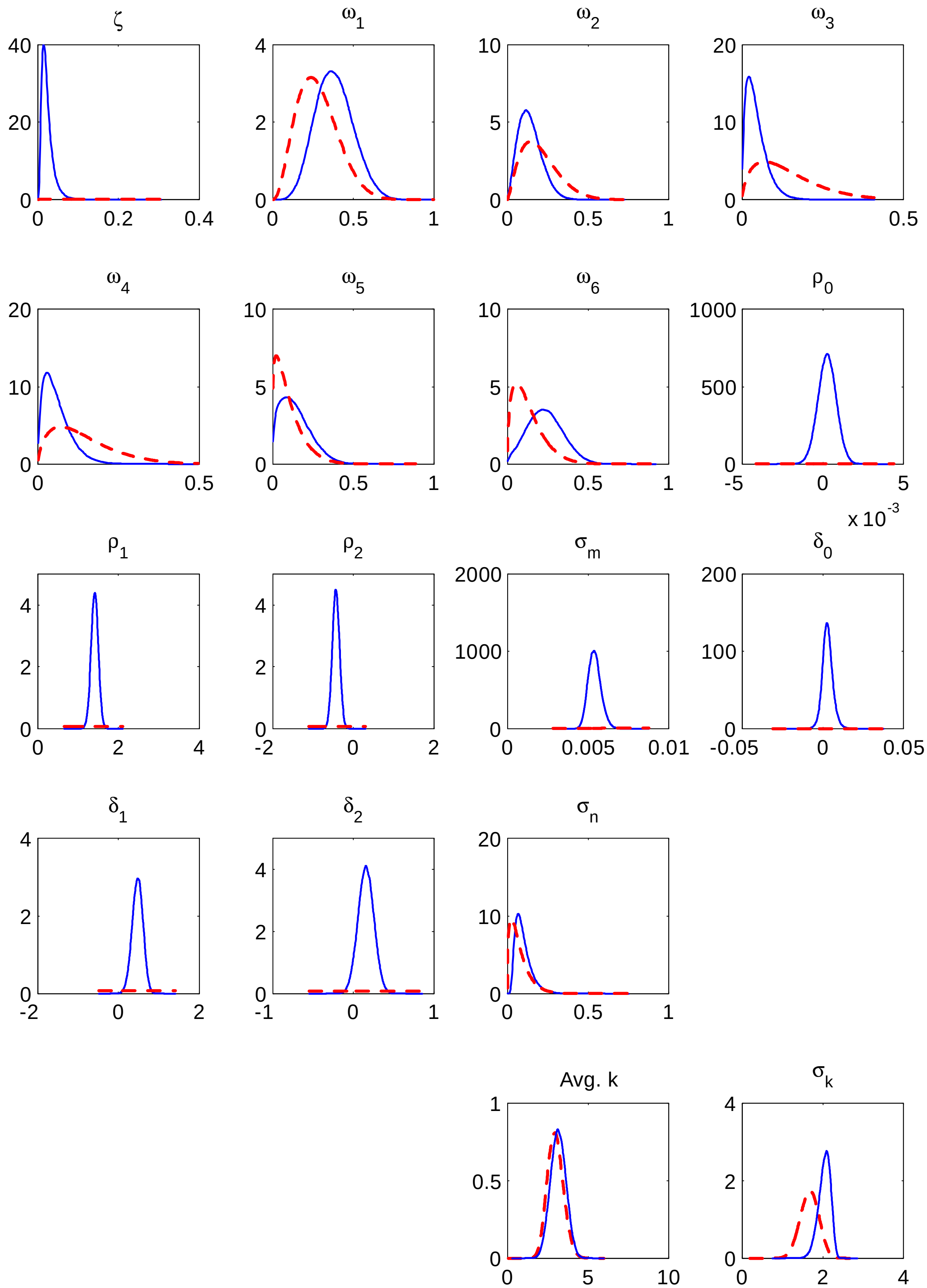

Figure 19: Bils-Klenow loose prior (dashed line) and posterior (solid line) marginal distributions, $K = 6$

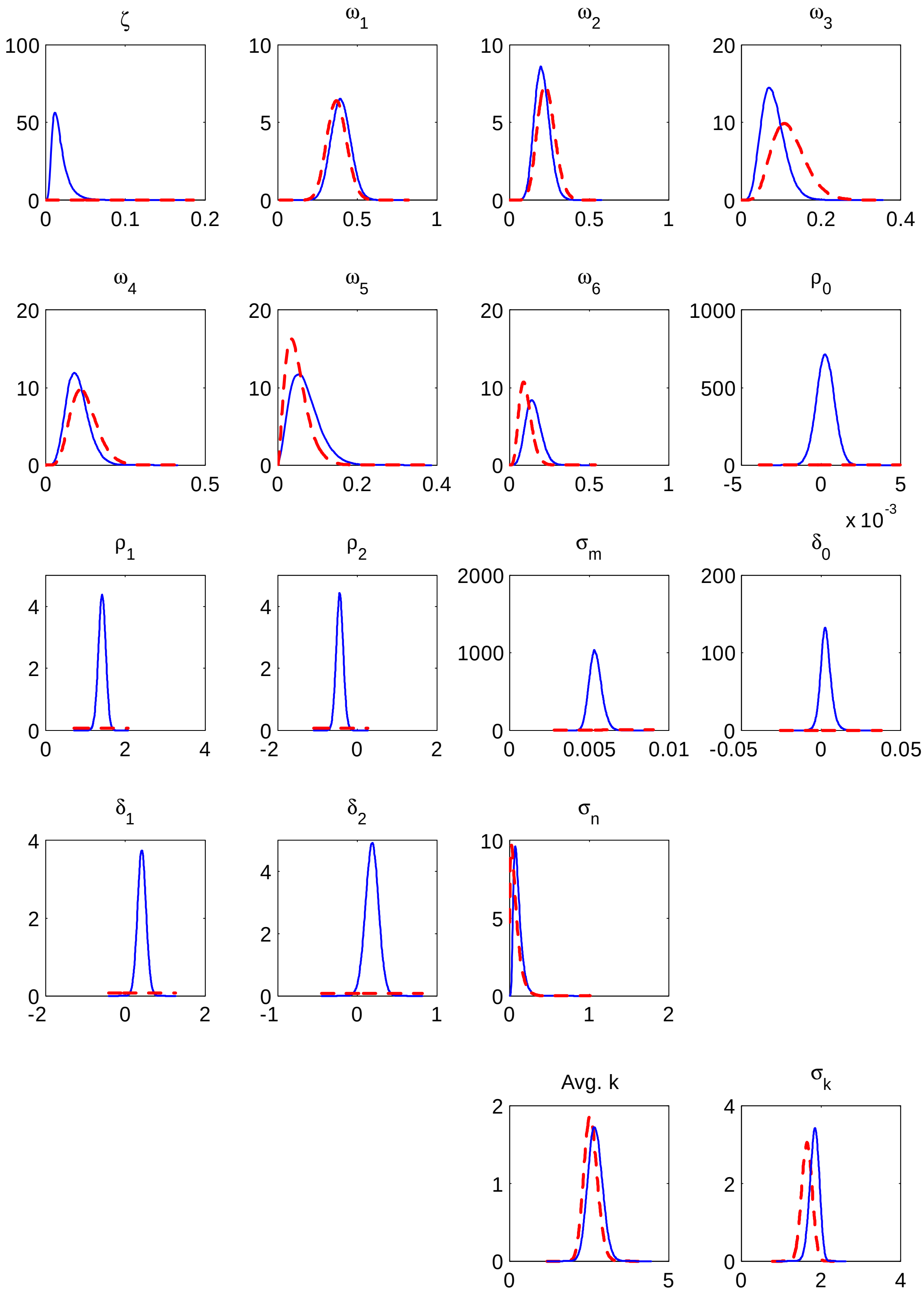

Figure 20: Bils-Klenow tight prior (dashed line) and posterior (solid line) marginal distributions, $K = 6$

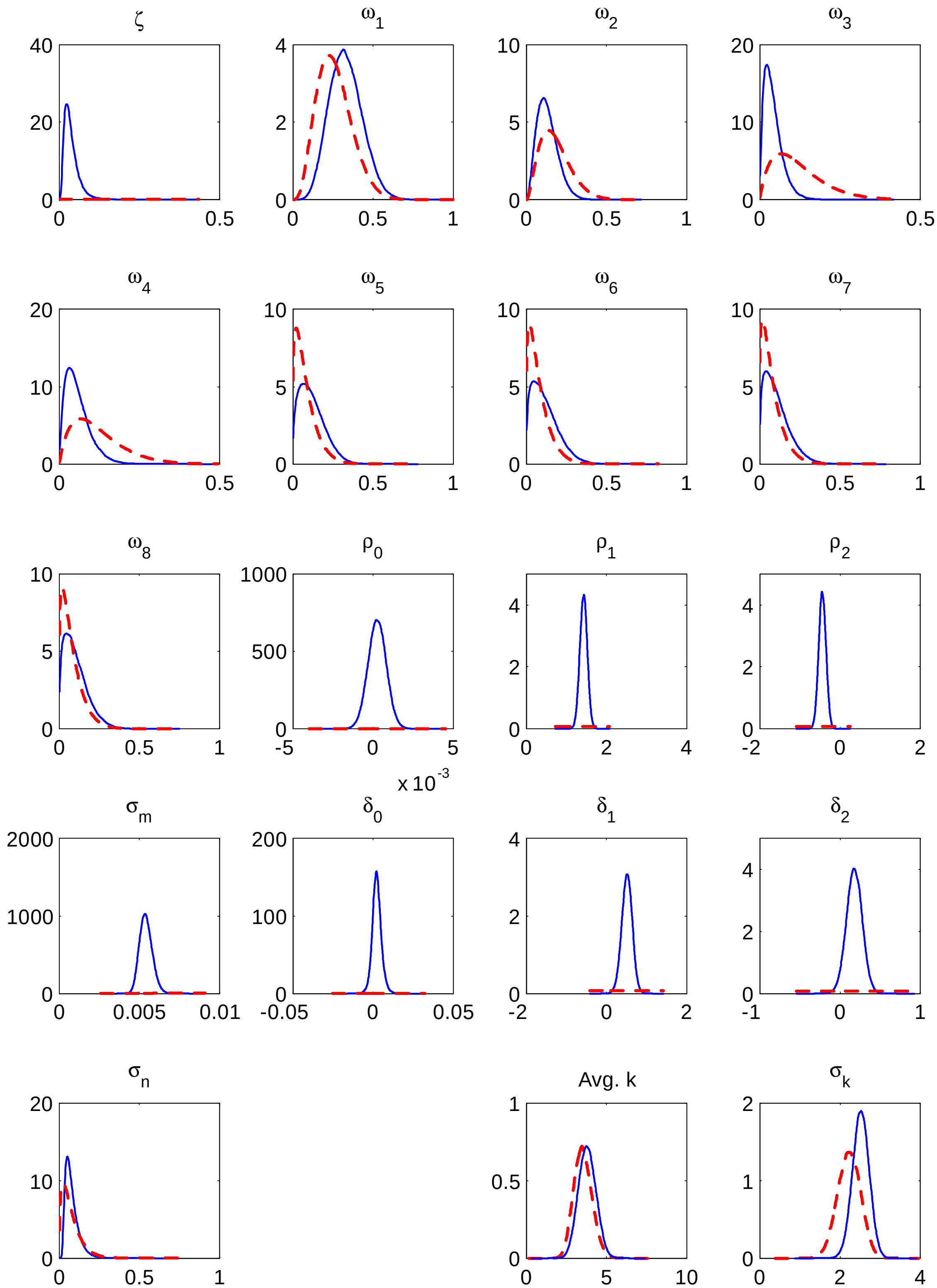

Figure 21: Bils-Klenow loose prior (dashed line) and posterior (solid line) marginal distributions, $K = 8$

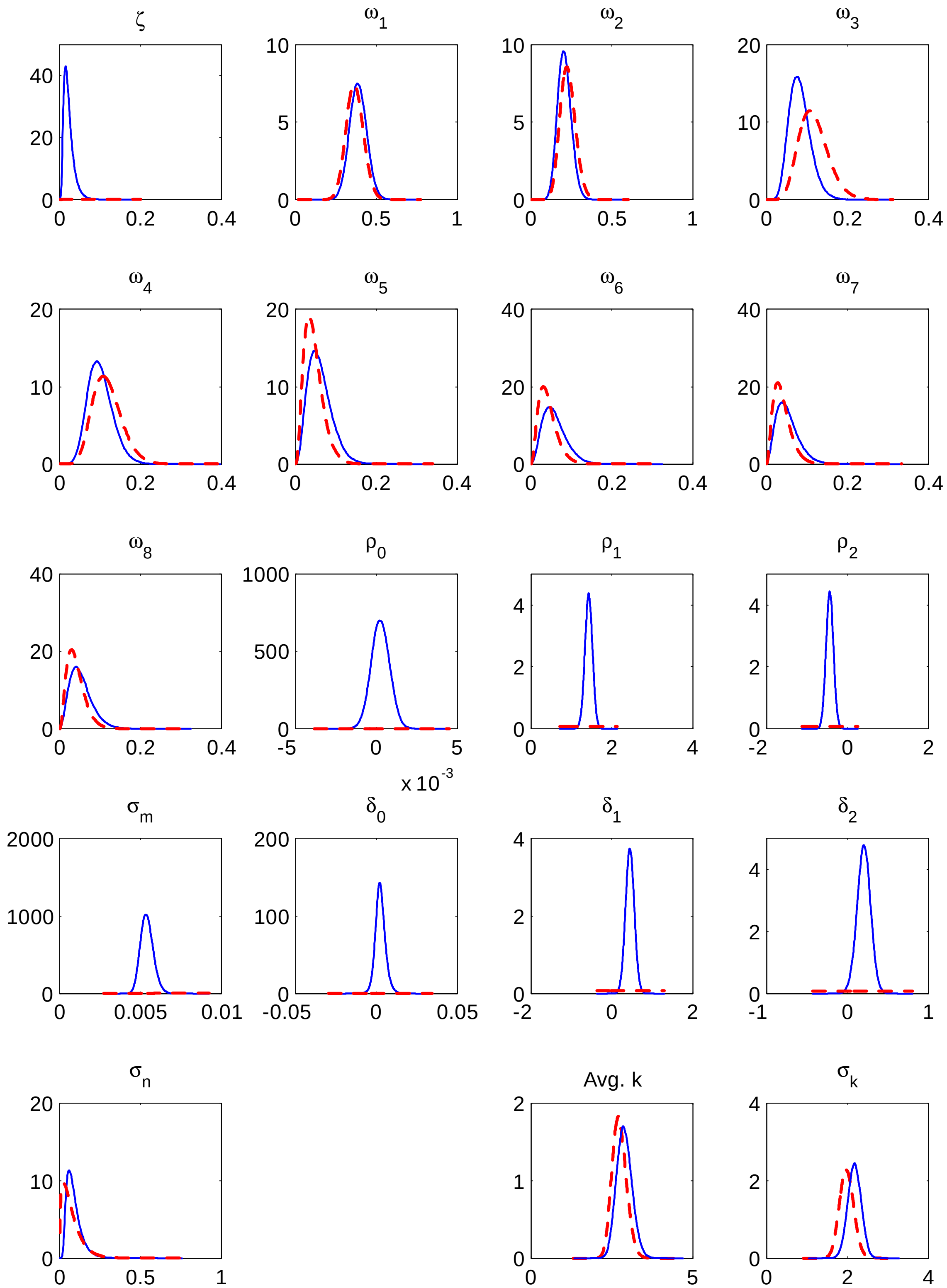

Figure 22: Bils-Klenow tight prior (dashed line) and posterior (solid line) marginal distributions, $K = 8$

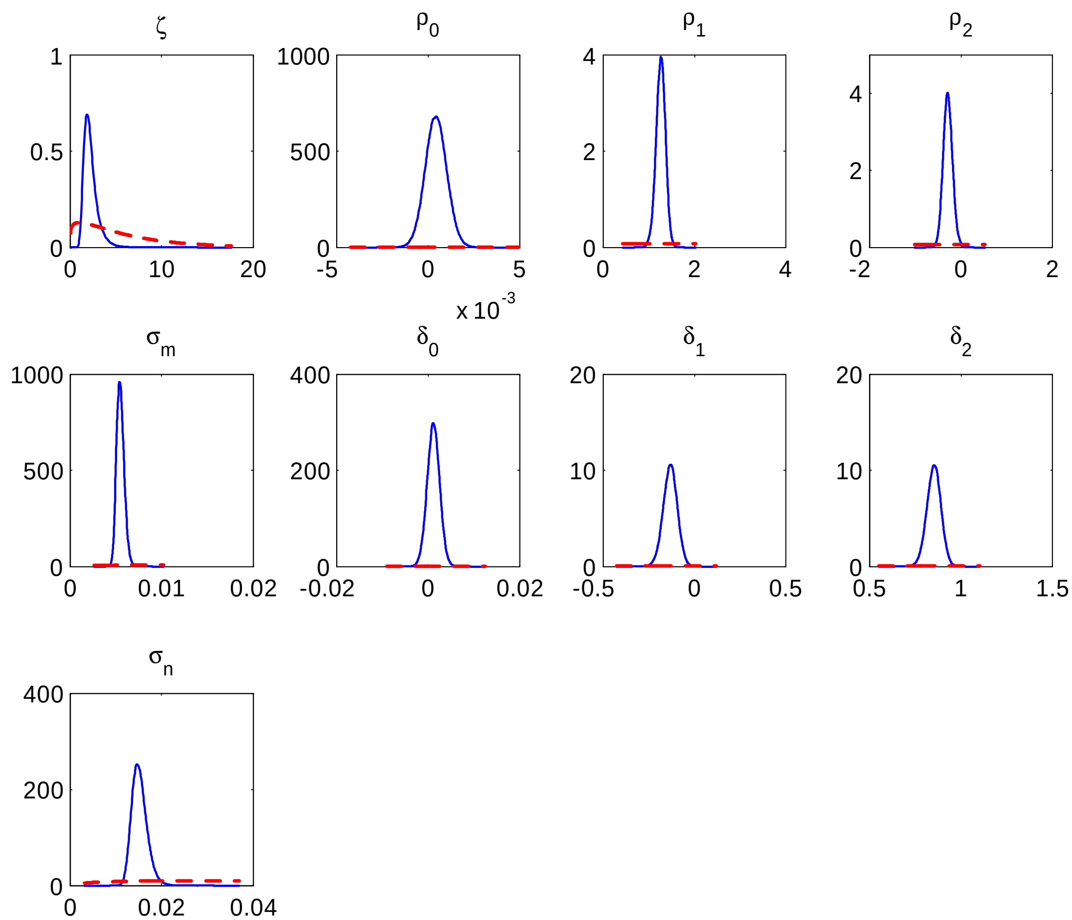

Figure 23: Prior (dashed line) and posterior (solid line) marginal distributions, 2-quarter price spells ($\omega_2 = 1$)

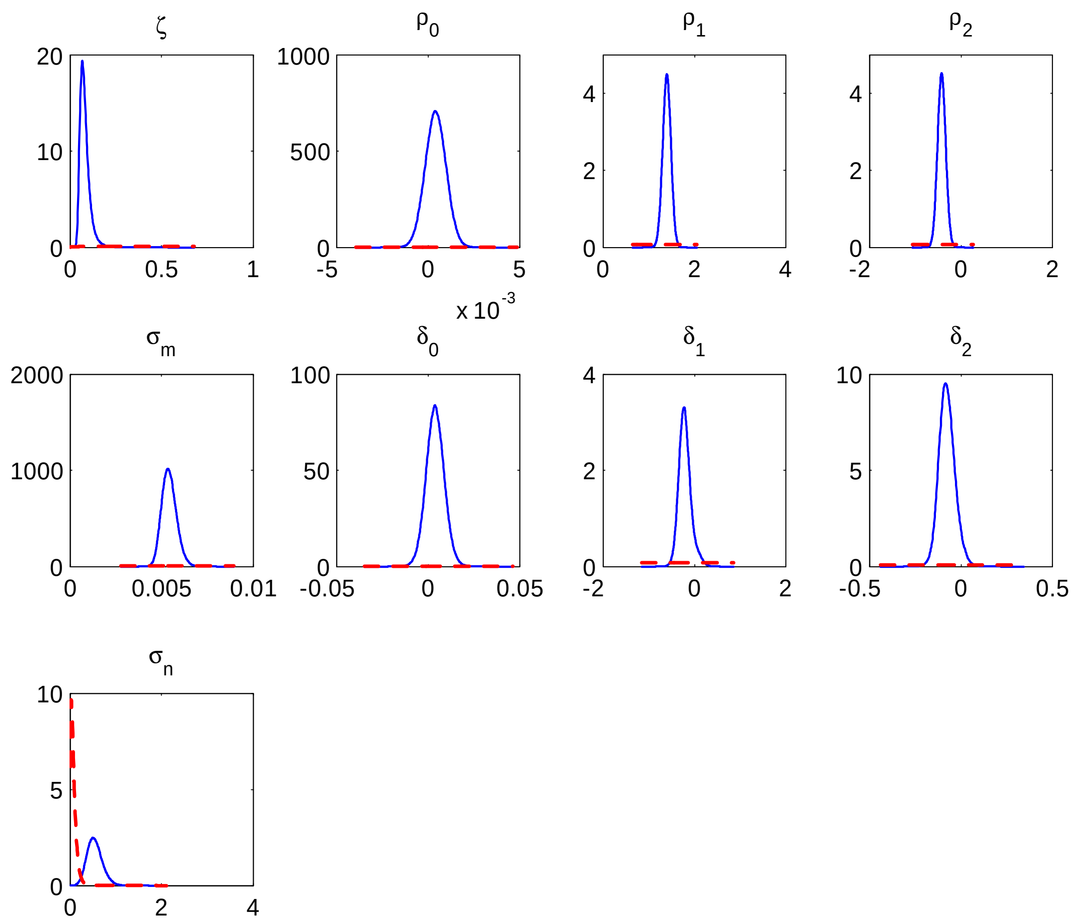

Figure 24: Prior (dashed line) and posterior (solid line) marginal distributions, 3-quarter price spells ($\omega_3 = 1$)

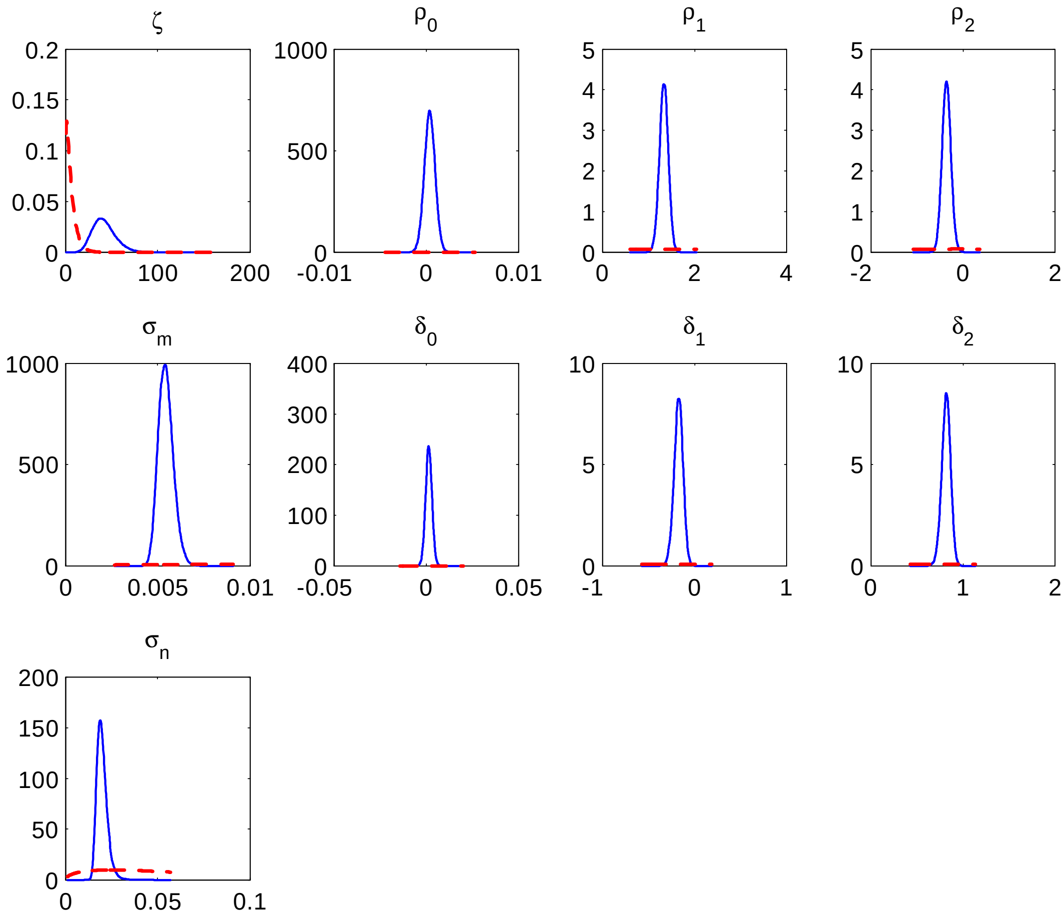

Figure 25: Prior (dashed line) and posterior (solid line) marginal distributions, 4-quarter price spells ($\omega_4 = 1$)

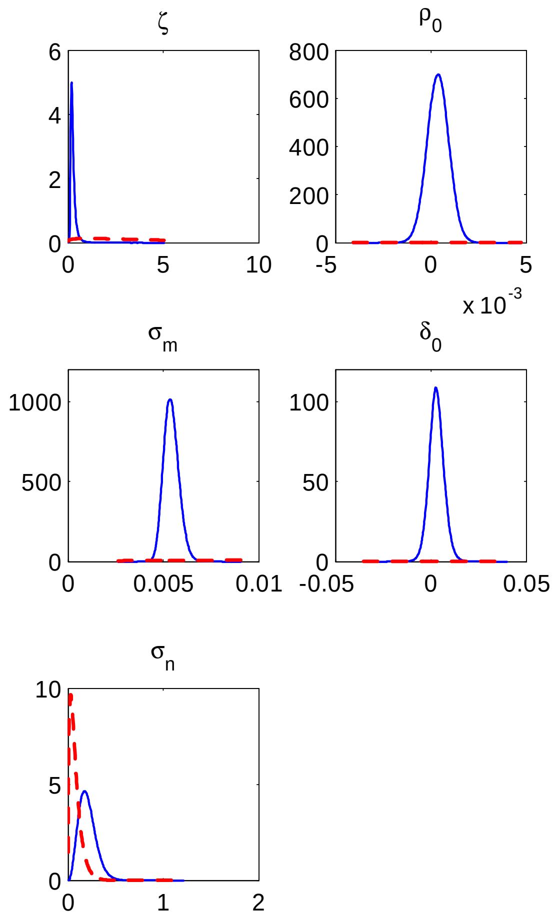

Figure 26: Prior (dashed line) and posterior (solid line) marginal distributions, 5-quarter price spells ($\omega_5 = 1$)

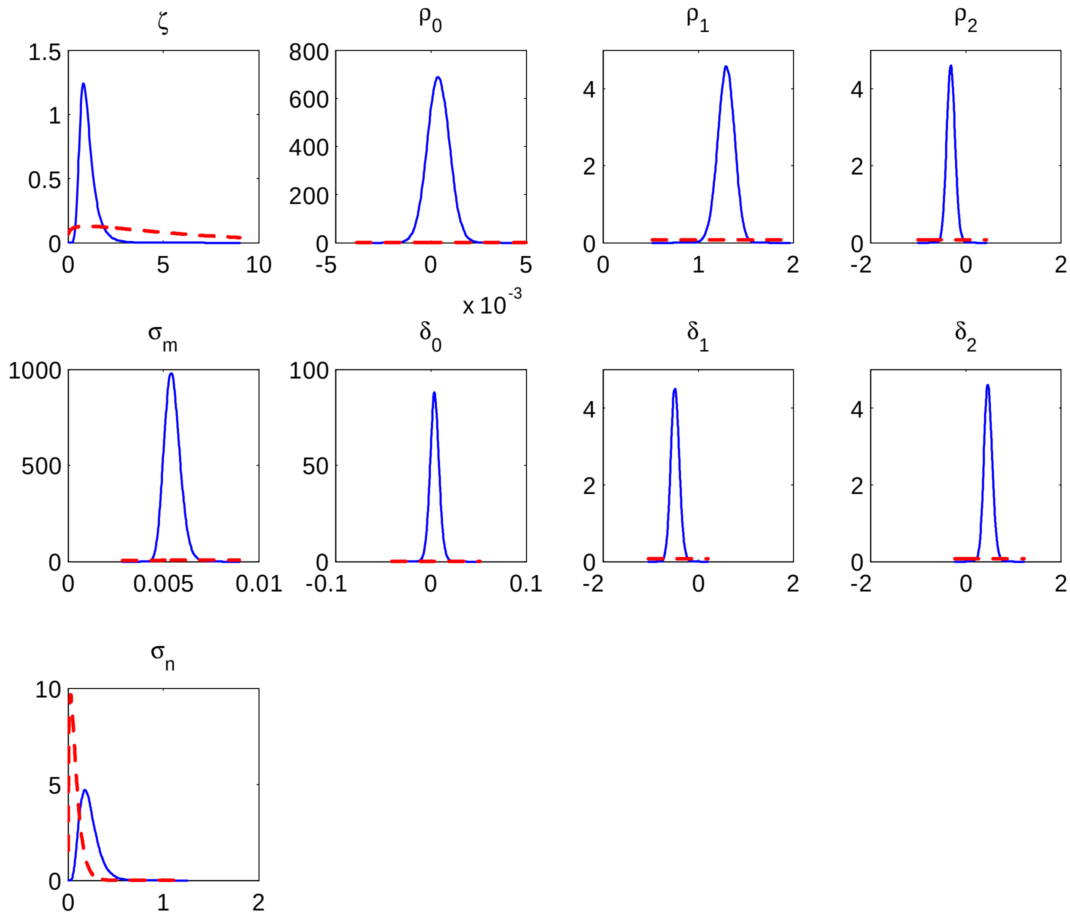

Figure 27: Prior (dashed line) and posterior (solid line) marginal distributions, 6-quarter price spells ($\omega_6 = 1$)

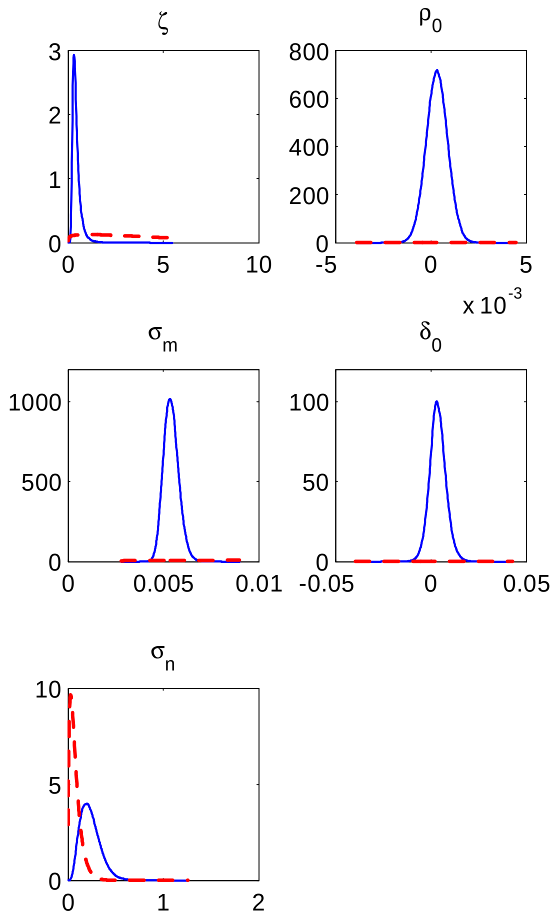

Figure 28: Prior (dashed line) and posterior (solid line) marginal distributions, 7-quarter price spells ($\omega_7 = 1$)

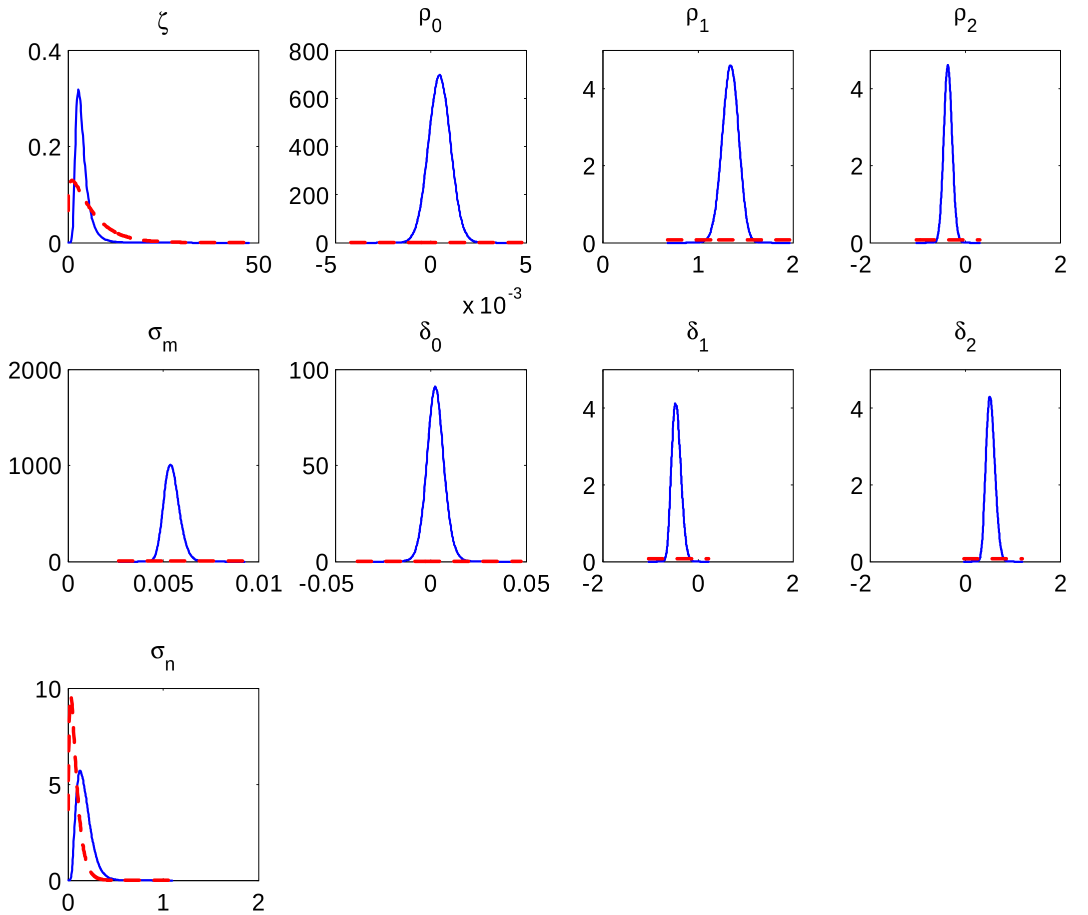

Figure 29: Prior (dashed line) and posterior (solid line) marginal distributions, 8-quarter price spells ($\omega_8 = 1$)