Sectoral Price Data and Models of Price Setting*

Bartosz Maćkowiak
European Central Bank and CEPR

Emanuel Moench
Federal Reserve Bank of New York

Mirko Wiederholt
Northwestern University

First draft: December 2008. This draft: June 2009.

Abstract

We estimate impulse responses of sectoral price indexes to aggregate shocks and to sector-specific shocks. In the median sector, 100 percent of the long-run response of the sectoral price index to a sector-specific shock occurs in the month of the shock. The standard Calvo model and the standard sticky-information model can match this finding only under extreme assumptions concerning the profit-maximizing price. The rational-inattention model of Maćkowiak and Wiederholt (2009a) can match this finding without an extreme assumption concerning the profit-maximizing price. Furthermore, there is little variation across sectors in the speed of response of sectoral price indexes to sector-specific shocks. The rational-inattention model matches this finding, while the Calvo model predicts too much cross-sectional variation.

JEL: C11, D21, D83, E31.

Keywords: Bayesian dynamic factor model, Calvo model, menu cost, sticky information, rational inattention.

*We thank for comments Sergio Rebelo, an anonymous referee, our discussants at the JME conference in Gerzenee Ricardo Reis and Karl Walentin, other participants in this conference, Gianni Amisano, Domenico Giannone, Marek Jarociński, Ariel Shwayder and seminar participants at Bundesbank and Humboldt University Berlin. The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the European Central Bank or the Federal Reserve. E-mail addresses: bartosz.mackowiak@ecb.int, emanuel.moench@ny.frb.org, m-wiederholt@northwestern.edu.

1 Introduction

Over the last twenty years, there has been a surge in research on macroeconomic models with price stickiness. In these models, price stickiness arises either from adjustment costs (e.g. the Calvo model and the menu cost model) or from some form of information friction (e.g. the sticky-information model and the rational-inattention model). Models of price stickiness are often evaluated by looking at aggregate data. Recently models of price stickiness have been evaluated by looking at micro data. This paper evaluates models of price stickiness by studying sectoral data. A statistical model for sectoral inflation rates is estimated and used to compute impulse responses of sectoral price indexes to aggregate shocks and to sector-specific shocks. The paper proceeds by analyzing whether different models of price setting can match the empirical impulse responses.

The statistical model that is estimated is the following. The inflation rate in a sector equals the sum of two components, an aggregate component and a sector-specific component. The parameters in the aggregate component and in the sector-specific component may differ across sectors. An innovation in the aggregate component may affect the inflation rates in all sectors. An innovation in the sector-specific component affects only the inflation rate in this sector. The statistical model is estimated using monthly sectoral consumer price data from the U.S. economy for the period 1985-2005. The data are compiled by the Bureau of Labor Statistics (BLS). From the estimated statistical model, one can compute impulse responses of the price index for a sector to an innovation in the aggregate component and to an innovation in the sector-specific component.

The median impulse responses have the following shapes. After a sector-specific shock, 100 percent of the long-run response of the sectoral price index occurs in the month of the shock, and the response equals the long-run response in all months following the shock. By contrast, after an aggregate shock, only 15 percent of the long-run response of the sectoral price index occurs in the month of the shock, and the response gradually approaches the long-run response in the months following the shock. Another way of summarizing the median impulse responses is as follows. The sector-specific component of the sectoral inflation rate is essentially a white noise process, while the aggregate component of the sectoral inflation rate is positively autocorrelated.

The paper proceeds by studying whether the standard Calvo model, the standard sticky-information model, and the rational-inattention model developed in Maćkowiak and Wiederholt (2009a) can match the median impulse response of sectoral price indexes to sector-specific shocks. The focus is on the response to sector-specific shocks, because it is well known that all three models can match the median impulse response of sectoral price indexes to aggregate shocks, for reasonable parameter values. In fact, the models have been developed to explain the slow response of prices to aggregate shocks. What we find interesting is that these models emphasize different reasons for why the response of prices to aggregate shocks is slow: infrequent price adjustment (Calvo model) and information frictions (sticky-information model and rational-inattention model). This paper evaluates the plausibility of the reason emphasized by a given model by asking whether the model can match the median impulse response of sectoral price indexes to sector-specific shocks.

Recall that this impulse response looks like the impulse response function of a random walk: the sectoral price index jumps on impact of a sector-specific shock, and stays there. Proposition 1 shows that the standard Calvo model can match the median impulse response of sectoral price indexes to sector-specific shocks only under an extreme assumption concerning the response of the profit-maximizing price to sector-specific shocks. After a sector-specific shock, the profit-maximizing price needs to jump by about $\left(1/\lambda^2\right) x$ in the month of the shock, and then has to jump back to x in the month following the shock to generate a response equal to x of the sectoral price index on impact and in all months following the shock. Here λ denotes the fraction of firms that can adjust their prices in a month. Proposition 2 provides a similar, though less extreme, result for the standard sticky-information model developed in Mankiw and Reis (2002). After a sector-specific shock, the profit-maximizing price needs to jump by $(1/\lambda)\, x$ in the month of the shock, and then has to decay slowly to x to generate a response equal to x of the sectoral price index on impact and in all months following the shock. Here λ denotes the fraction of firms that can update their pricing plans in a month. By contrast, the rational-inattention model developed in Maćkowiak and Wiederholt (2009a) matches the median impulse response of sectoral price indexes to sector-specific shocks without an extreme assumption concerning the response of the profit-maximizing price to sector-specific shocks. The reason is simple.

According to the estimated statistical model, sector-specific shocks are on average much larger than aggregate shocks. Under these circumstances, the theoretical model predicts that decision-makers in firms pay significantly more attention to sector-specific conditions than to aggregate conditions, implying that prices respond quickly to sector-specific shocks and slowly to aggregate shocks.

The different models of price setting are also evaluated on their ability to predict the right amount of variation across sectors in the speed of response of sectoral price indexes to sector-specific shocks. According to the estimated statistical model, there is little variation across sectors in the speed of response of sectoral price indexes to sector-specific shocks. It turns out that a multi-sector Calvo model calibrated to the sectoral monthly frequencies of price changes reported in Bils and Klenow (2004) predicts too much cross-sectional variation in the speed of response to sector-specific shocks. By contrast, the rational-inattention model developed in Maćkowiak and Wiederholt (2009a) correctly predicts little cross-sectional variation in the speed of response to sector-specific shocks. The reason is as follows. According to the theoretical model, decision-makers in firms in the median sector are already paying so much attention to sector-specific conditions that they track sector-specific conditions almost perfectly. Paying even more attention to sector-specific conditions has little effect on the speed of response of prices to sector-specific shocks.

This paper is related to Boivin, Giannoni, and Mihov (2009). They use a factor augmented vector autoregressive model to study sectoral data published by the Bureau of Economic Analysis (BEA) on personal consumption expenditure. Boivin, Giannoni, and Mihov (2009) find that sectoral price indexes respond quickly to sector-specific shocks and slowly to aggregate shocks, and that sector-specific shocks account for a dominant share of the variance in sectoral inflation rates. This paper differs from Boivin, Giannoni, and Mihov (2009) in several ways. First of all, the statistical model, estimation methodology, and dataset are different. Second, this paper characterizes the conditions under which the standard Calvo model, the standard sticky-information model, and the rational-inattention model developed in Maćkowiak and Wiederholt (2009a) can match the median impulse response of sectoral price indexes to sector-specific shocks. Third, this paper estimates the cross-sectional distribution of the speed of response to aggregate shocks and the cross-

sectional distribution of the speed of response to sector-specific shocks. These cross-sectional distributions are useful for evaluating models of price setting. Fourth, this paper studies the distribution of sector-specific shocks and discusses the relationship to recent menu cost models.

This paper is also related to Reis and Watson (2007a, 2007b) who use a dynamic factor model to study sectoral data published by the BEA on personal consumption expenditure. The focus of Reis and Watson (2007a, 2007b) is on estimating the numeraire (defined as a common component in prices that has an equiproportional effect on all prices). Furthermore, this paper is related to Kehoe and Midrigan (2007) who study data from Europe and the United States on sectoral real exchange rates. Kehoe and Midrigan (2007) find much less heterogeneity in the persistence of sectoral real exchange rates in the data than predicted by the Calvo model.

The statistical model in this paper belongs to the class of dynamic factor models. Dynamic factor models have been estimated using maximum-likelihood methods, non-parametric methods based on principal components, and Bayesian methods.[1] This paper uses Bayesian methods. Section 2 explains the contribution to the literature on estimation of dynamic factor models. Section 2 also describes how the statistical model and estimation methodology differ from the work of Boivin, Giannoni, and Mihov (2009).

The paper is organized as follows. Section 2 presents the statistical model and estimation methodology. Section 3 describes the data. Sections 4 and 5 present the results from the statistical model. Section 6 discusses robustness of the results. Section 7 studies whether the model of Calvo (1983), the sticky-information model of Mankiw and Reis (2002), and the rational-inattention model developed in Maćkowiak and Wiederholt (2009a) can match the estimated impulse responses. Section 8 concludes. Appendix A gives econometric details.

[1] Maximum likelihood estimation: in frequency domain (Geweke, 1977; Sargent and Sims, 1977; Geweke and Singleton, 1981); in time domain (Engle and Watson, 1981; Stock and Watson, 1989; Quah and Sargent, 1992; Reis and Watson, 2007a, 2007b); quasi-maximum likelihood in time domain (Doz, Giannone, and Reichlin, 2006). Non-parametric estimation based on principal components (Forni, Hallin, Lippi, and Reichlin, 2000; Stock and Watson, 2002a, 2002b; Bernanke, Boivin, and Eliasz, 2005; Boivin, Giannoni, and Mihov, 2009). Bayesian estimation (Otrok and Whiteman, 1998; Kim and Nelson, 1999; Kose, Otrok, and Whiteman, 2003; Del Negro and Otrok, 2007).

Appendices B and C contain proofs of theoretical results.[2]

2 Statistical Model and Estimation Methodology

Consider the statistical model

$$\pi_{nt} = \mu_n + A_n(L)\, u_t + B_n(L)\, v_{nt}, \tag{1}$$

where π_{nt} is the month-on-month inflation rate in sector n in period t, μ_n are constants, $A_n(L)$ and $B_n(L)$ are square summable polynomials in the lag operator, u_t is an unobservable factor following a unit-variance Gaussian white noise process, and each v_{nt} follows a unit-variance Gaussian white noise process. The processes v_{nt} are pairwise independent and independent of the process u_t.

It is straightforward to generalize equation (1) such that u_t follows a vector Gaussian white noise process with covariance matrix identity. In estimation, this paper considers the case when u_t follows a scalar process and the case when u_t follows a vector process.

Let π_{nt}^A denote the aggregate component of the inflation rate in sector n, that is,

$$\pi_{nt}^A = A_n(L)\, u_t.$$

The aggregate component of the inflation rate in sector n is parameterized as a finite-order moving average process. The order of the polynomials $A_n(L)$ is chosen to be as high as computationally feasible. Specifically, the order of the polynomials $A_n(L)$ is set to twenty four, that is, u_t and twenty four lags of u_t enter equation (1).

Let π_{nt}^S denote the sector-specific component of the inflation rate in sector n, that is,

$$\pi_{nt}^S = B_n(L)\, v_{nt}.$$

To reduce the number of parameters to estimate, the sector-specific component of the inflation rate in sector n is parameterized as an autoregressive process:

$$\pi_{nt}^S = C_n(L)\, \pi_{nt}^S + B_{n0} v_{nt},$$

[2] Data and replication code are available from the authors.

where $C_n(L)$ is a polynomial in the lag operator satisfying $C_{n0} = 0$. In estimation, this paper considers the case when the order of the polynomials $C_n(L)$ equals six and the case when the order of the polynomials $C_n(L)$ equals twelve.

Before estimation, the sectoral inflation rates are demeaned. Furthermore, the sectoral inflation rates are normalized to have unit variance. These adjustments imply that the estimated model is

$$\tilde{\pi}_{nt} = a_n(L)\, u_t + b_n(L)\, v_{nt},$$

where $\tilde{\pi}_{nt} = [(\pi_{nt} - \mu_n)/\sigma_{\pi_n}]$ is the normalized inflation rate in sector n in period t, and $a_n(L)$ and $b_n(L)$ are square summable polynomials in the lag operator. Here σ_{π_n} is the standard deviation of the inflation rate in sector n. The following relationships hold: $A_n(L) = \sigma_{\pi_n} a_n(L)$ and $B_n(L) = \sigma_{\pi_n} b_n(L)$. This normalization makes it easier to compare impulse responses across sectors. In what follows, the paper refers to coefficients appearing in the polynomials $a_n(L)$ and $b_n(L)$ as "normalized impulse responses".

This paper uses Bayesian methods to estimate the model. In particular, the Gibbs sampler with a Metropolis-Hastings step is used to sample from the joint posterior density of the factors and the model's parameters. Taking as given a Monte Carlo draw of the model's parameters, one samples from the conditional posterior density of the factors given the model's parameters. Here the paper follows Carter and Kohn (1994) and Kim and Nelson (1999). Afterwards, taking as given a Monte Carlo draw of the factors, one samples from the conditional posterior density of the model's parameters given the factors. Here the paper follows Chib and Greenberg (1994). The following prior is used. The prior has zero mean for each factor loading and for each autoregressive coefficient in the sector-specific component of the inflation rate in sector n. The prior starts out loose and becomes gradually tighter at more distant lags.[3]

The paper contributes to the branch of the literature on estimation of dynamic factor models using Bayesian methods.[4] The extant papers in this branch assume that factors follow independent autoregressive processes and that the loading of each variable on each factor is a scalar. Instead, here it is assumed that factors follow independent white noise

[3] See Appendix A for econometric details, including details of the prior.

[4] See Footnote 1.

processes and that the loadings of each variable on each factor form a polynomial in the lag operator. See equation (1). The former setup implies that, for any pair of variables i and j, the impulse response function of variable i to an innovation in a factor is proportional to the impulse response function of variable j to the same innovation. The latter setup implies no such restriction.[5] Note that the statistical model of Boivin, Giannoni, and Mihov (2009) has the same implication for impulse response functions as the former setup.[6] We believe it is important to allow for the possibility that the impulse response function of a sectoral price index to an aggregate shock differs in shape across sectors. Therefore, we prefer the latter setup.

The use of Bayesian methods offers a specific advantage in the context of analyses like Boivin, Giannoni, and Mihov's and ours. When one estimates regression relationships using variables derived from the dynamic factor model, Bayesian methods allow one to quantify easily the uncertainty concerning the regression relationships. See Sections 5-6. Without Bayesian methods, one typically proceeds as if the point estimate of, say, the standard deviation of sectoral inflation due to sector-specific shocks derived from the dynamic factor model were the truth.[7]

[5] An unpublished paper by Justiniano (2004) uses the latter setup and Bayesian methods, like this paper. This paper differs from Justiniano (2004) in that this paper includes a Metropolis-Hastings step in the Gibbs sampler while Justiniano does not. This difference means that, in sampling from the conditional posterior density of the model's parameters given the factors, this paper uses the full likelihood function while Justiniano uses the likelihood function conditional on initial observations.

[6] Boivin, Giannoni, and Mihov (2009) estimate their factor augmented VAR model using a two-step approach. First, they estimate multiple common factors using principal components methods. Second, they estimate a joint VAR of the estimated factors and the federal funds rate. This setup implies that the impulse responses of inflation rates in different sectors to an innovation in a given factor are proportional. Afterwards, they fit a univariate autoregressive process to the common component of each inflation rate. Here the common component is a weighted sum of the factors and the federal funds rate, where the weights are the factor loadings. Since the restrictions from the factor augmented VAR model are not taken into account in this estimation step and the factor loadings may differ across sectors, the impulse responses in different sectors to an innovation in the common component need not be proportional.

[7] Bayesian estimation of a dynamic factor model also offers a general advantage compared with estimation based on principal components. One obtains the joint posterior density of the factors and the model's parameters.

The advantage of principal-component-based estimation of a dynamic factor model, as in Boivin, Giannoni, and Mihov (2009), is that it is straightforward, from the computational point of view, to add more variables. For example, Boivin, Giannoni, and Mihov add sectoral data on quantities and macroeconomic data.

3 Data

This paper uses the data underlying the consumer price index (CPI) for all urban consumers in the United States. The data are compiled by the Bureau of Labor Statistics (BLS). The data are monthly sectoral price indexes. The sectoral price indexes are available at four different levels of aggregation: from least disaggregate (8 "major groups") to most disaggregate (205 sectors).[8] This paper focuses on the most disaggregate sectoral price indexes. For some sectors, price indexes are available for only a short period, often starting as recently as in 1998. This paper focuses on the 79 sectors for which monthly price indexes are available from January 1985. These sectors comprise 68.1 percent of the CPI. Each "major group" is represented. The sample used here ends in May 2005.

The median standard deviation of sectoral inflation in the cross-section of sectors in this paper's dataset is 0.0068. For comparison, the standard deviation of the CPI inflation rate in this paper's sample period is 0.0017. In 76 out of 79 sectors, the sectoral inflation rate is more volatile than the CPI inflation rate.

To gain an idea about the degree of comovement in this paper's dataset, one can compute principal components of the normalized sectoral inflation rates. The first few principal components explain only little of the variation in the normalized sectoral inflation rates. In particular, the first principle component explains 7 percent of the variation, and the first five principle components together explain 20 percent of the variation.

These observations suggest that changes in sectoral price indexes are caused mostly by sector-specific shocks.

[8] The "major groups" are (with the percentage share in the CPI given in brackets): food and beverages (15.4), housing (42.1), apparel (4.0), transportation (16.9), medical care (6.1), recreation (5.9), education and communication (5.9), other goods and services (3.8).

4 Responses of Sectoral Price Indexes to Sector-Specific Shocks and to Aggregate Shocks

This section reports results from the estimated dynamic factor model (1). The focus is on the benchmark specification in which the factor u_t follows a scalar process and the order of the polynomials $C_n(L)$ equals six. Two other specifications were also estimated: (i) a specification in which the order of the polynomials $C_n(L)$ equals twelve, and (ii) a specification in which u_t follows a bivariate vector process and the order of the polynomials $C_n(L)$ equals six. It turned out that the specification in which u_t follows a scalar process and the order of the polynomials $C_n(L)$ equals six forecasts better out-of-sample compared with the other two specifications. Therefore, this specification was chosen as the benchmark specification. Section 6 discusses the results from the other two specifications. Furthermore, the out-of-sample forecast performance of the dynamic factor model was compared with that of simple, autoregressive models for sectoral inflation. It turned out that: (i) the benchmark dynamic factor model forecasts better than the AR(6) model, and (ii) the dynamic factor model in which the order of the polynomials $C_n(L)$ equals twelve forecasts better than the AR(12) model. The forecast results show that the dynamic factor model fits the data well.[9]

To begin consider the variance decomposition of sectoral inflation into aggregate shocks and sector-specific shocks. Sector-specific shocks account for a dominant share of the variance in sectoral inflation. In the median sector, the share of the variance in sectoral inflation due to sector-specific shocks equals 90 percent. The sectoral distribution is tight. In the sector that lies at the 5th percentile of the sectoral distribution, the share of the variance in sectoral inflation due to sector-specific shocks equals 79 percent, and in the sector that lies at the 95th percentile of the sectoral distribution, the share of the variance in sectoral inflation due to sector-specific shocks equals 95 percent.

[9] The out-of-sample forecast exercise consisted of the following steps. (1) For each specification of the dynamic factor model and for each sector: (i) compute the forecast of the normalized sectoral inflation rate one-step-ahead in the last twenty four periods in the dataset, and (ii) save the average root mean squared error of the twenty four forecasts. (2) Performe the same exercise using an AR model for the normalized sectoral inflation rate. The AR model was estimated separately for each sector by OLS, with the number of lags equal to, alternatively, six and twelve.

Next, consider the impulse responses of sectoral price indexes to sector-specific shocks and to aggregate shocks. Figure 1 shows the cross-section of the normalized impulse responses of sectoral price indexes to sector-specific shocks (top panel) and to aggregate shocks (bottom panel). Each panel shows the posterior density taking into account both variation across sectors and parameter uncertainty. Specifically, for each sector, 7500 draws are made from the posterior density of the normalized impulse response of the sectoral price index to a given shock.[10] Afterwards, 1000 draws are selected at random. Since there are 79 sectors, this procedure gives a sample of 79000 impulse responses. Each panel in Figure 1 is based on 79000 impulse responses. The median impulse response of a sectoral price index to a sector-specific shock has the following shape. After a sector-specific shock, 100 percent of the long-run response of the sectoral price index occurs in the month of the shock, and the response equals the long-run response in all months following the shock. The median impulse response of a sectoral price index to an aggregate shock has a very different shape. After an aggregate shock, only 15 percent of the long-run response of the sectoral price index occurs in the month of the shock, and the response gradually approaches the long-run response in the months following the shock. Another way of summarizing the median impulse responses is as follows. The sector-specific component of the sectoral inflation rate is essentially a white noise process, while the aggregate component of the sectoral inflation rate is positively autocorrelated with an autocorrelation coefficient equal to 0.35.[11]

It is useful to compute a simple measure of the speed of the response of a price index to a given type of shock. Specifically, consider the absolute response to the shock in the short run divided by the absolute response to the shock in the long run. Take the short run to be between the impact of the shock and five months after the impact of the shock. Take the long run to be between 19 months and 24 months after the impact of the shock. Formally, let β_{nm} denote the impulse response of the price index for sector n to a sector-specific shock m periods after the shock. The speed of response of the price index for sector n to

[10] See Appendix A for details of the Gibbs sampler.

[11] Regressing the median impulse response of a sectoral inflation rate on its own lag yields a coefficient of 0.35.

sector-specific shocks is defined as:

$$\Lambda_n^S \equiv \frac{\frac{1}{6}\sum_{m=0}^{5} \mid \beta_{nm} \mid}{\frac{1}{6}\sum_{m=19}^{24} \mid \beta_{nm} \mid}.$$

Furthermore, let α_{nm} denote the impulse response of the price index for sector n to an aggregate shock m periods after the shock. The speed of response of the price index for sector n to aggregate shocks is defined as:

$$\Lambda_n^A \equiv \frac{\frac{1}{6}\sum_{m=0}^{5} \mid \alpha_{nm} \mid}{\frac{1}{6}\sum_{m=19}^{24} \mid \alpha_{nm} \mid}.$$

Figure 2 shows the cross-section of Λ_n^S (top panel) and the cross-section of Λ_n^A (bottom panel). Each panel shows the posterior density taking into account both variation across sectors and parameter uncertainty. Figure 2 has two main features. The median speed of response of a sectoral price index to sector-specific shocks is much larger than the median speed of response of a sectoral price index to aggregate shocks. The median speed of response of a sectoral price index to sector-specific shocks equals 1.01. The median speed of response of a sectoral price index to aggregate shocks equals 0.41.[12] Furthermore, the cross-section of the speed of response to sector-specific shocks is tight, while the cross-section of the speed of response to aggregate shocks is dispersed. 68 percent of the posterior probability mass of Λ_n^S lies between 0.89 and 1.05. 68 percent of the posterior probability mass of Λ_n^A lies between 0.2 and 1.12. There is little cross-sectional variation in the speed of response to sector-specific shocks, while there is considerable cross-sectional variation in the speed of response to aggregate shocks.[13]

[12] One can also look at the speed of response to shocks sector by sector. In 76 out of 79 sectors, the median speed of response of the sectoral price index to sector-specific shocks is larger than the median speed of response of the sectoral price index to aggregate shocks. Furthermore, one can construct, in each sector, a posterior probability interval for the speed of response to sector-specific shocks and a posterior probability interval for the speed of response to aggregate shocks. When 68 percent posterior probability intervals are constructed, in 43 out of 79 sectors the posterior probability interval for the speed of response to sector-specific shocks lies strictly above the posterior probability interval for the speed of response to aggregate shocks.

[13] Alternative measures of the speed of response to shocks yielded the same conclusions.

5 Regression Analysis

The last section showed that there is little cross-sectional variation in the speed of response to sector-specific shocks and considerable cross-sectional variation in the speed of response to aggregate shocks. This section studies whether the cross-sectional variation in the speed of response to a given type of shock is related to sectoral characteristics that we have information on. All regressions reported below are motivated by models of price setting that are presented in more detail in Section 7.

5.1 The Speed of Response and the Frequency of Price Changes

A basic prediction of the Calvo model is that sectoral price indexes respond faster to shocks in sectors with a higher frequency of price changes (holding constant all other sectoral characteristics).

Bils and Klenow (2004) report the monthly frequency of price changes for 350 categories of consumer goods and services, based on data from the BLS for the period 1995-1997. We can match 75 out of our 79 sectors into the categories studied by Bils and Klenow (2004). Nakamura and Steinsson (2008) report the monthly frequency of price changes for 270 categories of consumer goods and services, based on data from the BLS for the period 1998-2005. We can match 77 out of our 79 sectors into the categories studied by Nakamura and Steinsson (2008).

The information on the speed of response of the price index for sector n to a given type of shock comes from the estimated dynamic factor model. Note that we do not know the speed of response for certain. Instead, we have a posterior density of the speed of response. To account for uncertainty about the regression relationship in the regressions below, the posterior density of the regression coefficient is reported.[14]

Consider two regressions. First, consider the regression of the speed of response of the price index for sector n to aggregate shocks (Λ_n^A) on the sectoral monthly frequency of price

[14] Many draws are made from the posterior density of the speed of response. For each draw, the posterior density of the regression coefficient conditional on this draw is constructed and a draw is made from this density. This procedure yields the marginal posterior density of the regression coefficient, with the speed of response integrated out. This marginal posterior density is reported.

changes from Bils and Klenow (2004) and, alternatively, on the sectoral monthly frequency of regular price changes from Nakamura and Steinsson (2008).[15] The top two rows in Table 1 show that: (i) the posterior median of the regression coefficient is positive, (ii) the 90 percent posterior probability interval for the regression coefficient excludes zero, and (iii) the regression results using the Bils-Klenow frequencies differ little from the regression results using the Nakamura-Steinsson frequencies.

Second, consider the regression of the speed of response of the price index for sector n to sector-specific shocks (Λ_n^S) on the sectoral monthly frequency of price changes from Bils and Klenow (2004) and, alternatively, on the sectoral monthly frequency of regular price changes from Nakamura and Steinsson (2008). These results are in the bottom two rows in Table 1. With the Bils-Klenow frequencies, the regression coefficient is positive, but the regression coefficient is significantly smaller than the coefficient in the first regression. Furthermore, with the Nakamura-Steinsson frequencies of regular price changes, there is moderately strong support for a negative relationship.

5.2 The Speed of Response and the Standard Deviation of Shocks

In the rational-inattention model of Maćkowiak and Wiederholt (2009a), agents pay more attention to those shocks that on average cause more variation in the optimal decision. Therefore, the model predicts that sectoral price indexes respond faster to aggregate shocks in sectors with a larger standard deviation of sectoral inflation due to aggregate shocks. Similarly, the model predicts that sectoral price indexes respond faster to sector-specific shocks in sectors with a larger standard deviation of sectoral inflation due to sector-specific shocks.

Consider two regressions. First, consider the regression of the speed of response of the price index for sector n to aggregate shocks (Λ_n^A) on the standard deviation of sectoral inflation due to aggregate shocks. The results are in the top row in Table 2. The posterior median of the regression coefficient is positive. The 90 percent posterior probability interval excludes zero. Second, consider the regression of the speed of response of the price index for

[15] Regular price changes in Nakamura and Steinsson (2008) exclude price changes related to sales and product substitutions.

sector n to sector-specific shocks (Λ_n^S) on the standard deviation of sectoral inflation due to sector-specific shocks. The results are in the bottom row in Table 2. The posterior median of the regression coefficient is again positive. This time the 90 percent posterior probability interval includes zero but only barely. As one can see from the table, 94 percent of the posterior probability mass lies to the right of zero. One can conclude that the posterior evidence provides strong support for these predictions of the model. In addition, note that the 90 percent posterior probability intervals for the two coefficients barely overlap, suggesting that there is a difference in the magnitude of the two coefficients. Section 7 shows that the rational-inattention model of Maćkowiak and Wiederholt (2009a) predicts a difference in the magnitude of the two coefficients.[16]

Section 7 also shows another prediction of the rational-inattention model of Maćkowiak and Wiederholt (2009a). When the amount of information processed by price setters in firms is given exogenously or when price setters in firms can decide to process more information subject to a strictly convex cost function, there is a tension between attending to aggregate conditions and attending to sector-specific conditions. Under these circumstances, the model predicts that the speed of response of a sectoral price index to aggregate shocks is: (i) increasing in the standard deviation of sectoral inflation due to aggregate shocks, and (ii) decreasing in the standard deviation of sectoral inflation due to sector-specific shocks. The results for the corresponding regression are in the middle row in Table 2. There is moderately strong support for this prediction of the model: 92 percent of the posterior probability mass for the coefficient on the standard deviation of sectoral inflation due to aggregate shocks lies to the right of zero, and 80 percent of the posterior probability mass for the coefficient on the standard deviation of sectoral inflation due to sector-specific shocks lies to the left of zero.

[16] In the regressions reported in Table 2 both the regressand and the regressor (the regressors) are uncertain. Many draws are made from the joint posterior density of the regressand and the regressor (the regressors). For each joint draw, the posterior density of the regression coefficient conditional on this joint draw is constructed and a draw is made from this density. This procedure yields the marginal posterior density of the regression coefficient, with the regressand and the regressor (the regressors) integrated out. This marginal posterior density is reported.

5.3 The Frequency of Price Changes and the Standard Deviation of Shocks

A basic prediction of the menu cost model is that firms change prices more frequently in sectors with larger shocks (holding constant all other sectoral characteristics).

In the data, sector-specific shocks account for a dominant share of the variance in sectoral price indexes. Therefore, a simple way to investigate this prediction of the menu cost model is to look for a positive relationship between the sectoral monthly frequency of price changes and the standard deviation of sectoral inflation due to sector-specific shocks. Table 3 shows strong evidence for the positive relationship, in the case of the Bils-Klenow frequencies and in the case of the Nakamura-Steinsson frequencies.

The menu cost model also predicts a positive relationship between the frequency of price changes and the steady-state inflation rate. This prediction was investigated, but no relationship was found between the monthly frequency of price changes in a given sector and the mean inflation rate in that sector. It is plausible that more variation in mean inflation rates than is present in this paper's sample would be needed for a significant positive relationship to arise.

6 Robustness

This section considers three robustness checks.

6.1 The Distribution of Sector-Specific Shocks

This subsection examines the posterior density of sector-specific shocks, $(v_{n1}, ..., v_{nT})_{n=1}^{N}$, from the benchmark specification of the dynamic factor model. Specifically, the posterior density of skewness and the posterior density of kurtosis of the sector-specific shocks are examined. Each density suggests that the sector-specific shocks are slightly non-Gaussian. The posterior density of skewness has a median of zero but it has a sizable negative tail (the posterior mean is -0.1). The posterior density of kurtosis has a median of 3.7 and a mean of 4.2. The extent of non-Gaussianity fails to change when one allows for more lags in the sector-specific component of the sectoral inflation rate and when one adds another

factor. However, the negative skewness and the excess kurtosis come mainly from only a few sectors. These sectors are dropped from the sample and the benchmark specification of the dynamic factor model is reestimated.[17] The findings reported in Sections 4 and 5 remain unaffected.[18] Furthermore, the sector-specific shocks from the reestimated benchmark specification appear approximately Gaussian. The posterior density of skewness has a median of zero and a mean of -0.01. The posterior density of kurtosis has a median of 3.6 and a mean of 3.7. These results suggest that the findings reported in Sections 4 and 5 are not driven by a few sectors experiencing non-Gaussian sector-specific shocks.

6.2 More Lags and Quarterly Data

This subsection examines the possibility that the findings reported in Sections 4 and 5 are influenced by the fact that the sector-specific component of the sectoral inflation rate is approximated as an autoregressive process. If the sector-specific component of the sectoral inflation rate has a moving average root large in absolute value, one needs to allow for many lags in the autoregressive approximation for it to be accurate. First, a specification of the dynamic factor model is estimated that allows for more lags in the sector-specific component of the sectoral inflation rate. In particular, a specification is estimated in which the order of the polynomials $C_n(L)$ equals twelve. The findings reported in Sections 4 and 5 remain unaffected. Second, the benchmark specification of the dynamic factor model is reestimated using quarterly data.[19] Not surprisingly, in the median sector the share of the variance in sectoral inflation due to sector-specific shocks falls, to 71 percent from 90 percent with monthly data. The speed of response to aggregate shocks remains unaffected. The speed of response to sector-specific shocks falls somewhat, but it remains much higher than the speed of response to aggregate shocks. The support for the regression relationships predicted by the rational-inattention model of Maćkowiak and Wiederholt (2009a) actually

[17] Specifically, 11 sectors are dropped. The sample is reduced to 68 sectors.

[18] For example, the median speed of response to aggregate shocks equals 0.41, exactly as reported in Section 4. The median speed of response to sector-specific shocks equals 1.02, 0.01 higher than reported in Section 4. 68 percent of the posterior probability mass of Λ_n^S lies between 0.89 and 1.05, exactly as reported in Section 4. 68 percent of the posterior probability mass of Λ_n^A lies between 0.2 and 1.05, almost exactly as reported in Section 4.

[19] The order of the polynomials $A_n(L)$ equals eight and the order of the polynomials $C_n(L)$ equals two.

strengthens. See Table 4 which reproduces, based on quarterly data, the rational-inattention model regressions from Table 2.

6.3 Multiple Factors

The final robustness check is to estimate a specification of the dynamic factor model with two factors. In particular, a specification is estimated in which u_t follows a bivariate vector process and the order of the polynomials $C_n(L)$ equals six. The conclusion that sector-specific shocks account for a dominant share of the variance in sectoral price indexes remains unaffected. In the median sector, the share of the variance in sectoral inflation due to sector-specific shocks falls only a little, to 89 percent from 90 percent in the benchmark specification. The conclusion that sectoral price indexes respond quickly to sector-specific shocks and slowly to aggregate shocks also remains unaffected, although the speed of response to aggregate shocks increases somewhat. In the median sector, 15 percent of the long-run response of the sectoral price index occurs within one month of an innovation in one factor; and 30 percent of the long-run response of the sectoral price index occurs within one month of an innovation in the other factor. Most regression relationships reported in Section 5 become somewhat weaker. This is as expected given that many parameters are estimated in the specification with two factors. Note also that the specification with two factors performs worse in the out-of-sample forecast exercise than the benchmark specification. This difference makes us focus on the results from the benchmark specification.

7 Models of Price Setting

This section studies whether different models of price setting can match the empirical findings reported in Sections 4-6. Four models of price setting are considered: the Calvo model, a menu-cost model, the sticky-information model developed in Mankiw and Reis (2002), and the rational-inattention model developed in Maćkowiak and Wiederholt (2009a). Since several of the empirical findings reported in Sections 4-6 are about the response of sectoral price indexes to sector-specific shocks, versions of these four models with multiple sectors and sector-specific shocks are studied. To fix ideas, Section 7.1 presents a specific multi-

sector setup. Later it is shown that the main theoretical results do not depend on the details of the multi-sector setup.

7.1 Common Setup

Consider an economy with a continuum of sectors of mass one. In each sector, there is a continuum of firms of mass one. Sectors are indexed by n and firms within a sector are indexed by i. Each firm supplies a differentiated good and sets the price for the good.

The demand for good i in sector n in period t is given by[20]

$$C_{int} = \left(\frac{P_{int}}{P_{nt}}\right)^{-\theta} \left(\frac{P_{nt}}{P_t}\right)^{-\eta} C_t, \tag{2}$$

where P_{int} is the price of good i in sector n, P_{nt} is the sectoral price index, P_t is the aggregate price index and C_t is aggregate composite consumption. The parameters satisfy $\theta > 1$ and $\eta > 1$. The sectoral price index and the aggregate price index are given by

$$P_{nt} = \left(\int_0^1 P_{int}^{1-\theta} di\right)^{\frac{1}{1-\theta}}, \tag{3}$$

and

$$P_t = \left(\int_0^1 P_{nt}^{1-\eta} dn\right)^{\frac{1}{1-\eta}}. \tag{4}$$

Output of firm i in sector n in period t is given by

$$Y_{int} = Z_{nt} L_{int}^{\alpha}, \tag{5}$$

where Z_{nt} is sector-specific total factor productivity (TFP) and L_{int} is labor input of the firm. The parameter $\alpha \in (0, 1]$ is the elasticity of output with respect to labor input. In every period, firms produce the output that is required to satisfy demand

$$Y_{int} = C_{int}. \tag{6}$$

Finally, the real wage rate in period t is assumed to equal $w(C_t)$, where $w : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly increasing, twice continuously differentiable function.

[20] The demand function (2) with price indexes (3) and (4) can be derived from expenditure minimization by households when households have a CES consumption aggregator, where $\theta > 1$ is the elasticity of substitution between goods from the same sector and $\eta > 1$ is the elasticity of substitution between consumption aggregates from different sectors.

Substituting the demand function (2), the production function (5), equation (6) and the real wage rate $w(C_t)$ into the usual expression for nominal profits and dividing by the price level yields the real profit function. A log-quadratic approximation of the real profit function around the non-stochastic solution of the model yields the following expression for the profit-maximizing price in period t

$$p_{int}^{\diamond} = \underbrace{p_t + \frac{\omega + \frac{1-\alpha}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\theta} c_t}_{p_{int}^{\diamond A}} + \underbrace{\frac{\frac{1-\alpha}{\alpha}(\theta - \eta)}{1 + \frac{1-\alpha}{\alpha}\theta} \hat{p}_{nt} - \frac{\frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\theta} z_{nt}}_{p_{int}^{\diamond S}}, \tag{7}$$

where $p_{int} = \ln(P_{int})$, $p_t = \ln(P_t)$, $c_t = \ln\left(C_t/\bar{C}\right)$, $\hat{p}_{nt} = \ln(P_{nt}/P_t)$, and $z_{nt} = \ln\left(Z_{nt}/\bar{Z}\right)$. Here $\bar{C}$, $\bar{Z}$ and ω denote composite consumption, TFP and the elasticity of the real wage with respect to composite consumption at the non-stochastic solution. Note that the profit-maximizing price has an aggregate component, $p_{int}^{\diamond A}$, and a sector-specific component, $p_{int}^{\diamond S}$.[21] Furthermore, after the log-quadratic approximation of the real profit function, the profit loss in period t due to a deviation from the profit-maximizing price equals

$$\frac{\bar{C}(\theta - 1)\left(1 + \frac{1-\alpha}{\alpha}\theta\right)}{2}\left(p_{int} - p_{int}^{\diamond}\right)^2. \tag{8}$$

See Appendix A in Maćkowiak and Wiederholt (2009a) for the derivation of equation (8).

In addition to the log-quadratic approximation of the real profit function, log-linearization of the equations for the price indexes around the non-stochastic solution of the model yields

$$p_{nt} = \int_0^1 p_{int} di, \tag{9}$$

and

$$p_t = \int_0^1 p_{nt} dn, \tag{10}$$

where $p_{nt} = \ln(P_{nt})$.

In the following price-setting models, it is assumed that the profit-maximizing price equals (7), the profit loss due to a deviation from the profit-maximizing price equals (8), and the sectoral price index and the aggregate price index are given by equations (9) and (10), respectively.

[21] Introducing sector-specific shocks in the form of multiplicative demand shocks in (2) instead of multiplicative productivity shocks in (5) yields an equation for the profit-maximizing price that is almost identical to equation (7). The only difference is the coefficient in front of z_{nt}.

7.2 Calvo Model

In the Calvo model, a firm can adjust its price with a constant probability in any given period. Let λ_n denote the probability that a firm in sector n can adjust its price. Assume that the profit-maximizing price of good i in sector n in period t is given by equation (7), the price index for sector n in period t is given by equation (9), and a firm in sector n that can adjust its price in period t sets the price that minimizes

$$E_t \left[\sum_{s=t}^{\infty} \left[(1-\lambda_n)\,\beta \right]^{s-t} \frac{\bar{C}\,(\theta-1)\left(1+\frac{1-\alpha}{\alpha}\theta\right)}{2} \left(p_{int} - p_{ins}^{\diamondsuit} \right)^2 \right]. \tag{11}$$

In this model, the profit-maximizing price equals the sum of two components: an aggregate component and a sector-specific component. Furthermore, the aggregate component, $p_{int}^{\diamondsuit A}$, and the sector-specific component, $p_{int}^{\diamondsuit S}$, are the same for all firms within a sector. Formally, the profit-maximizing price of firm i in sector n in period t has the form

$$p_{int}^{\diamondsuit} = p_{nt}^{\diamondsuit A} + p_{nt}^{\diamondsuit S}. \tag{12}$$

A firm in sector n that can adjust its price in period t sets the price

$$p_{int}^{*} = \left[1 - (1-\lambda_n)\,\beta \right] E_t \left[\sum_{s=t}^{\infty} \left[(1-\lambda_n)\,\beta \right]^{s-t} p_{ins}^{\diamondsuit} \right]. \tag{13}$$

The price set by adjusting firms equals a weighted average of the current profit-maximizing price and future profit-maximizing prices. Finally, the price index for sector n in period t equals

$$p_{nt} = (1-\lambda_n)\,p_{nt-1} + \lambda_n p_{int}^{*}, \tag{14}$$

because the adjusting firms are drawn randomly and all adjusting firms in a sector set the same price.

Recall that the median impulse response of sectoral price indexes to sector-specific shocks reported in Figure 1 has the property that all of the response of the sectoral price index to a sector-specific shock occurs in the month of the shock. The following proposition answers the question of whether the standard Calvo model can match the median impulse response of sectoral price indexes to sector-specific shocks.

Proposition 1 *(Calvo model with sector-specific shocks) Suppose that the profit-maximizing price of firm i in sector n in period t is given by equation (12), the price set by adjusting*

firms is given by equation (13), and the sectoral price index is given by equation (14). Then, the impulse response of the price index for sector n to a shock equals x on impact of the shock and in all periods following the shock if and only if the impulse response of the profit-maximizing price to the shock equals: (i)

$$\frac{\frac{1}{\lambda_n} - (1-\lambda_n)\beta}{1-(1-\lambda_n)\beta} x, \tag{15}$$

on impact of the shock, and (ii) x thereafter.

Proof. See Appendix B. ■

In the Calvo model, there exists a unique impulse response of the profit-maximizing price to a sector-specific shock which implies that all of the response of the sectoral price index to the sector-specific shock occurs in the period of the shock. If prices are flexible ($\lambda_n = 1$), the sector-specific component of the profit-maximizing price has to follow a random walk. If prices are sticky ($0 < \lambda_n < 1$), the profit-maximizing price first needs to jump by expression (15) on impact of the shock and then has to jump back to x in the period following the shock to generate a response equal to x of the sectoral price index on impact of the shock and in all periods following the shock. Proposition 1 follows directly from equations (12)-(14). Note that the required extent of overshooting of the profit-maximizing price depends only on the two parameters λ_n and β.

To illustrate Proposition 1, consider the following three examples. In each example, one period equals one month. Therefore, set $\beta = 0.99^{1/3}$. First, suppose that $\lambda_n = (1/12)$. This value implies that firms adjust their prices on average once a year. Then the profit-maximizing response on impact has to overshoot the profit-maximizing response in the next month by a *factor* of 128. Second, suppose that $\lambda_n = 0.087$. This is the monthly frequency of regular price changes (i.e., excluding sales and item substitutions) reported by Nakamura and Steinsson (2008). Then the profit-maximizing response on impact has to overshoot the profit-maximizing response in the next month by a factor of 118. Third, suppose that $\lambda_n = 0.21$. This is the monthly frequency of price changes reported by Bils and Klenow (2004). Then the profit-maximizing response on impact has to overshoot the profit-maximizing response in the next month by a factor of 19. All three examples are depicted in Figure 3. For the sake of clarity, the impulse response of the sectoral price

index in Figure 3 is normalized to one.

Going a step further, consider the impulse response of sector-specific productivity that yields the impulse response of the profit-maximizing price described in Proposition 1. When the profit-maximizing price is given by equation (7), the sector-specific component of the profit-maximizing price equals

$$p_{nt}^{\diamondsuit S} = \frac{\frac{1-\alpha}{\alpha}(\theta - \eta)}{1 + \frac{1-\alpha}{\alpha}\theta}\hat{p}_{nt} - \frac{\frac{1}{\alpha}}{1 + \frac{1-\alpha}{\alpha}\theta}z_{nt}. \tag{16}$$

Solving the last equation for sector-specific productivity yields

$$z_{nt} = -\frac{1 + \frac{1-\alpha}{\alpha}\theta}{\frac{1}{\alpha}}\left[p_{nt}^{\diamondsuit S} - \frac{\frac{1-\alpha}{\alpha}(\theta - \eta)}{1 + \frac{1-\alpha}{\alpha}\theta}\hat{p}_{nt}\right]. \tag{17}$$

Substituting the impulse response of the profit-maximizing price described in Proposition 1 and the impulse response of the sectoral price index into equation (17) delivers the impulse response of sector-specific productivity that yields the impulse response of the profit-maximizing price described in Proposition 1. For the parameter values $\alpha = (2/3)$, $\theta = 4$ and $\eta = 2$, Figure 4 shows the impulse responses of sector-specific productivity that yield the impulse responses of the profit-maximizing price depicted in Figure 3.

We interpret the results presented in Figure 1, Proposition 1 and Figure 3 as saying that the standard Calvo model has difficulties matching the median empirical response of sectoral price indexes to sector-specific shocks. To match the median empirical response of sectoral price indexes to sector-specific shocks, one needs a value for the Calvo parameter that is close to one in a monthly model or one has to make an extreme assumption concerning the response of the profit-maximizing price to sector-specific shocks. One could try to modify the Calvo model. Since Proposition 1 follows directly from equations (12)-(14), one has to modify at least one of these three equations to change this property of the model. Consider two potential modifications. First, one could assume that the profit-maximizing price differs across firms within a sector. However, the only change in Proposition 1 is that the proposition becomes a statement about the response of the average profit-maximizing price in the sector. For some firms the profit-maximizing price can respond less but then for other firms the profit-maximizing price has to respond more. Second, one could assume that with probability λ_n^A a firm can adjust *only* the aggregate component of its price and

with probability λ_n^S a firm can adjust *only* the sector-specific component of its price. If one assumes in addition that the parameter λ_n^A is small and the parameter λ_n^S is large, the model can generate a slow response of the sectoral price index to aggregate shocks and a quick response of the sectoral price index to sector-specific shocks. However, it seems difficult to justify these assumptions in the context of the Calvo model.

Next, consider whether the standard Calvo model can match the cross-sectional distribution of the speed of response to sector-specific shocks. First, consider the case: $\theta = \eta$ and z_{nt} following a random walk. In this case, the sector-specific component of the profit-maximizing price (7) is independent of the prices set by other firms and follows a random walk. The impulse response of the price index for sector n to a sector-specific shock then has the property that the fraction of the long-run response that has occurred, say, three periods after the shock simply equals the fraction of firms that have adjusted their prices in the last four periods:

$$\sum_{j=0}^{3} \lambda_n (1-\lambda_n)^j = 1 - (1-\lambda_n)^4 . \tag{18}$$

For $\lambda_n = 0.1$, $\lambda_n = 0.25$ and $\lambda_n = 0.5$, expression (18) equals 0.34, 0.68 and 0.94, respectively. Furthermore, according to Table A1 in Bils and Klenow (2004), these three values for λ_n correspond roughly to the 1st decile, the median and the 9th decile of the cross-sectional distribution of the monthly frequency of price changes in our sample of sectors. Hence, expression (18) and the cross-sectional distribution of the frequency of price changes imply substantial cross-sectional variation in the speed of response to sector-specific shocks. By contrast, the empirical part of this paper finds little cross-sectional variation in the speed of response to sector-specific shocks. See Figure 2. Expression (18) is derived assuming that $\theta = \eta$. When $\theta > \eta$, there is strategic complementarity in pricing in response to sector-specific shocks, which amplifies cross-sectoral differences in λ_n. By contrast, when $\theta < \eta$, there is strategic substitutability in pricing in response to sector-specific shocks, which mutes cross-sectoral differences in λ_n. Hence, to reduce the cross-sectional variation in the speed of response to sector-specific shocks in the standard Calvo model, one could assume $\theta < \eta$. However, this assumption seems implausible because $\theta < \eta$ means that the elasticity of substitution within sectors is smaller than the elasticity of substitution across sectors. Expression (18) is also based on the assumption that z_{nt} follows a random walk.

Thus, the other possibility of reducing the cross-sectional variation in the speed of response to sector-specific shocks in the standard Calvo model is to assume a different z_{nt} process. If sector-specific productivity "overshoots" on impact of a sector-specific shock and the extent of "overshooting" is larger in sectors with a smaller frequency of price changes, then there is less cross-sectional variation in the speed of response to sector-specific shocks. In fact, if in all sectors the impulse response of the profit-maximizing price to a sector-specific shock equals the one described in Proposition 1, then all sectoral price indexes respond fully on impact to sector-specific shocks and there is no cross-sectional variation in the speed of response to sector-specific shocks. However, this requires a very specific variation of the extent of "overshooting" with the frequency of price changes. For example, according to equation (15), the extent of "overshooting" in a sector with $\lambda_n = 0.1$ has to equal 30 times the extent of "overshooting" in a sector with $\lambda_n = 0.5$.

7.3 Sticky-Information Model

In the sticky-information model developed in Mankiw and Reis (2002), a firm can update its pricing plan with a constant probability in any given period. A pricing plan specifies a price path (i.e. a price as a function of time). The difference with the Calvo model is that firms choose a price path instead of a price. To understand the implications of this model for the impulse responses of sectoral price indexes to aggregate shocks and to sector-specific shocks consider a multi-sector version of the model with sector-specific shocks.

Let λ_n denote the probability that a firm in sector n can update its pricing plan. Assume that the profit-maximizing price of good i in sector n in period t is given by equation (7), the price index for sector n in period t is given by equation (9), and a firm in sector n that can update its pricing plan in period t chooses the price path that minimizes

$$E_t \left[\sum_{s=t}^{\infty} \beta^{s-t} \frac{\bar{C}\left(\theta - 1\right)\left(1 + \frac{1-\alpha}{\alpha}\theta\right)}{2} \left(p_{ins} - p_{ins}^{\diamond}\right)^2 \right]. \tag{19}$$

In this model, the profit-maximizing price of firm i in sector n in period t has the form

$$p_{int}^{\diamond} = p_{nt}^{\diamond A} + p_{nt}^{\diamond S}, \tag{20}$$

a firm that can update its pricing plan in period t chooses a price for period $s \geq t$ that

equals the conditional expectation of the profit-maximizing price in period s

$$p_{ins|t} = E_t \left[p_{ins}^{\diamondsuit} \right], \tag{21}$$

and the price index for sector n in period t equals

$$p_{nt} = \sum_{j=0}^{\infty} \lambda_n (1 - \lambda_n)^j E_{t-j} \left[p_{int}^{\diamondsuit} \right], \tag{22}$$

because a fraction $\lambda_n (1 - \lambda_n)^j$ of firms in sector n last updated their pricing plans j periods ago and these firms set a price equal to $E_{t-j} \left[p_{int}^{\diamondsuit} \right]$.

Comparing equations (13)-(14) and equations (21)-(22) shows two differences between the Calvo model and the sticky-information model. First, in the Calvo model firms front-load expected future changes in the profit-maximizing price, while in the sticky-information model firms wait with the price adjustment until the expected change in the profit-maximizing price actually occurs. Second, in the Calvo model inflation (i.e. a change in the price level) only comes from the fraction λ_n of firms that can adjust their prices in the current period, while in the sticky-information model inflation may also come from the fraction $(1 - \lambda_n)$ of firms that cannot update their pricing plans in the current period. Mankiw and Reis (2002) show that these two differences have interesting implications for the response of inflation and output to nominal shocks and to (anticipated and unanticipated) disinflations.[22]

To understand the implications of the standard sticky-information model for the impulse responses of sectoral price indexes to aggregate shocks and to sector-specific shocks, note the following property of impulse response functions in the standard sticky-information model. Firms that have updated their pricing plans since a shock occurred respond perfectly to the shock. All other firms do not respond at all to the shock. Furthermore, the fraction of firms that have updated their pricing plans over the last τ periods in sector n equals

$$\sum_{j=0}^{\tau} \lambda_n (1 - \lambda_n)^j = 1 - (1 - \lambda_n)^{\tau+1}. \tag{23}$$

Thus, the response of the price index for sector n in period t to a shock that occurred τ

[22] These statements refer to the Calvo model without indexation. The Calvo model with indexation to past inflation is more similar to the sticky-information model because in the Calvo model with indexation by setting a price the firm effectively chooses a price path.

periods ago simply equals $\left[1-(1-\lambda_n)^{\tau+1}\right]$ times the response of the profit-maximizing price in sector n in period t to the same shock. This is true for any shock.

To illustrate this result, consider the following example. Suppose that in sector n the aggregate component of the profit-maximizing price follows a random walk with a standard deviation of the innovation equal to σ_A and the sector-specific component of the profit-maximizing price follows a random walk with a standard deviation of the innovation equal to σ_S. Furthermore, suppose that firms in sector n update their pricing plans on average once a year, as assumed in Mankiw and Reis (2002). In a monthly model, this means $\lambda_n = (1/12)$. Figure 5 shows the impulse responses of the sectoral price index to aggregate shocks and to sector-specific shocks implied by the model for $\sigma_A = 0.48$ and $\sigma_S = 0.88$.[23] The impulse responses of the sectoral price index to the two shocks have an identical shape, independent of the standard deviation of the two shocks. The reason is that the impulse responses of the profit-maximizing price to the two shocks have an identical shape. For comparison, Figure 5 also reproduces from Figure 1 the median empirical response of sectoral price indexes to aggregate shocks as well as the median empirical response of sectoral price indexes to sector-specific shocks.

The following proposition answers the question of whether the standard sticky-information model can match the median impulse response of sectoral price indexes to sector-specific shocks reported in Figure 1.

Proposition 2 *(Sticky-information model with sector-specific shocks) Suppose that the profit-maximizing price of firm i in sector n in period t is given by equation (20) and the sectoral price index is given by equation (22). Then, the impulse response of the price index for sector n to a shock equals x on impact of the shock and in all periods following the shock if and only if, for all $\tau = 0, 1, 2, \ldots$, the impulse response of the profit-maximizing price τ periods after the shock equals*

$$\frac{1}{1-(1-\lambda_n)^{\tau+1}}x. \tag{24}$$

[23] The median impulse response of a sectoral price index to aggregate shocks reported in Figure 1 equals 0.48 in the long run. The median impulse response of a sectoral price index to sector-specific shocks reported in Figure 1 equals 0.88 in the long run.

Proof. This result follows directly from the sentence below equation (23). ■

In the standard sticky-information model, there exists a unique impulse response of the profit-maximizing price to a sector-specific shock which implies that all of the response of the sectoral price index to a sector-specific shock occurs in the period of the shock. If firms update their pricing plans every period ($\lambda_n = 1$), the sector-specific component of the profit-maximizing price has to follow a random walk. If firms update their pricing plans infrequently ($0 < \lambda_n < 1$), the profit-maximizing price first needs to jump by $(1/\lambda_n)\,x$ in the period of the shock and then has to decay slowly to x in the periods following the shock to generate a response equal to x of the sectoral price index on impact of the shock and in all periods following the shock. Proposition 2 follows directly from equations (20) and (22). Note that the required extent of overshooting of the profit-maximizing price depends only on the parameter λ_n.

To illustrate Proposition 2, consider the following example. Suppose that firms update their pricing plans on average once a year, as assumed in Mankiw and Reis (2002). In a monthly model, this means $\lambda_n = (1/12)$. Then the profit-maximizing response on impact of a sector-specific shock has to overshoot the profit-maximizing response in the long run by a *factor* of twelve. See Figure 6. Again one can compute from equation (17) the impulse response of sector-specific productivity that yields this impulse response of the profit-maximizing price. Note that less overshooting is necessary in the sticky-information model than in the Calvo model for the same value of λ_n, but the extent of overshooting is still large.

We interpret the results presented in Figure 5, Proposition 2 and Figure 6 as saying that the standard sticky-information model has difficulties matching the median empirical response of sectoral price indexes to sector-specific shocks. To match the median empirical response of sectoral price indexes to sector-specific shocks, one needs a value for λ_n close to one in a monthly model or one has to make an extreme assumption about the response of the profit-maximizing price to sector-specific shocks. One could modify the sticky-information model. Since Proposition 2 follows directly from equations (20) and (22), one has to modify at least one of these two equations to change this property of the model. Consider two potential modifications. First, one could assume that the profit-maximizing price differs

across firms within a sector. The only change in Proposition 2 is that the proposition becomes a statement about the response of the average profit-maximizing price in the sector. Second, one could assume that with probability λ_n^A a firm updates only its information concerning aggregate conditions and with probability λ_n^S a firm updates only its information concerning sector-specific conditions. If one assumes in addition that the parameter λ_n^A is small and the parameter λ_n^S is large, the model can generate a slow response of the sectoral price index to aggregate shocks and a quick response of the sectoral price index to sector-specific shocks. These assumptions seem plausible in the context of the sticky-information model. In particular, Reis (2006) shows that a model with a fixed cost of obtaining perfect information can provide a microfoundation for the sticky-information model of Mankiw and Reis (2002). One could envision a modification of the Reis (2006) model with the property that there is one fixed cost of obtaining perfect information concerning aggregate conditions and another fixed cost of obtaining perfect information concerning sector-specific conditions. This would make the model more similar to the model presented in the next paragraph.

7.4 Rational-Inattention Model of Maćkowiak and Wiederholt (2009a)

In the rational-inattention model developed in Maćkowiak and Wiederholt (2009a), price setters in firms have limited attention and decide what to focus on. Price setters face a trade-off between paying attention to aggregate conditions and paying attention to idiosyncratic conditions. Following Sims (2003), limited attention is modeled as a constraint on information flow. To understand the implications of this model for the impulse responses of sectoral price indexes to aggregate shocks and to sector-specific shocks, consider a simple multi-sector version of this model with sector-specific shocks.

The profit-maximizing price of good i in sector n in period t is given by equation (7). In period zero, the decision-maker in a firm chooses the precision of the signals that he or she will receive in the following periods. In each period $t \geq 1$, the decision-maker receives the signals and sets a price equal to the conditional expectation of the profit-maximizing price. In each period $t \geq 1$, the expectation is formed given the sequence of all signals that the decision-maker has received up to that point in time. The sectoral price index for sector n in period t is given by equation (9).

To make the results for this model as transparent as possible, this subsection presents an analytical solution for the price index for sector n in the case when the aggregate component and the sector-specific component of the profit-maximizing price each follow a Gaussian random walk. It is straightforward to compute numerical solutions of the model when the profit-maximizing price follows some other Gaussian process.[24]

Formally, in period zero, the decision-maker in a firm chooses the precision of the signals so as to minimize the expected discounted sum of losses in profits due to deviations of the actual price from the profit-maximizing price:

$$\min_{(\sigma_\varepsilon,\sigma_\psi)\in\mathcal{R}_+^2} E\left[\sum_{t=1}^{\infty}\beta^t \frac{\bar{C}\left(\theta-1\right)\left(1+\frac{1-\alpha}{\alpha}\theta\right)}{2}\left(p_{int}-p_{int}^{\diamondsuit}\right)^2\right], \tag{25}$$

subject to: (i) the process for the profit-maximizing price

$$p_{int}^{\diamondsuit} = p_{int}^{\diamondsuit A} + p_{int}^{\diamondsuit S}, \tag{26}$$

with

$$p_{int}^{\diamondsuit A} = p_{int-1}^{\diamondsuit A} + \sigma_A u_t, \tag{27}$$

and

$$p_{int}^{\diamondsuit S} = p_{int-1}^{\diamondsuit S} + \sigma_S v_{nt}, \tag{28}$$

where u_t and v_{nt} follow independent, unit-variance Gaussian white noise processes; (ii) the optimal price setting decision in period t given information in period t

$$p_{int} = E\left[p_{int}^{\diamondsuit}|s_{in}^t\right], \tag{29}$$

where $s_{in}^t = \left(s_{in}^0, s_{in1}, s_{in2}, \ldots, s_{int}\right)$ is the sequence of all signals that the decision-maker in firm i in sector n has received up to period t; (iii) an assumption concerning the set of signal vectors that the decision-maker can choose from

$$s_{int} = \begin{pmatrix} s_{int}^A \\ s_{int}^S \end{pmatrix} = \begin{pmatrix} p_{int}^{\diamondsuit A} \\ p_{int}^{\diamondsuit S} \end{pmatrix} + \begin{pmatrix} \sigma_\varepsilon \varepsilon_{int} \\ \sigma_\psi \psi_{int} \end{pmatrix}, \tag{30}$$

where ε_{int} and ψ_{int} follow idiosyncratic, unit-variance Gaussian white noise processes that are independent of the u process and the v_n process as well as independent of each other;

[24] See Maćkowiak and Wiederholt (2009a) and Maćkowiak and Wiederholt (2009b).

and (iv) the constraint on information flow

$$\forall t = 1, 2, \ldots : \underbrace{H\left(p_{int}^{\diamondsuit A} | s_{in}^{t-1}\right) - H\left(p_{int}^{\diamondsuit A} | s_{in}^{t}\right)}_{\kappa^A} + \underbrace{H\left(p_{int}^{\diamondsuit S} | s_{in}^{t-1}\right) - H\left(p_{int}^{\diamondsuit S} | s_{in}^{t}\right)}_{\kappa^S} \leq \kappa. \quad (31)$$

Here $H(X|\mathcal{I})$ denotes the conditional entropy of X given the information set $\mathcal{I}$, which is a measure of the conditional uncertainty of X given $\mathcal{I}$. The difference $H\left(X_t | s_{in}^{t-1}\right) - H\left(X_t | s_{in}^{t}\right)$ is a measure of the reduction in uncertainty about X_t that is due to the new signal received in period t. Sims (2003, Section 5) suggests using this measure of uncertainty reduction to quantify the amount of information received by the decision-maker in period t. The information flow constraint (31) states that, in each period $t \geq 1$, the information flow is limited. The information flow concerning aggregate conditions, denoted κ^A, plus the information flow concerning sector-specific conditions, denoted κ^S, cannot exceed the value κ.

The optimal allocation of attention (i.e. a pair κ^A and κ^S with $\kappa^A + \kappa^S \leq \kappa$) is derived under two different assumptions about the value of the overall attention devoted to the price setting decision (i.e. κ). In the benchmark specification of the model, it is assumed that the decision-maker can choose the overall attention devoted to the price setting decision facing the cost function $c(\kappa) = \phi\kappa$, where $\phi > 0$ is the real marginal cost of devoting attention to the price setting decision. This cost can be interpreted as an opportunity cost (devoting more attention to the price setting decision means devoting less attention to some other decision) or a monetary cost (e.g. a wage payment). Formally, the term $[\beta/(1-\beta)]\, c(\kappa)$ is added to the objective (25) and the variable κ is added to the vector of choice variables. In an alternative specification of the model, it is assumed that κ is fixed. Similarities and differences of these two specifications are discussed below.

It is worth pointing out that the assumption that the noise terms in equation (30) are independent captures the idea that paying attention to aggregate conditions and paying attention to sector-specific conditions are separate activities. This assumption is discussed in detail and relaxed in Section VIIB of Maćkowiak and Wiederholt (2009a).

Finally, to abstract from transitional dynamics in conditional variances, it is assumed that at the end of period zero (i.e. after the decision-maker has chosen the precision of the signals) the decision-maker receives information such that the conditional variances of $p_{in1}^{\diamondsuit A}$

and $p_{in1}^{\diamondsuit S}$ given information in period zero equal the steady-state values of the conditional variances of $p_{int}^{\diamondsuit A}$ and $p_{int}^{\diamondsuit S}$ given information in period $t-1$. This simplifies computing the solution to problem (25)-(31).

To begin consider the price setting behavior for a given allocation of attention (i.e. for a given pair κ^A and κ^S). One can show that the price setting behavior for a given allocation of attention satisfies the following equation:[25]

$$\begin{aligned} p_{int}^{\diamondsuit} - p_{int} &= \sum_{l=0}^{\infty}\left[\left(2^{-2\kappa^A}\right)^{l+1}\sigma_A u_{t-l} - \left(2^{-2\kappa^A}\right)^{l}\left(2^{-\kappa^A}\right)\sigma_A \varepsilon_{int-l}\right] \\ &\quad + \sum_{l=0}^{\infty}\left[\left(2^{-2\kappa^S}\right)^{l+1}\sigma_S v_{nt-l} - \left(2^{-2\kappa^S}\right)^{l}\left(2^{-\kappa^S}\right)\sigma_S \psi_{int-l}\right], \end{aligned} \tag{32}$$

where $p_{int}^{\diamondsuit} - p_{int}$ is the difference between the profit-maximizing price and the actual price. The speed at which the gap $p_{int}^{\diamondsuit} - p_{int}$ closes after an innovation in the aggregate component, u_t, depends on the attention allocated to aggregate conditions, κ^A; and the speed at which the gap closes after an innovation in the sector-specific component, v_{nt}, depends on the attention allocated to sector-specific conditions, κ^S. Hence, if the decision-maker pays more attention to sector-specific conditions than to aggregate conditions ($\kappa^S > \kappa^A$), the price set by firm i in sector n responds faster to sector-specific shocks than to aggregate shocks.

The remaining question is the following. How much attention will the decision-maker devote to aggregate conditions and how much attention will the decision-maker devote to sector-specific conditions? Substituting the price setting behavior for a given allocation of attention into expression (25) yields the expected discounted sum of profit losses for a given allocation of attention:[26]

$$\frac{\beta}{1-\beta}\frac{\bar{C}\left(\theta-1\right)\left(1+\frac{1-\alpha}{\alpha}\theta\right)}{2}\left(\frac{\sigma_A^2}{2^{2\kappa^A}-1}+\frac{\sigma_S^2}{2^{2\kappa^S}-1}\right). \tag{33}$$

It is now straightforward to derive the optimal allocation of attention.

When the decision-maker in a firm chooses the overall attention devoted to the price setting decision facing the cost function $c(\kappa) = \phi\kappa$, the decision-maker equates the marginal value of attending to aggregate conditions to the marginal cost of attention. Furthermore,

[25] See Appendix C.

[26] See again Appendix C.

the decision-maker equates the marginal value of attending to sector-specific conditions to the marginal cost of attention. Formally,

$$\frac{\bar{C}\left(\theta-1\right)\left(1+\frac{1-\alpha}{\alpha}\theta\right)}{2}\frac{\sigma_A^2 2^{2\kappa^A}}{\left(2^{2\kappa^A}-1\right)^2}2\ln\left(2\right)=\phi, \tag{34}$$

and

$$\frac{\bar{C}\left(\theta-1\right)\left(1+\frac{1-\alpha}{\alpha}\theta\right)}{2}\frac{\sigma_S^2 2^{2\kappa^S}}{\left(2^{2\kappa^S}-1\right)^2}2\ln\left(2\right)=\phi. \tag{35}$$

Rearranging these two equations yields

$$2^{\kappa^A}-\frac{1}{2^{\kappa^A}}=\sigma_A\sqrt{\frac{\bar{C}\left(\theta-1\right)\left(1+\frac{1-\alpha}{\alpha}\theta\right)}{\phi}\ln\left(2\right)}, \tag{36}$$

and

$$2^{\kappa^S}-\frac{1}{2^{\kappa^S}}=\sigma_S\sqrt{\frac{\bar{C}\left(\theta-1\right)\left(1+\frac{1-\alpha}{\alpha}\theta\right)}{\phi}\ln\left(2\right)}. \tag{37}$$

The model predicts that price setters devote more attention to aggregate conditions when aggregate conditions are more volatile. Similarly, the model predicts that price setters devote more attention to sector-specific conditions when sector-specific conditions are more volatile. Dividing equation (37) by equation (36) yields

$$\frac{2^{\kappa^S}-2^{-\kappa^S}}{2^{\kappa^A}-2^{-\kappa^A}}=\frac{\sigma_S}{\sigma_A}. \tag{38}$$

One arrives at the following prediction of the model. When the sector-specific component of the profit-maximizing price is more volatile than the aggregate component of the profit-maximizing price ($\sigma_S > \sigma_A$), price setters devote more attention to sector-specific conditions than to aggregate conditions ($\kappa^S > \kappa^A$), implying that prices respond faster to sector-specific shocks than to aggregate shocks.

In the derivation above it was assumed that decision-makers in firms can choose the overall attention devoted to the price setting decision facing the cost function $c\left(\kappa\right)=\phi\kappa$. When one assumes instead that κ is fixed, the optimal allocation of attention is given by: (i) the optimality condition that the marginal value of attending to aggregate conditions has to equal the marginal value of attending to sector-specific conditions, and (ii) the constraint $\kappa^A+\kappa^S=\kappa$. The optimality condition is exactly equation (38). Hence, the predictions concerning the relative speed of response of prices to shocks are the same as before. The

difference is that now the attention devoted to aggregate conditions depends both positively on σ_A and negatively on σ_S. This can be seen by substituting the constraint $\kappa^A + \kappa^S = \kappa$ into the optimality condition (38). By contrast, with a constant marginal cost of attention, the attention devoted to aggregate conditions does not depend on σ_S. See equation (36).

Finally, integrating equation (32) over all i and using equations (9) and (26)-(28) yields an equation for the sectoral inflation rate that has the form of equation (1), which is the equation that we estimate.[27]

The remainder of this subsection summarizes several predictions of this model and points out additional predictions of the model. First, if in a sector the sector-specific component of the profit-maximizing price is more volatile than the aggregate component of the profit-maximizing price, then the sectoral price index responds faster to sector-specific shocks than to aggregate shocks.

Second, a sectoral price index responds faster to aggregate shocks the larger the standard deviation of the profit-maximizing price due to aggregate shocks; and a sectoral price index responds faster to sector-specific shocks the larger the standard deviation of the profit-maximizing price due to sector-specific shocks. Furthermore, when price setters in firms face an exogenous information-processing limit or when price setters in firms can decide to process more information subject to a strictly convex cost function, the speed of response of a sectoral price index to aggregate shocks depends both positively on the standard deviation of the profit-maximizing price due to aggregate shocks and negatively on the standard deviation of the profit-maximizing price due to sector-specific shocks.

Third, if on average across sectors the sector-specific component of the profit-maximizing price is more volatile than the aggregate component of the profit-maximizing price, then the cross-sectional variation in the speed of response of sectoral price indexes to sector-specific shocks is smaller than the cross-sectional variation in the speed of response of sectoral price indexes to aggregate shocks. Intuitively, when price setters already pay close attention to sector-specific shocks, increasing the standard deviation of sector-specific shocks has little

[27] In the Calvo model, the sectoral price level is given by equations (13)-(14). In the sticky-information model, the sectoral price level is given by equation (22). Hence, in these two models, the equation for the sectoral inflation rate also has the form of equation (1) when the profit-maximizing price (12) follows a Gaussian process with a time-invariant moving-average representation.

effect on the speed of response of prices to sector-specific shocks. Formally, equation (32) and the equations characterizing the optimal allocation of attention imply that the speed of response of prices to a given type of shock is concave in the standard deviation of the shock.

Fourth, if on average across sectors the sector-specific component of the profit-maximizing price is more volatile than the aggregate component of the profit-maximizing price, then the coefficient in the regression reported in the first row of Table 2 should be larger than the coefficient in the regression reported in the third row of Table 2. The reason is again that, according to this model, the speed of response of prices to a given type of shock is concave in the standard deviation of the shock.

It is remarkable that all these predictions are supported by the data.

Boivin, Giannoni, and Mihov (2009) find a positive coefficient in the regression of the speed of response of sectoral price indexes to aggregate shocks on the standard deviation of sectoral inflation due to sector-specific shocks. It is worth pointing out that this finding is consistent with the rational-inattention model described above. The reasons are as follows. First, the model predicts a positive coefficient in the regression of the speed of response to aggregate shocks on the standard deviation of sectoral inflation due to aggregate shocks. Second, the model predicts a negative coefficient in the regression of the speed of response to aggregate shocks on the standard deviation of sectoral inflation due to sector-specific shocks *after controling* for the standard deviation of sectoral inflation due to aggregate shocks. Third, in the data there is a strong positive relationship between the standard deviation of sectoral inflation due to aggregate shocks and the standard deviation of sectoral inflation due to sector-specific shocks. Therefore, a positive coefficient in the regression of the speed of response to aggregate shocks on the standard deviation of sectoral inflation due to sector-specific shocks is to be expected when one fails to control for the standard deviation of sectoral inflation due to aggregate shocks.[28]

[28] The model also offers two explanations for the positive relationship in the data between the standard deviation of sectoral inflation due to aggregate shocks and the standard deviation of sectoral inflation due to sector-specific shocks. First, there may simply be a positive relationship between the volatility of the profit-maximizing price due to aggregate shocks and the volatility of the profit-maximizing price due to sector-specific shocks. Second, the parameters in objective (25) may differ across sectors. In sectors where

7.5 Menu Cost Model

Since sectoral price indexes respond quickly to sector-specific shocks, which are large on average, and sectoral price indexes respond slowly to aggregate shocks, which are small on average, one could imagine that a menu cost model can match the empirical findings presented in Sections 4-5. However, in a menu cost model, when a firm changes its price, the firm responds to both aggregate and sector-specific conditions independent of what triggered the price change. Thus, when firms respond quickly to sector-specific shocks and sector-specific shocks hit frequently, then firms also respond quickly to aggregate shocks. Hence, it seems that a menu cost model that can match the empirical findings presented in Sections 4-5 would have to be a menu cost model with infrequent sector-specific shocks.[29] However, note that Section 6.1 provides evidence suggesting that: (i) after dropping a few outlier sectors, sector-specific shocks are approximately Gaussian, and (ii) the empirical findings presented in Sections 4-5 are not driven by the few sectors experiencing non-Gaussian sector-specific shocks.

8 Conclusions

In order to evaluate models of price setting, this paper estimates a dynamic factor model using sectoral price data. Three kinds of results emerge. First, the median impulse responses of sectoral consumer price indexes have the following shapes. 100 percent of the long-run response of a sectoral price index to a sector-specific shock occurs in the month of the shock. By contrast, only 15 percent of the long-run response of a sectoral price index to an aggregate shock occurs in the month of the shock. Second, there is little cross-sectional variation in the speed of response to sector-specific shocks, while there is considerable cross-sectional variation in the speed of response to aggregate shocks. Third, the results from several regressions are reported.

pricing mistakes are more costly, firms pay more attention to both aggregate and sector-specific conditions, implying that prices in those sectors respond faster to both aggregate and sector-specific shocks. This raises both the standard deviation of sectoral inflation due to aggregate shocks and the standard deviation of sectoral inflation due to sector-specific shocks.

[29] For a menu cost model with a similar assumption, see Gertler and Leahy (2008).

The rational-inattention model developed in Maćkowiak and Wiederholt (2009a) can match most of these empirical findings. The key features of this model are: (i) a quick response to sector-specific conditions does not imply a quick response to aggregate conditions, and (ii) the speed of response to a given type of shock is positively related to the volatility of the shock. The assumption that attending to aggregate conditions and attending to sector-specific conditions are separate activities is a sufficient condition, but no necessary condition, for the model to have these properties. See Section VIIB in Maćkowiak and Wiederholt (2009a).

The standard Calvo model and the standard sticky-information model have difficulties matching these empirical findings. We think that the way in which these models fail gives us a new perspective on these models and suggests ways to modify these models. In the future, it would be interesting to study more formally whether a menu cost model can match these findings. We conjecture that a menu cost model will have difficulties matching jointly the empirical findings reported above and the empirical distribution of sector-specific shocks reported in Section 6.1.

We hope that the empirical findings reported in this paper guide the development of models of pricing and/or information in the future.

A Econometric Details

Consider the dynamic factor model

$$\tilde{\pi}_{nt} = a_n'(L)\, u_t + \tilde{\pi}_{nt}^S, \tag{39}$$

$$\tilde{\pi}_{nt}^S = c_n(L)\, \tilde{\pi}_{nt}^S + \epsilon_{nt}, \tag{40}$$

where: (i) $\tilde{\pi}_{nt}$ is the zero mean, unit-variance, month-on-month inflation rate in sector n in period t, $n = 1, ..., N$, $t = 1, ..., T$; (ii) $u_t = (u_{1t}; ...; u_{Kt})'$ is a $K \times 1$ vector of factors satisfying $u_{kt} \sim N(0, 1)$ for each $k = 1, ..., K$; (iii) $a_n(L)$ is a $K \times 1$ polynomial in the lag operator of order M; (iv) $c_n(L)$ is a polynomial in the lag operator of order S satisfying $c_{n0} = 0$; and (v) $\epsilon_{nt} \sim N(0, \sigma_n^2)$, for each n.

The raw data, described in Section 3 of the paper, are monthly sectoral price indexes. The following adjustments of the raw data were made: The log of the price index in each sector was seasonally adjusted. The month-on-month inflation rate was constructed. The mean was subtracted from each sector's inflation rate. Each sector's inflation rate was divided by its standard deviation. Outliers were eliminated. The reason why outliers were eliminated was that spikes in individual inflation rates were picked up by the factors in preliminary estimation. Specifically, for each sector, the observations falling outside four times the standard deviation of the inflation rate were replaced by the mean inflation rate over the rest of the sample period. This procedure labeled as outliers about three-fourth of one percent of all observations, with 36 sectors having no outliers.

When one assembles N equations, with each equation having the form of equation (39), one obtains

$$\tilde{\pi}_t = a(L)\, u_t + \tilde{\pi}_t^S,$$

where $\tilde{\pi}_t$ and $\tilde{\pi}_t^S$ and vectors of length N, and each matrix appearing in the polynomial $a(L)$ has size $N \times K$. Let $\tilde{a}_0$ denote the $K \times K$ matrix consisting of the first K rows of a_0. The matrix $\tilde{a}_0$ is assumed to be a lower triangular matrix with strictly positive entries on the main diagonal. This is a sufficient condition to estimate u_t uniquely. See Geweke and Zhou (1996). Consider the case when $K = 1$. In order to ensure that the factor is positively correlated with the CPI inflation rate, the CPI inflation rate is added to the dataset, and

the CPI inflation rate is ordered first. As it turns out, the model attributes almost all of the variance in the CPI inflation rate to innovations in the factor. In the case when $K = 2$, the variable ordered second (after the CPI inflation rate) is the variable best explained by the first principal component of the normalized inflation rates.

The Gibbs sampler with a Metropolis-Hastings step is used to sample from the joint posterior density of the factors $\{u_t\}$ and the model's parameters ($a_n(L)$, $c_n(L)$, and σ_n, for each n). Given a Monte Carlo draw of the model's parameters, one samples from the conditional posterior density of the factors given the model's parameters. This step is described in Section A.1. Afterwards, given a Monte Carlo draw of the factors, one samples from the conditional posterior density of the model's parameters given the factors. This step is described in Section A.2.

A.1 Sampling Factors Given Parameters

This section follows Carter and Kohn (1994) and Kim and Nelson (1999) in sampling from the conditional posterior density of the factors given the model's parameters ($a_n(L)$, $c_n(L)$, and σ_n, for each n). One begins by writing the model (39)-(40) in state space form. Equations (39)-(40) imply that

$$\tilde{\pi}^*_{nt} = g'_n(L)\, u_t + \epsilon_{nt}, \tag{41}$$

where $\tilde{\pi}^*_{nt} = (1 - c_n(L))\,\tilde{\pi}_{nt}$, and $g_n(L) = (1 - c_n(L))\, a_n(L)$, for each n. If one thinks of the coefficients appearing in the polynomial $g_n(L)$ as forming a vector G_n, and if one defines a vector F_t as $F_t = \left(u_t; u_{t-1}; ...; u_{t-(M+S)}\right)$, one can write equation (41) as

$$\tilde{\pi}^*_{nt} = G'_n F_t + \epsilon_{nt}.$$

Note that vectors G_n and F_t have length $l = K(M + S + 1)$. When one assembles N equations of this form, one arrives at the observation equation:

$$\tilde{\pi}^*_t = GF_t + \epsilon_t, \tag{42}$$

where $\tilde{\pi}^*_t$ and ϵ_t are vectors of length N, and G is a matrix of size $N \times l$. Let R denote the variance-covariance matrix of ϵ_t. Note that R is an $N \times N$ diagonal matrix with diagonal elements $\sigma_1^2, ..., \sigma_N^2$. The state equation is

$$F_{t+1} = JF_t + \tilde{u}_{t+1}, \tag{43}$$

where J is an $l \times l$ matrix defined as

$$J = \begin{bmatrix} 0_{K\times K} & 0_{K\times K(M+S)} \\ \mathcal{I}_{K(M+S)} & 0_{K(M+S)\times K} \end{bmatrix},$$

and $\tilde{u}_{t+1}$ is a vector of length l defined as $\tilde{u}_{t+1} = \left(u_{t+1}; 0_{K(M+S)\times 1}\right)$. Let Q denote the variance-covariance matrix of $\tilde{u}_{t+1}$. Note that Q is an $l \times l$ diagonal matrix, with the first K diagonal elements equal to unity and all other elements equal to zero.

One runs Kalman filter iterations from period $t = 1$ to period $t = T$ to obtain

$$F_{t|t} = F_{t|t-1} + P_{t|t-1}G'\left(GP_{t|t-1}G' + R\right)^{-1}\left(\tilde{\pi}_t^* - GF_{t|t-1}\right), \tag{44}$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}G'\left(GP_{t|t-1}G' + R\right)^{-1}GP_{t|t-1}, \tag{45}$$

$$F_{t+1|t} = JF_{t|t},$$

and

$$P_{t+1|t} = JP_{t|t}J' + Q.$$

The unconditional density of the state vector is used to initialize the Kalman filter, where $F_{1|0}$ is a zero vector and

$$vec\left(P_{1|0}\right) = \left[\mathcal{I}_{l^2} - (J \otimes J)\right]^{-1} vec\left(Q\right).$$

Next, one samples from the probability density function of F_T given the data until period T. This is a Gaussian density with mean $F_{T|T}$ and variance-covariance matrix $P_{T|T}$. Subsequently, for each $t = T-1, T-2, ..., 1$, one samples from the probability density function of F_t given the data until period t and F_{t+1}. This is a Gaussian density with mean $F_{t|t,F_{t+1}}$ and variance-covariance matrix $P_{t|t,F_{t+1}}$, where

$$F_{t|t,F_{t+1}} = F_{t|t} + P_{t|t}J'P_{t+1|t}^{-1}\left(F_{t+1} - F_{t+1|t}\right),$$

and

$$P_{t|t,F_{t+1}} = P_{t|t} - P_{t|t}J'P_{t+1|t}^{-1}JP_{t|t}.$$

In practice, since in the model in this paper Q is a singular matrix, one must modify the densities one samples from slightly. Let Q^* be the matrix of size $l^* \times l^*$ (where $l^* < l$)

obtained after removing from Q each row that contains nothing but zeros. Furthermore, let F_{t+1}^* be the corresponding $l^* \times 1$ state vector, and let J^* be the corresponding $l^* \times l$ matrix in the state equation. For each $t = T-1, T-2, ..., 1$, one samples from the probability density function of F_t given the data until period t and F_{t+1}^*. This is a Gaussian density with mean $F_{t|t,F_{t+1}^*}$ and variance-covariance matrix $P_{t|t,F_{t+1}^*}$, where

$$F_{t|t,F_{t+1}^*} = F_{t|t} + P_{t|t}J^{*\prime}\left(J^*P_{t|t}J^{*\prime} + Q^*\right)^{-1}\left(F_{t+1}^* - J^*F_{t|t}\right),$$

and

$$P_{t|t,F_{t+1}^*} = P_{t|t} - P_{t|t}J^{*\prime}\left(J^*P_{t|t}J^{*\prime} + Q^*\right)^{-1}J^*P_{t|t}.$$

In the model in this paper, it turns out that $F_{t|t,F_{t+1}^*} = F_{t|t}$ and $P_{t|t,F_{t+1}^*} = P_{t|t}$, so that the Gaussian probability density functions one samples from are characterized fully by expressions (44) and (45). The reason is that the factors in the model in this paper follow a vector white noise process.

A.2 Sampling Parameters Given Factors

This section follows Chib and Greenberg (1994) in sampling from the conditional posterior density of the model's parameters ($a_n(L)$, $c_n(L)$, and σ_n, for each n) given the factors. Taking the factors as given and collecting N pairs of equations (39)-(40) yields N independent Gaussian regressions, each regression with autoregressive errors of order S. Therefore, given the factors, one can analyze each pair of equations (39)-(40) using the results of Chib and Greenberg (1994) on Gaussian regression with autoregressive errors. Note that, like Chib and Greenberg (1994), this paper uses the full likelihood function without conditioning on initial observations.

In the rest of this section, for simplicity, it is convenient to drop the subscript n, to think of the coefficients appearing in the polynomial $a(L)$ as forming a vector θ, and to think of the coefficients appearing in the polynomial $c(L)$ as forming a vector ϕ. Furthermore, it is useful to define a vector x_t according to $x_t = (u_t; u_{t-1}; ...; u_{t-M})$, and it is useful to define a vector $\tilde{x}_t$ according to $\tilde{x}_t = \left(\tilde{\pi}_{t-1}^S; ...; \tilde{\pi}_{t-S}^S\right)$. Then, one can write the model (39)-(40) as

$$\tilde{\pi}_t = x_t'\theta + \tilde{\pi}_t^S,$$

$$\tilde{\pi}_t^S = \tilde{x}_t'\phi + \epsilon_t,$$

and, in matrix notation,

$$\tilde{\Pi} = X\theta + \tilde{\Pi}^S,$$

$$\tilde{\Pi}^S = \tilde{X}\phi + \epsilon.$$

Note that vectors x_t and θ have length $K(M+1)$, vectors $\tilde{x}_t$ and ϕ have length S, vectors $\tilde{\Pi}$, $\tilde{\Pi}^S$, and ϵ have length T, matrix X has size $T \times K(M+1)$, and matrix $\tilde{X}$ has size $T \times S$.

Let Φ be an $S \times S$ matrix satisfying

$$\Phi = \begin{bmatrix} \tilde{\phi}' & \phi_S \\ \mathcal{I}_{S-1} & 0_{(S-1)\times 1} \end{bmatrix},$$

where $\tilde{\phi} = \left(\phi_1; ...; \phi_{S-1}\right)$ is a vector of length $S-1$. The following joint prior density of θ, ϕ and σ^2 is assumed:

$$\left[\theta, \phi, \sigma^2\right] = [\theta]\,[\phi]\left[\sigma^2\right] = \mathcal{N}\left(\theta_0, \Theta_0^{-1}\right)\left[\mathcal{N}\left(\phi_0, \Phi_0^{-1}\right)\mathcal{I}(\Phi)\right]\mathcal{IG}\left[\nu_0/2, \delta_0/2\right],$$

where $\mathcal{N}$ denotes the normal density, $\mathcal{IG}$ denotes the inverse gamma density, and $\mathcal{I}(\Phi)$ is an indicator function equal to one when all eigenvalues of Φ are less than one in modulus.

The following notation is useful. $\tilde{\Pi}$ is partitioned so that $\tilde{\Pi} = \left(\tilde{\Pi}_1; \tilde{\Pi}_2\right)$, where $\tilde{\Pi}_1$ has length S and $\tilde{\Pi}_2$ has length $T-S$. Analogously, X is partitioned so that $X = (X_1; X_2)$, where X_1 has size $S \times K(M+1)$ and X_2 has size $T-S \times K(M+1)$. Σ is defined as the $S \times S$ matrix satisfying

$$\Sigma = \Phi\Sigma\Phi' + \left(1; 0_{(S-1)\times 1}\right)\left(1; 0_{(S-1)\times 1}\right)',$$

where $\left(1; 0_{(S-1)\times 1}\right)$ is a vector of length S. That is,

$$vec(\Sigma) = \left[\mathcal{I}_{S^2} - (\Phi \otimes \Phi)\right]^{-1} vec\left[\left(1; 0_{(S-1)\times 1}\right)\left(1; 0_{(S-1)\times 1}\right)'\right].$$

Let $chol(\Sigma)$ denote the lower triangular Choleski square root of Σ. Define $\tilde{\Pi}^* = \left(\tilde{\Pi}_1^*; \tilde{\Pi}_2^*\right)$ and $X^* = (X_1^*; X_2^*)$, where: (i) $\tilde{\Pi}_1^* = [chol(\Sigma)]^{-1}\tilde{\Pi}_1$; (ii) $X_1^* = [chol(\Sigma)]^{-1} X_1$; (iii) $\tilde{\Pi}_2^*$ is a vector of length $T-S$ with t'th row given by $[1 - c(L)]\,\tilde{\pi}_t$; and (iv) X_2^* is a matrix of size $T-S \times K(M+1)$ with t'th row given by $[1 - c(L)]\,x_t$. Let $e_t = \tilde{\pi}_t - x_t'\theta$, for each

$t = S+1, ..., T$; let e be a vector of length $T-S$ satisfying $e = (e_{S+1}; ...; e_T)$; and let E be a matrix of size $T-S \times S$ with t'th row given by $(e_{t-1}, ..., e_{t-S})$, for each $t \geq S+1$.

Chib and Greenberg (1994) derive the following conditional posterior densities:

$$\theta \mid \phi, \sigma^2 \sim \mathcal{N}\left[\Theta_T^{-1}\left(\Theta_0\theta_0 + \sigma^{-2}X^{*\prime}\tilde{\Pi}^*\right), \Theta_T^{-1}\right], \tag{46}$$

$$\sigma^2 \mid \phi, \theta \sim \mathcal{IG}\left[\left(\frac{\nu_0 + T}{2}\right), \frac{\delta_0 + \left(\tilde{\Pi}^* - X^*\theta\right)'\left(\tilde{\Pi}^* - X^*\theta\right)}{2}\right], \tag{47}$$

$$\phi \mid \theta, \sigma^2 \propto \Psi(\phi) \times \mathcal{N}\left[\Phi_T^{-1}\left(\Phi_0\phi_0 + \sigma^{-2}E'e\right), \Phi_T^{-1}\right]\mathcal{I}(\Phi), \tag{48}$$

where

$$\Theta_T = \Theta_0 + \sigma^{-2}X^{*\prime}X^*,$$

$$\Phi_T = \Phi_0 + \sigma^{-2}E'E,$$

$$\Psi(\phi) = \mid \Sigma(\phi) \mid^{-1/2} \exp\left[-\frac{1}{2\sigma^2}\left(\tilde{\Pi}_1 - X_1\theta\right)'\Sigma(\phi)^{-1}\left(\tilde{\Pi}_1 - X_1\theta\right)\right],$$

and ϕ in brackets in the last expression reminds us that Σ depends on ϕ.

The conditional density of θ and the conditional density of σ^2 are standard, but the conditional density of ϕ cannot be sampled from directly. Following Chib and Greenberg (1994), this paper samples from the density of ϕ using the Metropolis-Hastings algorithm. See also Otrok and Whiteman (1998). At each iteration j of the Gibbs sampler, one generates a candidate draw ϕ^* from the density $\mathcal{N}\left[\Phi_T^{-1}\left(\Phi_0\phi_0 + \sigma^{-2}E'e\right), \Phi_T^{-1}\right]\mathcal{I}(\Phi)$. One then sets $\phi^j = \phi^*$ with probability

$$\rho = \min\left[\frac{\Psi(\phi^*)}{\Psi\left(\phi^{j-1}\right)}, 1\right],$$

and one sets $\phi^j = \phi^{j-1}$ with probability $1-\rho$.

Values for the prior hyperparameters are chosen following the Minnesota prior. It is assumed that $\theta_0 = 0_{K(M+1)\times 1}$. Furthermore, Θ_0 is assumed to be a diagonal matrix of size $K(M+1) \times K(M+1)$ with the following entries on the main diagonal: each of the first K entries equals 1, each of the subsequent K entries also equals 1, each of the K entries after that equals 4, and so on until the last K entries on the main diagonal of Θ_0 equal M^2. These assumptions imply that the prior mean on each loading equals zero, the standard

deviation of the prior on a contemporaneous loading equals 1, and the standard deviation of the prior on a loading on an m'th lag of the factor equals $(1/m)$. Next, it is assumed that $\phi_0 = 0_{S\times 1}$. Furthermore, Φ_0 is assumed to be a diagonal matrix of size $S \times S$ with entries on the main diagonal given by $(s/0.2)^2$, $s = 1, ..., S$. These assumptions imply that the prior mean on each autoregressive coefficient equals zero, the standard deviation of the prior on a coefficient on a first lag equals 0.2, and the standard deviation of the prior on a coefficient on an s'th lag equals $(0.2/s)$. Finally, it is assumed that $\nu_0 = 4$ and $\delta_0 = 0.1$. Note that a tighter prior is used in the case when $K = 2$. The standard deviation of the prior on a loading on an m'th lag of a factor equals $\left(1/m^{1.5}\right)$. In the autoregressive component of the model, the standard deviation of the prior on a coefficient on an s'th lag equals $\left(0.2/s^{1.5}\right)$.

To initialize the Gibbs sampler, the data are regressed on, alternatively, the first K principal components of the data or randomly generated K "indexes" (current and lagged). The regression coefficients are used as initial values for the model's parameters, $\left\{\theta_n, \phi_n, \sigma_n^2\right\}^{j=0}$, $n = 1, ..., N$. Each j'th iteration of the Gibbs sampler proceeds as follows. Given $\left\{\theta_n, \phi_n, \sigma_n^2\right\}^{j-1}$, $n = 1, ..., N$, make a draw of the factors $\{u_t\}^j$, as described in Section A.1. Next, given $\{u_t\}^j$, make a draw of the model's parameters $\left\{\theta_n, \phi_n, \sigma_n^2\right\}^j$, $n = 1, ..., N$, as described in this section above. Here, begin by drawing θ_n^j given $\left\{\phi_n, \sigma_n^2\right\}^{j-1}$ from density (46), afterwards draw $\left(\sigma_n^2\right)^j$ given θ_n^j and ϕ_n^{j-1} from density (47), and finally draw ϕ_n^j given $\left\{\theta_n, \sigma_n^2\right\}^j$ from density (48). 20000 draws are made. The initial 5000 draws are discarded. Every second draw is saved out of the remaining 15000 draws. This procedure yields 7500 draws from the posterior density.

A.3 Convergence of the Gibbs Sampler

The paper uses formal and informal diagnostics to assess convergence of the Gibbs sampler. Two formal convergence diagnostics are computed: the Raftery-Lewis measure of the number of draws required to achieve a certain precision of the sampler (see Raftery and Lewis, 1992); and the Geweke relative numerical efficiency indicator (see Geweke, 1992). The parameters for the Raftery-Lewis diagnostic are set as follows: quantile = 0.025; desired accuracy = 0.0125; required probability of attaining the desired accuracy = 0.95. Note that, since we compute the Raftery-Lewis diagnostic and the Geweke diagnostic for each

parameter in the model, we obtain a cross-section of the Raftery-Lewis diagnostics and a cross-section of the Geweke diagnostics across the parameters. Table A.1 summarizes the cross-section of the Raftery-Lewis diagnostics across the parameters. Table A.2 summarizes the cross-section of the Geweke diagnostics across the parameters. Both tables refer to the benchmark specification. For ease of exposition, in both tables the parameters are grouped into the loadings (the θ's), the autoregressive parameters (the ϕ's), and the standard deviations (the σ's). Consider Table A.1. For 99 percent of the parameters, the Raftery and Lewis diagnostic suggests that one should make 4278 draws or fewer for the sampler to be precise. 4278 draws are a lot less than 20000 draws actually made here. Only for two parameters the Raftery and Lewis diagnostic suggests that one should make many more draws than 20000. Both of these parameters are autoregressive parameters in a single sector, "Tires". In this sector, the acceptance rate in the Metropolis-Hastings step of the Gibbs sampler turns out to be relatively low. Next, consider Table A.2. With only few exceptions, the Geweke indicator lies well below 20, which is the value of the Geweke indicator considered as small enough to signal good mixing properties of the sampler. See, for example, Primiceri (2005). Convergence of the Gibbs sampler was also monitored informally, by plotting the evolution of draws for a set of randomly selected parameters. Furthermore, initializing the Gibbs sampler at random points yielded very similar results. The Gibbs sampler was also run on artificial data. The estimated factors and the estimated parameters were very close to the true ones. All of this give us confidence that the Markov chain used here has converged to its ergodic distribution.

B Proof of Proposition 1

First, let x denote the long-run response of the price index for sector n to a shock in period t. Second, the price index for sector n in period t satisfies $p_{nt} = p_{nt-1} + x$ if and only if the price set by adjusting firms in sector n in period t satisfies

$$p^*_{int} = p_{nt-1} + \frac{1}{\lambda_n} x. \tag{49}$$

This follows from equation (14). Third, the sectoral price index satisfies $p_{nt+\tau} = p_{nt}$ for all $\tau \geq 1$ if and only if the price set by adjusting firms satisfies

$$\forall \tau \geq 1 : p^*_{int+\tau} = p_{nt}. \tag{50}$$

This follows from equation (14). Combining these two results yields that $p_{nt+\tau} = p_{nt-1} + x$ for all $\tau \geq 0$ if and only if

$$p^*_{int+\tau} = \begin{cases} p_{nt-1} + \frac{1}{\lambda_n} x & \text{for } \tau = 0 \\ p_{nt-1} + x & \text{for all } \tau \geq 1 \end{cases}. \tag{51}$$

Fourth, the adjustment price in period $t + \tau$ satisfies equation (51) for all $\tau \geq 1$ if and only if the profit-maximizing price in period $t + \tau$ satisfies

$$\forall \tau \geq 1 : p^{\diamondsuit}_{int+\tau} = p_{nt-1} + x. \tag{52}$$

This follows from equation (13). Fifth, given equation (52), the adjustment price in period t satisfies equation (51) if and only if the profit-maximizing price in period t satisfies

$$p^{\diamondsuit}_{int} = p_{nt-1} + \frac{\frac{1}{\lambda_n} - (1 - \lambda_n)\beta}{1 - (1 - \lambda_n)\beta} x. \tag{53}$$

This follows again from equation (13). Collecting results yields that $p_{nt+\tau} = p_{nt-1} + x$ for all $\tau \geq 0$ if and only if

$$p^{\diamondsuit}_{int+\tau} = \begin{cases} p_{nt-1} + \frac{\frac{1}{\lambda_n} - (1-\lambda_n)\beta}{1-(1-\lambda_n)\beta} x & \text{for } \tau = 0 \\ p_{nt-1} + x & \text{for all } \tau \geq 1 \end{cases}. \tag{54}$$

C Solving the Rational-Inattention Model

Kalman filtering: The state-space representation of the dynamics of the signal concerning aggregate conditions is

$$p^{\diamondsuit A}_{int} = p^{\diamondsuit A}_{int-1} + \sigma_A u_t, \tag{55}$$

$$s^A_{int} = p^{\diamondsuit A}_{int} + \sigma_\varepsilon \varepsilon_{int}. \tag{56}$$

The first equation is the state equation and the second equation is the observation equation. For ease of exposition, the following notation is used in this appendix: $X_t = p^{\diamondsuit A}_{int}$, $S_t = s^A_{int}$,

$X_{t|t} = E\left[X_t|S^t\right]$, $X_{t|t-1} = E\left[X_t|S^{t-1}\right]$, $\sigma^2_{t|t} = Var\left[X_t|S^t\right]$ and $\sigma^2_{t|t-1} = Var\left[X_t|S^{t-1}\right]$. The usual Kalman filter equations yield

$$X_{t|t} = X_{t|t-1} + \frac{\sigma^2_{t|t-1}}{\sigma^2_{t|t-1} + \sigma^2_\varepsilon}\left(S_t - X_{t|t-1}\right),$$

and

$$\sigma^2_{t|t} = \sigma^2_{t|t-1} - \frac{\sigma^2_{t|t-1}}{\sigma^2_{t|t-1} + \sigma^2_\varepsilon}\sigma^2_{t|t-1}.$$

Furthermore,

$$X_{t+1|t} = X_{t|t},$$

and

$$\sigma^2_{t+1|t} = \sigma^2_{t|t} + \sigma^2_A. \tag{57}$$

Substituting the last two equations into the two equations before yields

$$X_{t|t} = X_{t-1|t-1} + \frac{\sigma^2_{t-1|t-1} + \sigma^2_A}{\sigma^2_{t-1|t-1} + \sigma^2_A. + \sigma^2_\varepsilon}\left(S_t - X_{t-1|t-1}\right), \tag{58}$$

and

$$\sigma^2_{t|t} = \frac{\sigma^2_\varepsilon}{\sigma^2_{t-1|t-1} + \sigma^2_A + \sigma^2_\varepsilon}\left(\sigma^2_{t-1|t-1} + \sigma^2_A\right). \tag{59}$$

When $\sigma^2_{t|t} = \sigma^2_{t-1|t-1}$, the unique positive solution to the last equation is

$$\bar{\sigma}^2_{t|t} = \left(-\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\sigma^2_\varepsilon}{\sigma^2_A}}\right)\sigma^2_A. \tag{60}$$

The information flow constraint: When X_t and $S^t = \left(S^0, S_1, S_2, \ldots, S_t\right)$ have a multivariate Gaussian distribution, the conditional distribution of X_t given S^t is Gaussian. In this case, the conditional entropy of X_t given S^t is a simple function of the conditional variance of X_t given S^t

$$H\left(X_t|S^t\right) = \frac{1}{2}\log_2\left(2\pi e\sigma^2_{t|t}\right).$$

Similarly,

$$H\left(X_t|S^{t-1}\right) = \frac{1}{2}\log_2\left(2\pi e\sigma^2_{t|t-1}\right).$$

The equation

$$H\left(X_t|S^{t-1}\right) - H\left(X_t|S^t\right) = \kappa^A,$$

then reduces to

$$\frac{1}{2}\log_2\left(\frac{\sigma^2_{t|t-1}}{\sigma^2_{t|t}}\right) = \kappa^A.$$

Using equation (57) yields

$$\frac{1}{2}\log_2\left(\frac{\sigma^2_{t-1|t-1}+\sigma^2_A}{\sigma^2_{t|t}}\right) = \kappa^A. \tag{61}$$

When $\sigma^2_{t|t} = \sigma^2_{t-1|t-1}$, the last equation becomes

$$\bar{\sigma}^2_{t|t} = \frac{\sigma^2_A}{2^{2\kappa^A}-1}. \tag{62}$$

The assumption concerning initial information (s^0_{in}): To abstract from (purely deterministic) transitional dynamics in conditional variances, it is assumed that, after the decision-maker has chosen the allocation of attention (i.e. a pair κ^A and κ^S) in period zero, the decision-maker receives information in period zero such that the conditional variance of X_1 given information in period zero equals the steady-state value of the conditional variance of X_t given information in period $t-1$. This assumption implies that in period one the conditional variance of X_t given information in period t equals its steady-state value

$$\sigma^2_{1|1} = \bar{\sigma}^2_{t|t}. \tag{63}$$

This simplifies computing the solution to problem (25)-(31).

Variance of noise, value of the objective and pricing behavior for a given allocation of attention: Equating the right-hand sides of equations (60) and (62) yields the variance of noise in the signal concerning aggregate conditions, σ^2_ε, for given attention allocated to aggregate conditions, κ^A,

$$\begin{aligned}\sigma^2_\varepsilon &= \frac{1}{4}\left[\left(\frac{2^{2\kappa^A}+1}{2^{2\kappa^A}-1}\right)^2 - 1\right]\sigma^2_A \\ &= \frac{2^{2\kappa^A}}{\left(2^{2\kappa^A}-1\right)^2}\sigma^2_A. \end{aligned} \tag{64}$$

Similarly, the variance of noise in the signal concerning sector-specific conditions, σ^2_ψ, for given attention allocated to sector-specific conditions, κ^S, is

$$\sigma^2_\psi = \frac{2^{2\kappa^S}}{\left(2^{2\kappa^S}-1\right)^2}\sigma^2_S. \tag{65}$$

Furthermore, using the fact that

$$\begin{aligned} E\left[\left(p_{int}-p_{int}^{\diamondsuit}\right)^2\right] &= E\left[\left(p_{int}^{\diamondsuit}-E\left[p_{int}^{\diamondsuit}|s_{in}^t\right]\right)^2|s_{in}^t\right] \\ &= Var\left(p_{int}^{\diamondsuit}|s_{in}^t\right) \\ &= Var\left(p_{int}^{\diamondsuit A}|s_{in}^t\right)+Var\left(p_{int}^{\diamondsuit S}|s_{in}^t\right) \\ &= \frac{\sigma_A^2}{2^{2\kappa^A}-1}+\frac{\sigma_S^2}{2^{2\kappa^S}-1}, \end{aligned}$$

yields the value of objective (25) for a given allocation of attention:

$$\frac{\beta}{1-\beta}\frac{\bar{C}\left(\theta-1\right)\left(1+\frac{1-\alpha}{\alpha}\theta\right)}{2}\left(\frac{\sigma_A^2}{2^{2\kappa^A}-1}+\frac{\sigma_S^2}{2^{2\kappa^S}-1}\right). \tag{66}$$

Next, consider the pricing behavior for a given allocation of attention. Substituting equations (62) and (64) into equation (58) yields

$$X_{t|t}=\left(2^{-2\kappa^A}\right)X_{t-1|t-1}+\left(1-2^{-2\kappa^A}\right)S_t.$$

Furthermore, using equations (55) and (56) yields

$$\left(X_t-X_{t|t}\right)=\left(2^{-2\kappa^A}\right)\left(X_{t-1}-X_{t-1|t-1}\right)+\left(2^{-2\kappa^A}\right)\sigma_A u_t-\left(1-2^{-2\kappa^A}\right)\sigma_\varepsilon\varepsilon_{int}.$$

In addition, using equation (64) to substitute for σ_ε yields

$$\left(X_t-X_{t|t}\right)=\left(2^{-2\kappa^A}\right)\left(X_{t-1}-X_{t-1|t-1}\right)+\left(2^{-2\kappa^A}\right)\sigma_A u_t-\left(2^{-\kappa^A}\right)\sigma_A\varepsilon_{int}.$$

Solving this difference equation by repeated substitution, one arrives at

$$\begin{aligned} X_t-X_{t|t} &= \sum_{l=0}^{t-2}\left(2^{-2\kappa^A}\right)^l\left[\left(2^{-2\kappa^A}\right)\sigma_A u_{t-l}-\left(2^{-\kappa^A}\right)\sigma_A\varepsilon_{int-l}\right] \\ &\quad +\left(2^{-2\kappa^A}\right)^{t-1}\left(X_1-X_{1|1}\right). \end{aligned}$$

Thus,

$$\begin{aligned} p_{int}^{\diamondsuit A}-E\left[p_{int}^{\diamondsuit A}|s_{in}^t\right] &= \sum_{l=0}^{t-2}\left(2^{-2\kappa^A}\right)^l\left[\left(2^{-2\kappa^A}\right)\sigma_A u_{t-l}-\left(2^{-\kappa^A}\right)\sigma_A\varepsilon_{int-l}\right] \\ &\quad +\left(2^{-2\kappa^A}\right)^{t-1}\left(p_{in1}^{\diamondsuit A}-E\left[p_{in1}^{\diamondsuit A}|s_{in}^1\right]\right). \end{aligned} \tag{67}$$

It follows from equations (26), (29), (67) and the corresponding equation for $p_{int}^{\diamondsuit S} - E\left[p_{int}^{\diamondsuit S}|s_{in}^{t}\right]$ that

$$\begin{aligned} p_{int}^{\diamondsuit} - p_{int} &= \sum_{l=0}^{t-2}\left(2^{-2\kappa^A}\right)^l\left[\left(2^{-2\kappa^A}\right)\sigma_A u_{t-l} - \left(2^{-\kappa^A}\right)\sigma_A \varepsilon_{int-l}\right] \\ &+ \sum_{l=0}^{t-2}\left(2^{-2\kappa^S}\right)^l\left[\left(2^{-2\kappa^S}\right)\sigma_S v_{nt-l} - \left(2^{-\kappa^S}\right)\sigma_S \psi_{int-l}\right] \\ &+ \left(2^{-2\kappa^A}\right)^{t-1}\left(p_{in1}^{\diamondsuit A} - E\left[p_{in1}^{\diamondsuit A}|s_{in}^{1}\right]\right) + \left(2^{-2\kappa^S}\right)^{t-1}\left(p_{in1}^{\diamondsuit S} - E\left[p_{in1}^{\diamondsuit S}|s_{in}^{1}\right]\right) \end{aligned} \tag{68}$$

Equation (68) and equations (26)-(28) already pin down the response of p_{int} to various shocks. The following equation is easier to read than equation (68). As $t \to \infty$ (or for the right values of $p_{in1}^{\diamondsuit A} - E\left[p_{in1}^{\diamondsuit A}|s_{in}^{1}\right]$ and $p_{in1}^{\diamondsuit S} - E\left[p_{in1}^{\diamondsuit S}|s_{in}^{1}\right]$), equation (68) becomes

$$\begin{aligned} p_{int}^{\diamondsuit} - p_{int} &= \sum_{l=0}^{\infty}\left(2^{-2\kappa^A}\right)^l\left[\left(2^{-2\kappa^A}\right)\sigma_A u_{t-l} - \left(2^{-\kappa^A}\right)\sigma_A \varepsilon_{int-l}\right] \\ &+ \sum_{l=0}^{\infty}\left(2^{-2\kappa^S}\right)^l\left[\left(2^{-2\kappa^S}\right)\sigma_S v_{nt-l} - \left(2^{-\kappa^S}\right)\sigma_S \psi_{int-l}\right]. \end{aligned} \tag{69}$$

The optimal allocation of attention: The optimal allocation of attention is derived in the main text from equation (66).

References

[1] Bernanke, B.S., Boivin, J., Eliasz, P., 2005. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. Quarterly Journal of Economics 120, 387-422.

[2] Bils, M., Klenow, P.J., 2004. Some evidence on the importance of sticky prices. Journal of Political Economy 112, 947-985.

[3] Boivin, J., Giannoni, M., Mihov, I., 2009. Sticky prices and monetary policy: Evidence from disaggregated U.S. data. American Economic Review 99, 350-384.

[4] Calvo, G., 1983. Staggered prices in a utility-maximizing framework. Journal of Monetary Economics 12, 383-398.

[5] Carter, C.K., Kohn, R., 1994. On Gibbs sampling for state space models. Biometrika 81, 541-553.

[6] Chib, S., Greenberg, E., 1994. Bayes inference in regression models with ARMA(p,q) errors. Journal of Econometrics 64, 183-206.

[7] Del Negro, M., Otrok, C., 2007. 99 Luftballons: Monetary policy and the house price boom across U.S. states. Journal of Monetary Economics 54, 1962-1985.

[8] Doz, C., Giannone, D., Reichlin, L., 2006. A quasi maximum likelihood approach for large approximate dynamic factor models. European Central Bank Working Paper 674.

[9] Engle, R., Watson, M., 1981. A one-factor multivariate time series model of metropolitan wage rates. Journal of the American Statistical Association 76, 774-781.

[10] Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2000. The generalized dynamic-factor model: Identification and estimation. Review of Economics and Statistics 82, 540-554.

[11] Gertler, M., Leahy, J., 2008. A Phillips curve with an Ss foundation. Journal of Political Economy 116, 533-572.

[12] Geweke, J., 1977. The dynamic factor analysis of economic time series. In: Aigner, D., Goldberger, A. (Eds.), Latent Variables in Socioeconomic Models. North-Holland, Amsterdam, pp. 365-383.

[13] Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo, J.M., et al. (Eds.), Bayesian Statistics 4. Oxford University Press, Oxford, pp. 169-193.

[14] Geweke, J., Singleton, K.J., 1981. Maximum likelihood 'confirmatory' factor analysis of economic time series. International Economic Review 22, 37-54.

[15] Geweke, J., Zhou, G., 1996. Measuring the pricing error of the arbitrage pricing theory. Review of Financial Studies 9, 557-587.

[16] Justiniano, A., 2004. Estimation and model selection in dynamic factor analysis. Discussion paper, Princeton University.

[17] Kehoe, P.J., Midrigan, V., 2007. Sticky prices and sectoral real exchange rates. Federal Reserve Bank of Minneapolis Research Department Working Paper 656.

[18] Kim, C.J., Nelson, C.R., 1999. State-space models with regime switching. The MIT Press, Cambridge, Massachusetts.

[19] Kose, A.M., Otrok, C., Whiteman, C.H., 2003. International business cycles: World, region, and country-specific factors. American Economic Review 93, 1216-1239.

[20] Maćkowiak, B., Wiederholt, M., 2009a. Optimal sticky prices under rational inattention. American Economic Review 99, 769-803.

[21] Maćkowiak, B., Wiederholt, M., 2009b. Business cycle dynamics under rational inattention. Discussion paper, European Central Bank and Northwestern University.

[22] Mankiw, N.G., Reis, R., 2002. Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. Quarterly Journal of Economics 117, 1295-1328.

[23] Nakamura, E., Steinsson, J., 2008. Five facts about prices: A reevaluation of menu cost models. Quarterly Journal of Economics 123, 1415-1464.

[24] Otrok, C., Whiteman, C.H., 1998. Bayesian leading indicators: Measuring and predicting economic conditions in Iowa. International Economic Review 39, 997-1014.

[25] Primiceri, G.E., 2005. Time varying structural vector autoregressions and monetary policy. Review of Economic Studies 72, 821-852.

[26] Quah, D., Sargent, T.J., 1992. A dynamic index model for large cross sections. Institute for Empirical Macroeconomics Federal Reserve Bank of Minneapolis Discussion Paper 77.

[27] Raftery, A.E., Lewis, S.M., 1992. How many iterations in the Gibbs sampler? In: Bernardo, J.M., et al. (Eds.), Bayesian Statistics 4. Oxford University Press, Oxford, pp. 763-773.

[28] Reis, R., 2006. Inattentive producers. Review of Economic Studies 73, 793-821.

[29] Reis, R., Watson, M.W., 2007a. Measuring changes in the value of the numeraire. Discussion paper, Princeton University.

[30] Reis, R., Watson, M.W., 2007b. Relative goods' prices and pure inflation. National Bureau of Economic Research Working Paper 13615.

[31] Sargent, T.J., Sims, C.A., 1977. Business cycle modeling without pretending to have too much a priori economic theory. In: Sims, C.A. (Ed.), New Methods of Business Cycle Research. Federal Reserve Bank, Minneapolis, pp. 45-109.

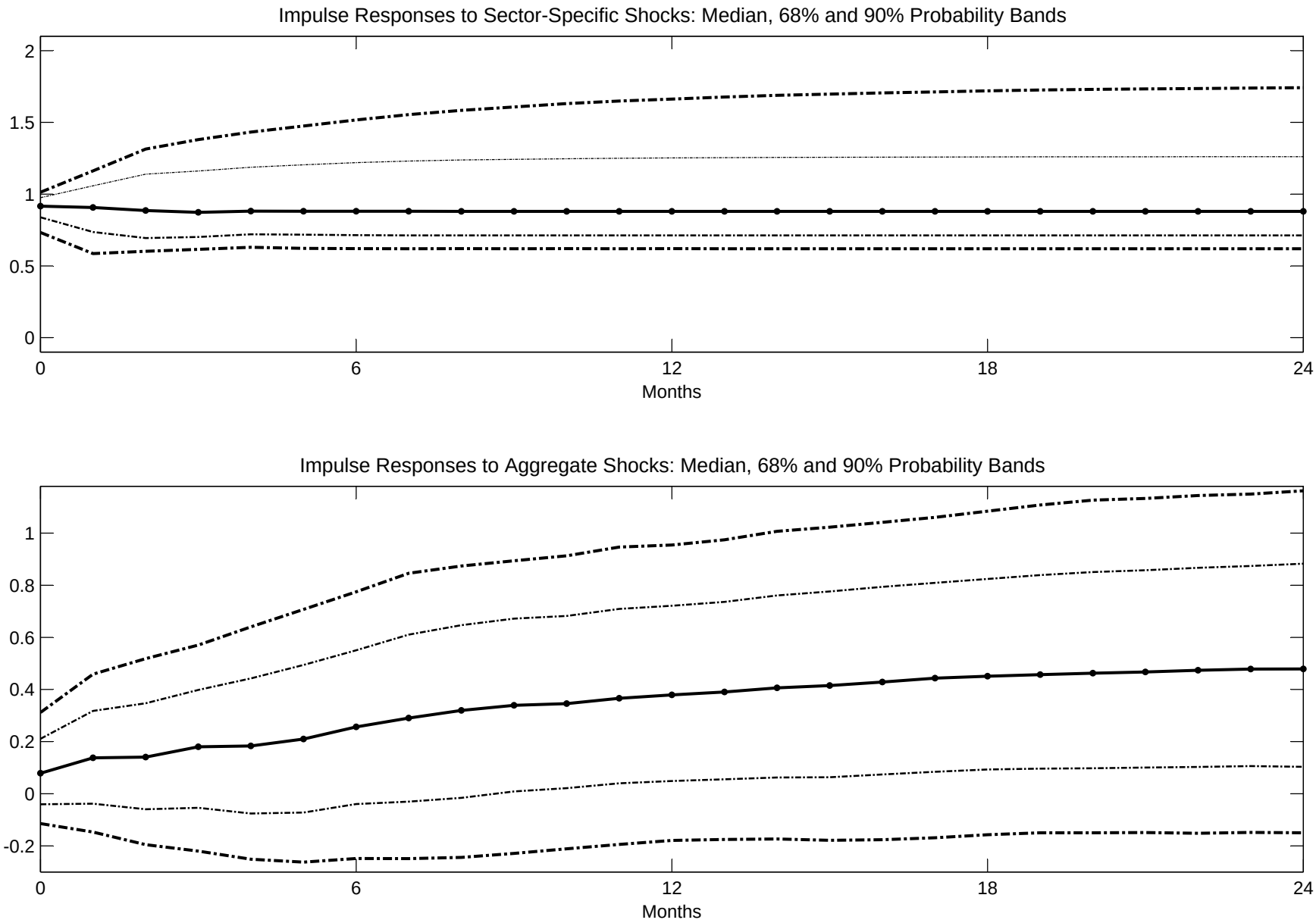
[32] Sims, C.A., 2003. Implications of rational inattention. Journal of Monetary Economics 50, 665-690.

[33] Stock, J.H., Watson, M.W., 1989. New indexes of coincident and leading economic indicators. National Bureau of Economic Research Macroeconomics Annual 4, 351-394.

[34] Stock, J.H., Watson, M.W., 2002a. Macroeconomic forecasting using diffusion indexes. Journal of Business and Economic Statistics 20, 147-162.

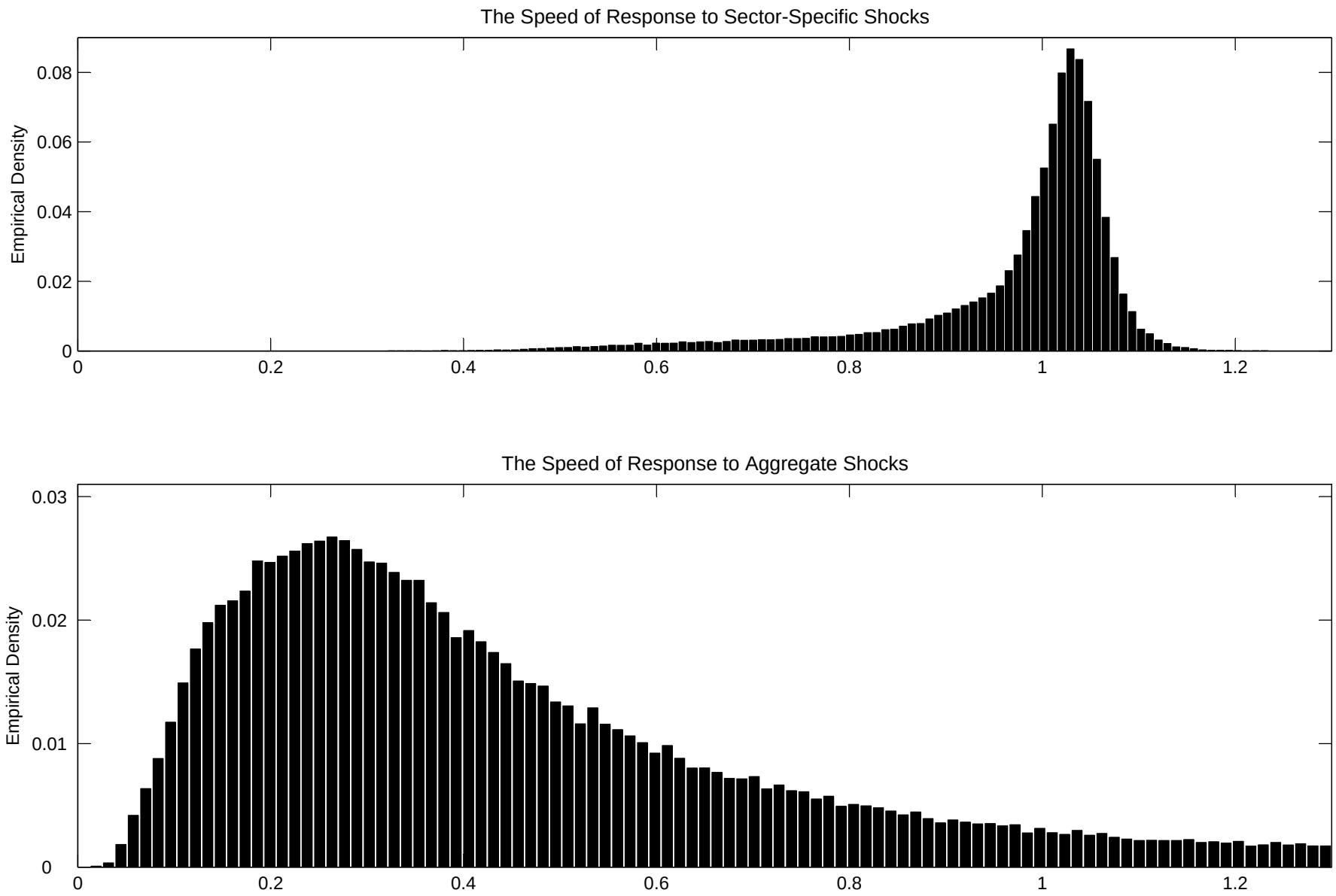
[35] Stock, J.H., Watson, M.W., 2002b. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167-1179.

Figure 1: The Cross-Section of the Normalized Impulse Responses of Sectoral Price Indexes



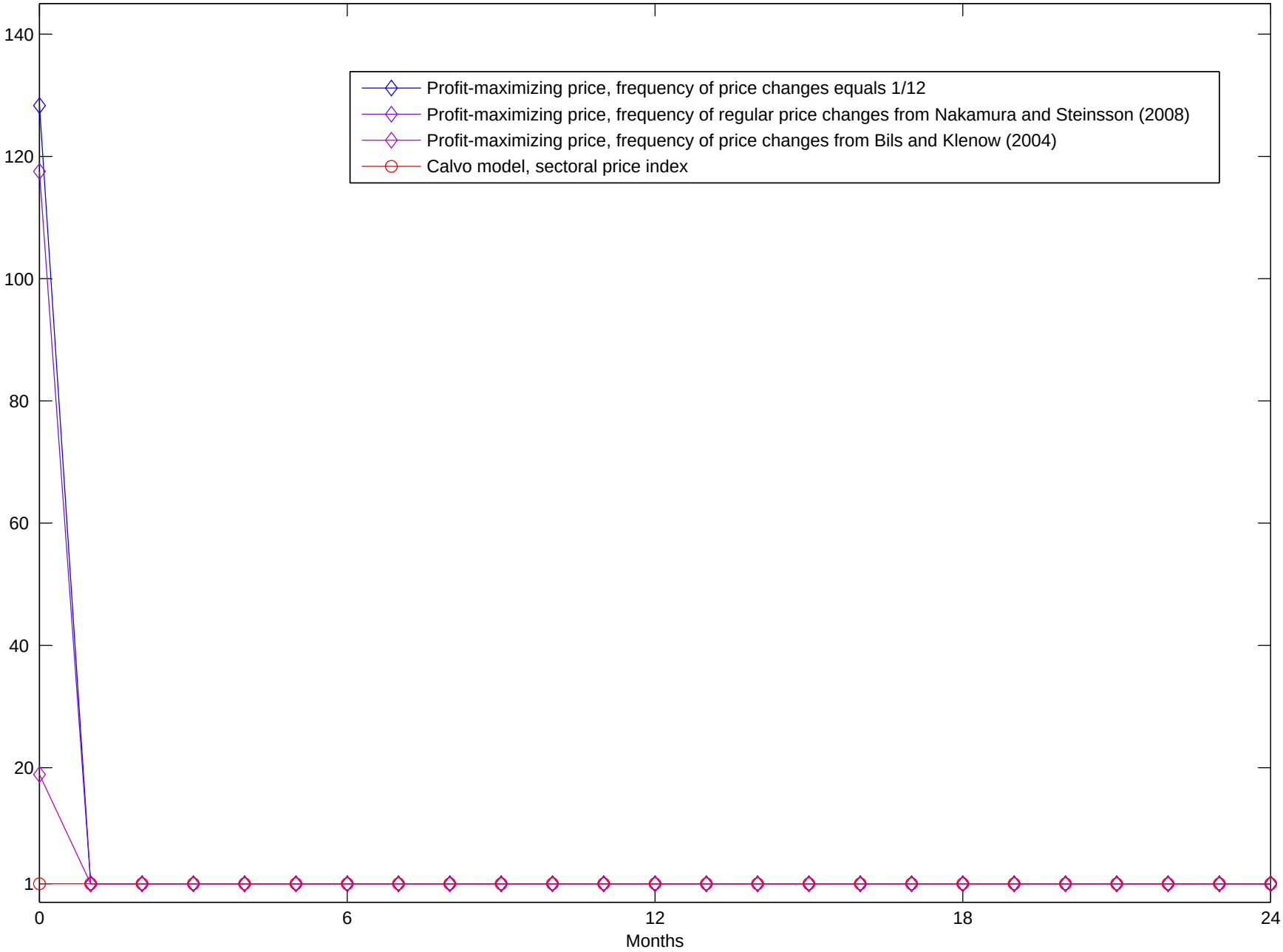
Note: Figure 1 shows the posterior density of the normalized impulse responses of sectoral price indexes to sector-specific shocks (top panel) and to aggregate shocks (bottom panel). The posterior density takes into account variation across sectors and parameter uncertainty. The results reported in Figure 1 are discussed in Section 4.

Figure 2: The Cross-Section of the Speed of Response of Sectoral Price Indexes to Shocks



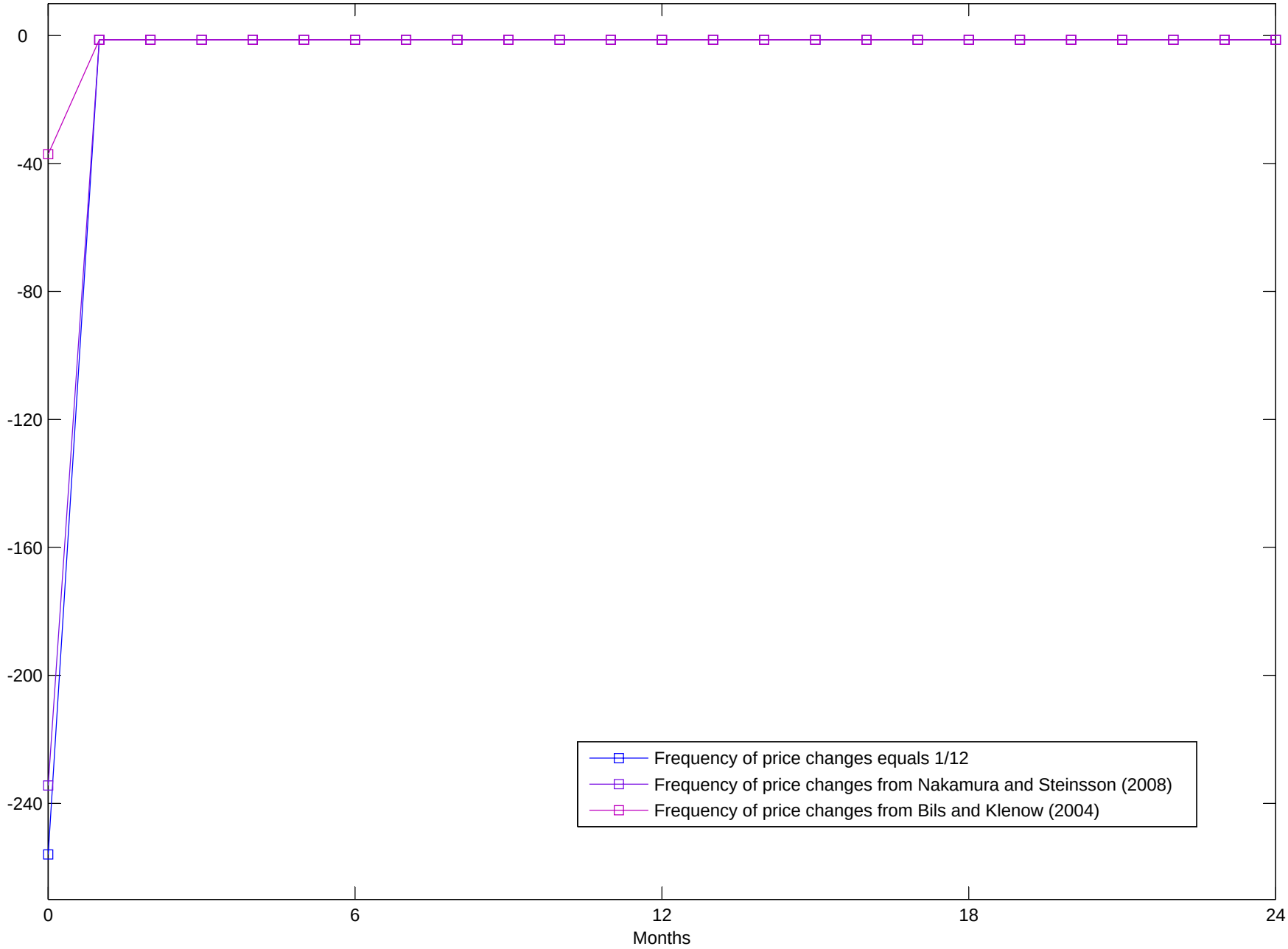
Note: Figure 2 shows the posterior density of the speed of response of sectoral price indexes to sector-specific shocks (top panel) and to aggregate shocks (bottom panel). The posterior density takes into account variation across sectors and parameter uncertainty. The speed of response is defined in Section 4.

Figure 3: Impulse Responses to Sector-Specific Shocks: Profit-Maximizing Price and Sectoral Price Index in the Calvo Model



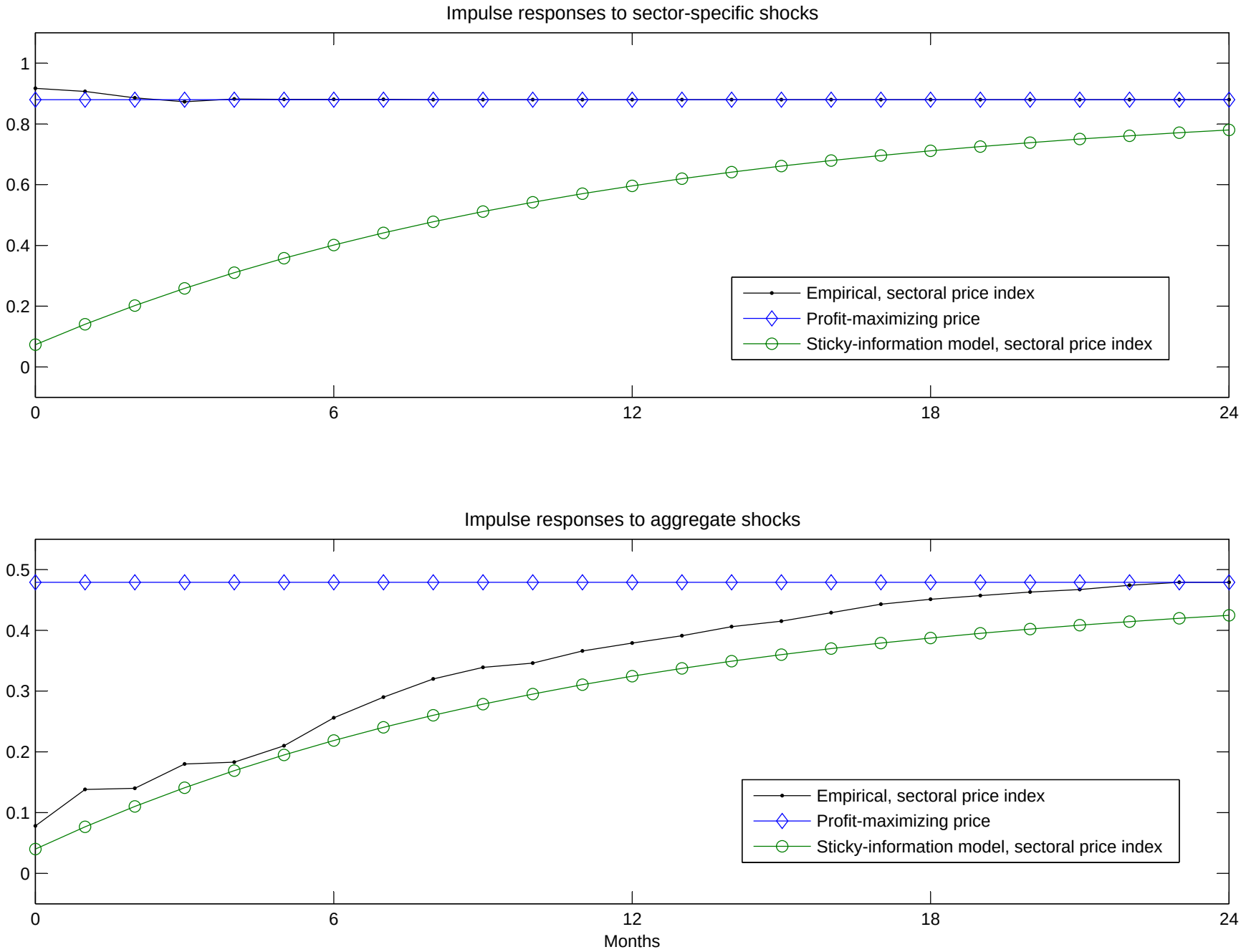
Note: Figure 3 shows the impulse response of the profit-maximizing price to a sector-specific shock (under three different assumptions concerning the frequency of price changes) and the impulse response of the sectoral price index in the Calvo model to the same shock (this impulse response is normalized to one).

Figure 4: Impulse Responses of Sector-Specific Productivity that Yield the Profit-Maximizing Impulse Responses in Figure 3



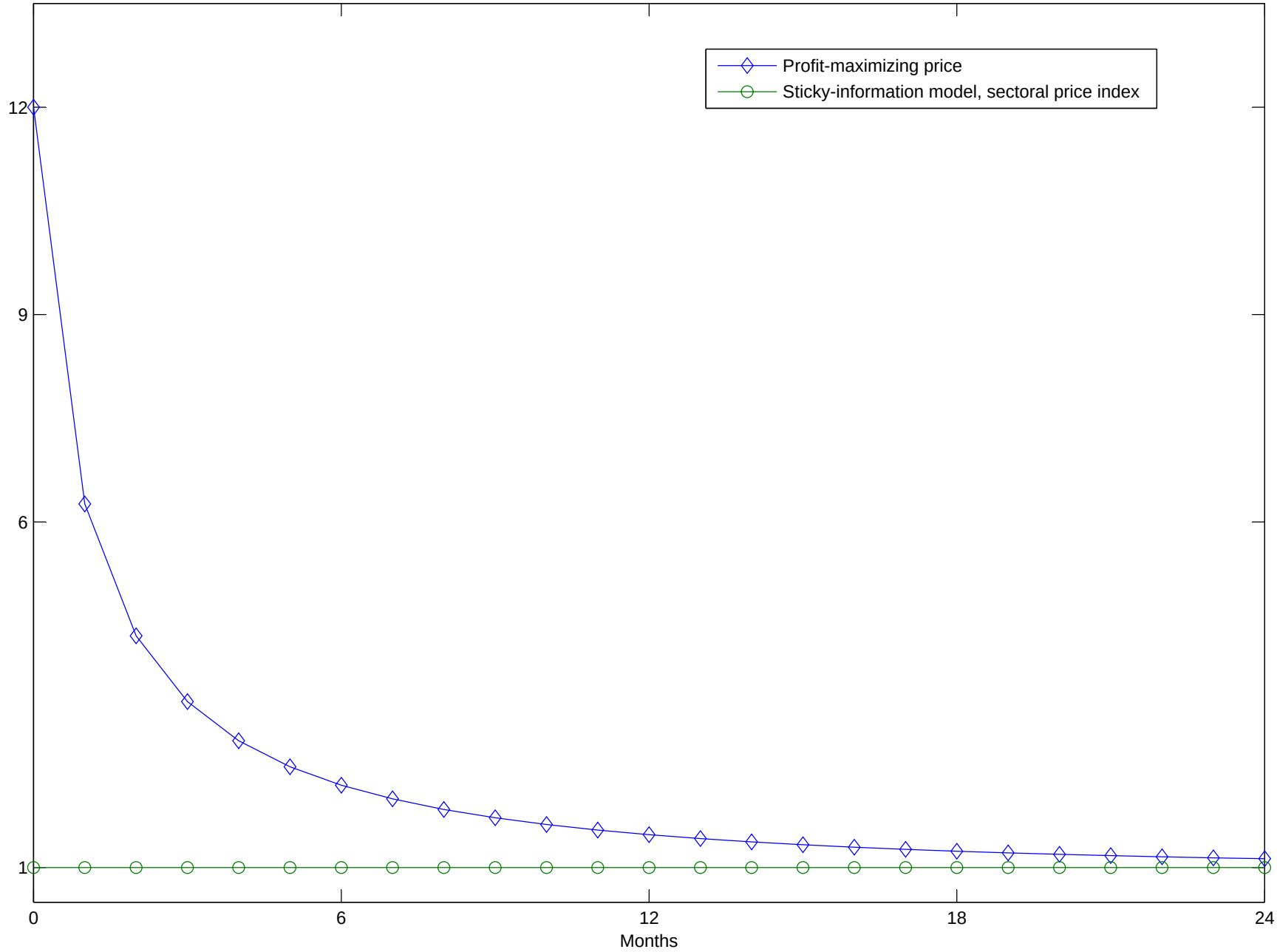
Note: Figure 4 shows the impulse responses of sector-specific productivity to an own shock (under three different assumptions concerning the frequency of price changes) that yield the impulse responses of the profit-maximizing price depicted in Figure 3.

Figure 5: Impulse Responses: Empirical, Profit-Maximizing, and Sticky-Information



Note: Figure 5 shows the impulse responses to sector-specific shocks (top panel) and to aggregate shocks (bottom panel) of three objects: (i) the sectoral price index in the data (the medians from Figure 1), (ii) the profit-maximizing price, and (iii) the sectoral price index in the sticky-information model.

Figure 6: Impulse Responses to Sector-Specific Shocks: Profit-Maximizing Price and Sectoral Price Index in the Sticky-Information Model



Note: Figure 6 shows the impulse response of the profit-maximizing price to a sector-specific shock (assuming updating of pricing plans once a year) and the impulse response of the sectoral price index in the sticky-information model to the same shock (this impulse response is normalized to one).

Table 1: The Speed of Response and the Frequency of Price Changes

Regressand	Regressor	Regressor
	Sectoral Monthly Frequency of Price Changes (Bils and Klenow, 2004)	**Sectoral Monthly Frequency of Regular Price Changes (Nakamura and Steinsson, 2008)**
Speed of Response of a Sectoral Price Index to Aggregate Shocks	**2.04** (0.29, 4.76)	
		1.55 (0.12, 4.52)
Speed of Response of a Sectoral Price Index to Sector-Specific Shocks	**0.14** (0.02, 0.28)	
		-0.09 (-0.21, 0.03)

Note: Table 1 presents the results from regressing the speed of response of a sectoral price index on the sectoral monthly frequency of price changes. Each number in bold type is the posterior median of the regression coefficient. The bracketed numbers show the 90 percent posterior probability interval for each regression coefficient. The regressions reported in Table 1 are discussed in Section 5.1. The speed of response of a sectoral price index is defined in Section 4.

Table 2: The Speed of Response and the Standard Deviation of Shocks

Regressand	Regressor	Regressor
	Standard Deviation of Sectoral Inflation due to Aggregate Shocks	**Standard Deviation of Sectoral Inflation due to Sector-Specific Shocks**
Speed of Response of a Sectoral Price Index to Aggregate Shocks	**31.42** (1.71, 85.11)	
	75.45 (-17.28, 251.50) 0.92	**-25.15** (-115.42, 26.49) 0.20
Speed of Response of a Sectoral Price Index to Sector-Specific Shocks		**1.18** (-0.04, 2.47) 0.94

Note: Table 2 presents the results from regressing the speed of response of a sectoral price index on the standard deviation(s) of sectoral inflation due to a given type of shock (given types of shocks). Each number in bold type is the posterior median of the regression coefficient. The bracketed numbers show the 90 percent posterior probability interval for each regression coefficient. When the interval includes zero, an additional number is reported. This number is the fraction of the posterior probability mass to the right of zero. The regressions reported in Table 2 are discussed in Section 5.2. The speed of response of a sectoral price index is defined in Section 4.

Table 3: The Frequency of Price Changes and the Standard Deviation of Shocks

Regressand	Regressor
	Standard Deviation of Sectoral Inflation due to Sector-Specific Shocks
Sectoral Monthly Frequency of Price Changes (Bils and Klenow, 2004)	**6.09**
	(4.62, 7.65)
Sectoral Monthly Frequency of Regular Price Changes (Nakamura and Steinsson, 2008)	**4.61**
	(2.53, 6.52)

Note: Table 3 presents the results from regressing the sectoral monthly frequency of price changes on the standard deviation of sectoral inflation due to sector-specific shocks. Each number in bold type is the posterior median of the regression coefficient. The bracketed numbers show the 90 percent posterior probability interval for each regression coefficient. The regressions reported in Table 3 are discussed in Section 5.3.

Table 4: The Speed of Response and the Standard Deviation of Shocks with Quarterly Data

Regressand	Regressor	Regressor
	Standard Deviation of Sectoral Inflation due to Aggregate Shocks	**Standard Deviation of Sectoral Inflation due to Sector-Specific Shocks**
Speed of Response of a Sectoral Price Index to Aggregate Shocks	**29.15** (6.44, 107.76)	
	61.21 (10.04, 297.30)	**-38.88** (-221.44, 8.98) 0.09
Speed of Response of a Sectoral Price Index to Sector-Specific Shocks		**2.19** (0.70, 3.66)

Note: Table 4 presents the results from reestimating, based on quarterly data, the regressions reported in Table 2. Each number in bold type is the posterior median of the regression coefficient. The bracketed numbers show the 90 percent posterior probability interval for each regression coefficient. When the interval includes zero, an additional number is reported. This number is the fraction of the posterior probability mass to the right of zero. The regressions reported in Table 4 are discussed in Section 6.2. The speed of response of a sectoral price index is defined in Section 4.

Table A.1: The Cross-Section of the Raftery-Lewis Diagnostics Across the Parameters

	Median	95th percentile	99th percentile	Max
θ's	623	695	1352	3736
φ's	697	1208	4278	478784
σ's	623	664	879	951

Note: Table A.1 reports the Raftery-Lewis diagnostics of convergence of the Gibbs sampler from the benchmark specification of the model. The Raftery-Lewis diagnostic was computed for each parameter of the model. This table summarizes the cross-section of the Raftery-Lewis diagnostics across all parameters. The parameters are grouped into the loadings (the θ's), the autoregressive parameters (the φ's), and the standard deviations (the σ's). The Raftery-Lewis diagnostics are discussed in Section A.3.

Table A.2: The Cross-Section of the Geweke Diagnostics Across the Parameters

	Median	95th percentile	99th percentile	Max
θ's	1.8	4.2	6.7	27.9
φ's	1.3	2.3	40.7	79.1
σ's	0.7	1.2	1.4	1.5

Note: Table A.2 reports the Geweke diagnostics of convergence of the Gibbs sampler from the benchmark specification of the model. The Geweke diagnostic was computed for each parameter of the model. This table summarizes the cross-section of the Geweke diagnostics across all parameters. The parameters are grouped into the loadings (the θ's), the autoregressive parameters (the φ's), and the standard deviations (the σ's). The Geweke diagnostics are discussed in Section A.3.