The Frequency of Price Adjustment and New Keynesian Business Cycle Dynamics*

Richard Dennis†
Federal Reserve Bank of San Francisco

May 2006
Preliminary. Comments welcome.

Abstract

The Calvo pricing model that lies at the heart of many New Keynesian business cycle models has been roundly criticized for being inconsistent both with time series data on inflation and with micro-data on the frequency of price changes. In this paper we show that the Galí and Gertler (1999) model, which allows for "rule-of-thumb" price setters, and whose structure can be interpreted in terms of menu costs and information gathering/processing costs, successfully resolves both criticisms. Moreover, the resulting Phillips curve shares the explanatory power of the partial-indexation model and dominates the full-indexation model and the Calvo model. Estimating a small-scale New Keynesian business cycle model, our results indicate that the share of firms that change prices each quarter is just over 60 percent, broadly in line with the Bils and Klenow (2004) study of Bureau of Labor Statistics price data. Reflecting the importance of information gathering/processing costs, we find that most firms that change prices are rule-of-thumb price setters. Finally, compared to specifications containing either the Calvo model or the full-indexation model, the data provide much greater support for the Galí-Gertler model.

Keywords: *Price adjustment, inflation indexation, Bayesian model averaging.*

JEL Classification: E52, E58.

*The views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of San Francisco or the Federal Reserve System.

†Address for Correspondence: Economic Research, Mail Stop 1130, Federal Reserve Bank of San Francisco, 101 Market St, CA 94105, USA. Email: richard.dennis@sf.frb.org.

1 Introduction

New Keynesian business cycle models have become the dominant framework for studying the design and conduct of monetary policy. The models formalize the rigidities and market imperfections that govern their behavior and are micro-founded, permitting welfare analysis and making policy experiments conducted within them immune to Lucas's (1976) critique. Prominent examples in the New Keynesian tradition include Rotemberg and Woodford (1997), Clarida, Galí, and Gertler (1999), McCallum and Nelson (1999), Walsh (2003), and Woodford (2003). One of the most important components in these models is the New Keynesian Phillips curve, the equation linking inflation to marginal costs that provides a stabilization role for monetary policy. The micro-structure that is most widely used to derive the New Keynesian Phillips curve is the Calvo model[1] (Calvo, 1983) and the defining feature of this model is that only a fixed (Calvo-) share of firms have the opportunity to optimize their price each period. This Calvo-share parameter governs the frequency with which firms change prices and determines the average duration between price changes.

Despite its popularity, the New Keynesian Phillips curve has attracted considerable criticism. Some criticisms are empirical; Estrella and Fuhrer (2002) argue that the New Keynesian Phillips curve provides a poor description of inflation dynamics because it asserts a correlation structure among inflation, the change in inflation, and marginal costs that prevents it from replicating the hump-shaped responses that are widely recognized to characterize inflation's behavior following shocks.[2] Similarly, Rudd and Whelan (2006) argue that the New Keynesian Phillips curve is incapable of describing inflation dynamics and suggest that there is little evidence of the type of forward-looking behavior required by the model. Other criticisms focus on whether estimates of the New Keynesian Phillips curve are economically plausible. In this vein, a prominent criticism is that Calvo-shares estimated from the New Keynesian Phillips curve imply a level of price rigidity that is inconsistent with micro-data on the frequency of price adjustment. For example, Sbordone (2002) estimates the Calvo-share to be around 0.8 for the United States, which implies that only 20 percent of firms change their prices each quarter and that firms change prices once every 15 months on average. But after examining Bureau of Labor Statistics data on price changes – the very price data that go into

[1] Roberts (1995) shows that Rotemberg's (1982) quadratic price adjustment costs model and Taylor's (1980) overlapping nominal wage contracts model give rise to closely related specifications, so the issues discussed in this paper apply equally to these models.

[2] In fact, the Estrella and Fuhrer (2002) criticisms apply to an entire class of rational expectations models, not just to the New Keynesian Phillips curve.

the consumer price index and the personal consumption expenditures price index – Bils and Klenow (2004) report that the average duration between price changes is just 4.3 months for the (weighted) median good in their sample. This disparity between estimates of the Calvo-share and micro-evidence on the frequency of price adjustment is worrisome, particularly since models built around the New Keynesian Phillips curve are routinely used to address issues as important as how to design a welfare-maximizing monetary policy (Erceg, Henderson, and Levin, 2000).

However, the parallels between the average duration between price changes implied by the Calvo-share and the average duration between price changes that would be estimated using a micro-dataset are less than exact. The Calvo-share describes the proportion of firms that make an optimal price change, whereas micro-data reveal whether a price has changed, but do not inform on whether the price change was optimal; some price changes may be suboptimal and it may be optimal not to change prices on some occasions. In addition, the Calvo model assumes that firms change prices once per period at most, something that is unlikely to hold true in practice, and it ignores the possibility that there may be heterogeneity in the frequency of price adjustment across firms. Each of these factors confound efforts to compare micro-data on the average duration between price changes to those implied by the Calvo-share.

Of course, these complicating factors are not a defense of the Calvo model, nor are they an argument for shielding New Keynesian business cycle models from the scrutiny of micro-data. However, they do mean that care needs to be taken when comparing macro-models to micro-data, and they call for a model of price setting that does not require a price change and an optimal price change to be the same thing. With regard to the latter, Galí and Gertler (1999) develop a model of price setting that builds on what they describe to be "rule-of-thumb" price setters. The essential feature of their model is that in each period a share of firms have the opportunity to change their prices, but they do not necessarily make an optimal price change. Instead, among those firms that change prices a fraction makes an optimal price change, while the remainder employ a rule-of-thumb pricing strategy.

Why is the Galí-Gertler price-setting environment attractive? Where traditional models of price adjustment have emphasized physical costs to changing prices, such as menu costs, as the source of price rigidity (Mankiw, 1985), recent literature has emphasized the costs that firms face when gathering (Mankiw and Reis, 2002) and processing (Sims, 2002) the information they require in order to set prices optimally. In fact, evidence suggests that costs to gathering and processing information may be much more important for price setting than traditional menu cost factors (Zbaraki, Ritson, Levy, Dutta, and Bergin, 2004).

What is attractive about the Galí-Gertler model, then, is that it provides an environment in which both costs can play a role. Menu costs – which are incurred whether or not a price change is optimal – are associated with the share of firms that can change prices. When these menu costs are large, a smaller share of firms will change their prices. Similarly, costs to gathering and processing information are associated with the share of price-changers that use rule-of-thumb pricing. When the costs to gathering and/or processing information are high, a larger share of price-changing firms will resort to a rule-of-thumb pricing strategy. Clearly, the Galí-Gertler model has a structure that allows it to be compared more readily to micro-data than the Calvo model. Moreover, since the rule-of-thumb is one in which firms index their prices to last period's inflation, the model contains a mechanism to generate intrinsic inflation persistence.

In this paper, we show how micro-data on durations between price changes can be used to construct an estimate of the discrete-time frequency of price adjustment that allows for heterogeneity and for multiple price changes per period. Applying this estimator to the Bils-Klenow data-set we estimate a discrete-time frequency of price adjustment equal to about 0.50 for quarterly data, a value considerably lower than estimates of the Calvo-share obtained from the New Keynesian Phillips curve. Next, we derive the Phillips curve associated with the Galí-Gertler model and highlight its connections to the New Keynesian Phillips curve, the full-indexation Phillips curve developed by Christiano, Eichenbaum, and Evans (2005), and the partial-indexation Phillips curve developed by Smets and Wouters (2003). We prove that these alternatives are all special cases of the Galí-Gertler Phillips curve, and argue that the Galí-Gertler model's micro-structure makes it superior to these alternatives as a consequence. Subsequently, we develop several New Keynesian business cycle models, considering specifications that interact the Galí-Gertler, the Calvo, and the full-indexation models of price setting with internal and external habit formation on the part of households. These models are deliberately kept small, focusing attention on price setting.

We estimate the models using full information maximum likelihood and Bayesian methods, and employ Bayesian predictive densities and posterior model probabilities for comparison purposes. The results are striking. First, whereas our estimate of the Calvo-share implies a mean duration between price changes that is clearly inconsistent with Bureau of Labor Statistics price data, the Galí-Gertler model does much better. In fact, our results place the share of firms that change prices each quarter at just over 60 percent, broadly in line with the Bils and Klenow findings, and a considerable improvement on the Calvo model. Second, although we find that roughly 60 percent of firms change their prices each quarter, we also find that the majority of these firms use rule-of-thumb pricing, supporting the view that information gathering/processing costs are more important for price-setting than traditional menu costs. Third, constructing predictive densities and using Bayesian model averaging, we quantify the economy's response to technology shocks, monetary policy shocks, and consumption preference shocks, revealing the counterfactual behavior of the Calvo model, establishing that the Galí-Gertler model generates hump-shaped impulse responses, and illustrating the behavioral similarities between internal and external habit formation.

We begin by describing the New Keynesian Phillips curve and illustrating the empirical disparity between the Calvo-share and the frequency of price adjustment implied by micro-data, emphasizing the study by Bils and Klenow in section 2. Section 3 outlines the economic environment that underlies the Galí-Gertler model, derives the associated Phillips curve, and compares it to the Calvo model, the full-indexation model, and the partial-indexation model. Section 3 also proves that the partial-indexation model and the Galí-Gertler model are isomorphic. Section 4 places the Galí-Gertler Phillips curve in a small-scale New Keynesian business cycle model suitable for estimation. Section 5 describes the data and discusses the estimation strategy. Section 6 presents and interprets the parameter estimates and the posterior model probabilities associated with each specification. Section 7 constructs predictive densities and uses Bayesian model averaging to summarize how consumption, inflation, and interest rates respond to shocks. Section 8 concludes.

2 The New Keynesian Phillips curve and price rigidity

As noted in the introduction, the centerpiece to much business cycle and policy analysis is the New Keynesian Phillips curve

$$\widehat{\pi}_t = \beta \mathrm{E}_t \widehat{\pi}_{t+1} + \frac{(1-\xi)(1-\beta\xi)}{\xi}\widehat{mc}_t, \tag{1}$$

where $\widehat{\pi}_t$ and $\widehat{mc}_t$ represent the percentage point deviation of inflation, π_t, and the percent deviation of real marginal costs, mc_t, around their zero-inflation nonstochastic steady state values, respectively. The economic environment that gives rise to this Phillips curve is one in which firms are monopolistically competitive, renting capital and labor and setting their prices to maximize profits subject to a constant elasticity of substitution demand curve, a Cobb-Douglas production technology, and a price rigidity, á la Calvo (1983). In equation (1),

$\beta \in (0,1)$ is the subjective discount factor and $\xi \in (0,1)$ is the Calvo-share, the share of firms that cannot optimize their prices each period.

With regard to suitable values for ξ, a touchstone in the literature is Blinder (1994), who surveyed firms on the frequency of their price changes. Based on Blinder's (1994) survey, Rotemberg and Woodford (1997) set $\xi = 0.66$, which implies an average duration between price changes of 9 months. But many calibration studies have assumed that prices change somewhat less frequently than this. For example, Erceg, Henderson, and Levin (2000) and Liu and Phaneuf (2005) each set $\xi = 0.75$, implying an average duration between price changes of 12 months.

Among studies that estimate ξ, a popular approach is to apply a GMM estimator to the moment condition[3]

$$\mathrm{E}_t\left[\left(\widehat{\pi}_t - \beta\widehat{\pi}_{t+1} - \frac{(1-\xi)(1-\beta\xi)}{\xi}\widehat{mc}_t\right)\mathbf{z}_t\right] = 0, \tag{2}$$

where $\mathbf{z}_t$ is a vector containing econometric instruments. This is the approach taken by Galí and Gertler (1999), Galí, Gertler, and López-Salido (2001), Eichenbaum and Fisher (2004), Jung and Yun (2005), and Ravenna and Walsh (2006). An alternative method is to iterate forward over equation (1) and combine the result with an evolution process for real marginal costs to produce an estimable expression relating inflation to real marginal costs (Sbordone, 2002). A range of estimates of ξ for the U.S. are displayed in Table 1.[4]

Table 1: Estimates of the New Keynesian Phillips Curve

Study	Sample	ξ
Galí & Gertler (1999)	1960:1 – 1997:4	0.829 – 0.884
Galí, Gertler & López-Salido (2001)	1970:1 – 1998:4	0.845 – 0.867
Sbordone (2002)	1960:2 – 1997:1	0.792
Eichenbaum and Fisher (2004)	1959:1 – 2001:4	0.87 – 0.91
Jung and Yun (2005)	1967:1 – 2004:4	0.910
Ravenna & Walsh (2006)	1960:1 – 2001:1	0.758 – 0.911

The estimates of ξ shown in Table 1 vary from a low of 0.758 to a high of 0.911. While $\xi = 0.758$ is broadly on par with the value used in calibration studies, a value such as $\xi = 0.911$

[3] An alternative moment condition that is often used is equation (2) multiplied through by ξ. Some of the estimates shown in Table 1 come from this alternative moment condition.

[4] All of the estimates reported in Table 1 have been made consistent with a Cobb-Douglas production technology and rental markets for capital and labor, facilitating comparison across studies by making the estimates invariant to particular assumptions about the steady state markup and labor's share of income. However, the values shown may differ from those reported in the original papers as a consequence. With respect to Sbordone's estimates, the best-fitting specification in Sbordone (2002, Table 2) has a coefficient on real marginal costs equaling $\frac{1}{18.3}$. Using Sbordone's assumption about the discount factor and assuming a rental market for capital, the implied value for ξ is 0.792.

is much larger than either the values used in calibration exercises or the value implied by Blinder's (1994) study. The average value for ξ in Table 1 is in the order of 0.85, suggesting that firms only change prices once every 20 months. The estimates in Table 1 highlight what has become an important criticism of the New Keynesian Phillips curve, which is that estimates of ξ are too large, implying mean durations between price changes that are inconsistent with micro-evidence on the frequency of price adjustment. For instance, in what is probably the most comprehensive study of micro-data to date,[5] Bils and Klenow (2004) analyze Bureau of Labor Statistics (BLS) data on goods prices and find that the (weighted) average duration between price changes is 6.6 months and that the average duration between price changes is only 4.3 months for the (weighted) median good, durations that are somewhat lower than those implied by the estimates of ξ in Table 1.

However, as stressed in the introduction, caution must be exercised when translating estimates of the Calvo-share into implied average durations between price changes. After all, the Calvo model makes no distinction between price changes and optimal price changes and may understate the share of firms that change their prices, overstating the average duration between price changes as a consequence. Moreover, the discrete-time Calvo model, the model that gives rise to equation (1), assumes that firms make at most one price change each period, potentially understating the number of price changes that occur and overstating the average duration between price changes. Finally, a direct comparison between average durations between price changes calculated using macro-models and micro-data may be misleading if there is heterogeneity in the frequency of price adjustment across firms. The first of these issues can only be addressed using a model that distinguishes between price changes and optimal price changes. The remainder of this section spells out how the remaining two issues might be addressed.

2.1 The frequency of price adjustment and implied durations

The Calvo model assumes that it is a draw from a Bernoulli distribution that determines whether or not a firm can change its prices, where the Bernoulli distribution is given by

$$p(x|\xi) = (1-\xi)^x \xi^{(1-x)}, \tag{3}$$

[5] Of course, there are other notable studies that look at micro-data on the frequency of price adjustment, including Cecchetti, (1986), Carlton (1986), and Kashyap (1995). We focus on the Bils and Klenow (2004) study because of its comprehensive nature. The study by Carlton (1986) looks at producer prices rather than consumer prices.

with $\xi \in (0,1)$. A firm that draws $x = 0$ cannot change its price, while a firm that draws $x = 1$ can. Equation (3) can be thought of as the discrete-time arrival process for the Calvo-signal, the signal that firms receive indicating whether they can change their price or not. According to the Bernoulli distribution, the discrete-time frequency of a price changing during the period, which is equivalent to the share of firms that can optimize their prices, is $p(x = 1|\xi) = 1 - \xi$. However, if firms that can change prices actually make decisions continuously, then they may make multiple price changes during any discrete time period. In a continuous-time setting, the arrival process for the Calvo-signal more naturally follows a Poisson distribution,

$$p(x|\mu) = \frac{(1-\mu)^x e^{-(1-\mu)}}{x!}, \tag{4}$$

with $\mu \in (0,1)$, where $x \in \{1, 2, 3, ...\}$ indicates the number of times a given firm can change prices.

Proposition 1: If the arrival process is Poisson, but it is modeled as Bernoulli, and if the Calvo-share satisfies $\xi \in \left(e^{-1}, 1\right)$, then (i) the continuous-time frequency of price adjustment, $1 - \mu$, and the Calvo-share are related according to $1 - \mu = -\ln(\xi)$, and (ii) the average and median durations between price changes are overstated.

Proof: From the Bernoulli distribution, the probability that a firm's price will change during a period is $1 - \xi$. From the Poisson distribution, the probability that a firm's price will change one or more times during a period is $1 - e^{-(1-\mu)}$. It follows that

$$1 - \mu = -\ln(\xi), \tag{5}$$

where $\mu \in (0,1)$ requires $\xi \in \left(e^{-1}, 1\right)$. Now, for the Bernoulli arrival process the average, $\overline{d}_B$, and median, $\underline{d}_B$, durations between price changes are given by

$$\overline{d}_B = \sum_{i=1}^{\infty} i(1-\xi)\xi^{(i-1)} = \frac{1}{1-\xi}, \tag{6}$$

$$\underline{d}_B = max\left\{\frac{\ln\left(\frac{1}{2}\right)}{\ln(\xi)}, 1\right\}, \tag{7}$$

whereas for the Poisson arrival process the average, $\overline{d}_P$, and median, $\underline{d}_P$, durations between price changes are given by

$$\overline{d}_P = \int_0^{\infty} h(1-\mu)e^{-(1-\mu)h}dh = \frac{1}{1-\mu}, \tag{8}$$

$$\underline{d}_P = -\frac{\ln\left(\frac{1}{2}\right)}{1-\mu}. \tag{9}$$

Substituting equation (5) into equation (7) establishes that $\underline{d}_B \geq \underline{d}_P$. Finally, equation (5) can be written as $\mu = 1 + \ln(\xi)$, which, by inspection, implies that $\mu \leq \xi$ for all $\xi \in (e^{-1}, 1)$. Therefore, $\overline{d}_B \geq \overline{d}_P$, which completes the proof. ■

Proposition 1 has two main implications. First, it establishes that the standard mapping from the Calvo-share to the average duration between price changes will generally overstate the average duration between price changes because it does not allow for the possibility that firms may make multiple price changes during a period. However, if one is prepared to model the arrival process for the Calvo-signal with a Poisson distribution, then Proposition 1 shows how to translate the Calvo-share into a continuous-time frequency of price adjustment, from which the implied average duration between price changes can be more accurately calculated. Second, it shows that a more robust comparison between micro-data and the Calvo-share is afforded by using median durations between price changes rather than average durations between price changes.

2.2 Heterogeneity and implied durations

Now consider the possibility that there is heterogeneity in the frequency of price adjustment across firms.[6] To allow for such heterogeneity, assume that the continuous-time frequency of price adjustment, $1 - \mu$, is distributed across firms according to a Beta density, a density known for its flexibility and generality. With the Beta density given by

$$p\left[(1-\mu)\,|a,b\right] = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}\,(1-\mu)^{(a-1)}\,\mu^{(b-1)} \tag{10}$$

we have

Proposition 2: For the class of Beta densities that satisfies $a > 1$ and $b > 0$, the implied average duration between price changes evaluated at the average continuous-time frequency of price adjustment understates the average duration between price changes.

Proof: The average of the Beta density equals $\frac{a}{a+b}$, which implies that the average duration between price changes evaluated at the average continuous-time frequency of price adjustment equals $\frac{a+b}{a}$.

[6] We consider firms that produce a single good. As a consequence, heterogeneity in the frequency of price adjustment across firms is equivalent to heterogeneity in the frequency of price adjustment across goods.

Next, the average duration between price changes, $\overline{d}$, is given by

$$\begin{aligned}
\overline{d} &= \int_0^1 \frac{1}{1-\mu} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} (1-\mu)^{(a-1)} \mu^{(b-1)} d\mu \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 (1-\mu)^{(a-2)} \mu^{(b-1)} d\mu \\
&= \frac{\Gamma(a+b)}{\Gamma(a)} \frac{\Gamma(a-1)}{\Gamma(a-1+b)} \int_0^1 \frac{\Gamma(a-1+b)}{\Gamma(a-1)\Gamma(b)} (1-\mu)^{(a-2)} \mu^{(b-1)} d\mu \\
&= \frac{\Gamma(a-1)}{\Gamma(a)} \frac{\Gamma(a+b)}{\Gamma(a-1+b)} \\
&= \frac{a-1+b}{a-1}.
\end{aligned}$$

Clearly, $\frac{b}{a-1} \geq \frac{b}{a}$ for all $a > 1, b > 0$, which completes the proof. ■

Proposition 2 shows that ignoring heterogeneity in the continuous-time frequency of price adjustment across firms causes the average duration between price changes to be understated. However, to the extent that a Beta distribution can usefully approximate how the continuous-time frequency of price adjustment is distributed across firms, and to the extent that the arrival process for the Calvo-signal is better described by a Poisson distribution than a Bernoulli distribution, Propositions 1 and 2 enable one to estimate the discrete-time frequency of price adjustment, $1-\xi$, from micro-data on durations between price changes.

Bils and Klenow (2004) report the durations between price changes, along with their weights, for 350 goods in the consumers' price index. The (weighted) average of these durations is just under 6.6 months and the (weighted) standard deviation of these durations is just over 7.1 months. From the Beta distribution, these two moments are given by

$$\overline{d} = \frac{a-1+b}{a-1}, \tag{11}$$

$$var(d) = \frac{(a+b-1)}{(a-1)} \left[\frac{(a+b-2)}{(a-2)} - \frac{(a+b-1)}{(a-1)} \right]. \tag{12}$$

Using equations (11) and (12), $a = 2.724$ and $b = 9.601$ match the average duration and the standard deviation of durations in the Bils-Klenow data-set. From these estimates of a and b, the average continuous-time frequency of price adjustment is 0.221, at a monthly rate. Employing the transformation provided by Proposition 1, the continuous-time frequency of price adjustment translates into a discrete-time frequency of price adjustment of 0.198, again at a monthly rate. How does this estimate compare to the Bils and Klenow direct estimate of the frequency of price adjustment? In the Bils-Klenow data-set, the (weighted) median frequency of price adjustment is 0.209 and the (weighted) average frequency of price adjustment is 0.226.[7] Thus, the Beta-distribution-based estimate is consistent with the direct estimates available from the Bils-Klenow data-set, which suggests that the approach is also likely to be useful in contexts where only durations between price changes are available.

Now, if the monthly discrete-time frequency of price adjustment is 0.198 (0.226), then the quarterly discrete-time frequency of price adjustment is approximately 0.485 (0.536), which suggests that a value for ξ of around 0.5 is appropriate for quarterly data. Clearly, the estimates in Table 1 place ξ much higher than 0.5, and it is on the basis of this that we conclude that even after correcting for heterogeneity, and after allowing for the possibility that firms may make multiple price changes per period, the New Keynesian Phillips curve remains at odds with micro-data on the frequency of price adjustment.

3 The Galí-Gertler model

We now turn to a model that distinguishes between price changes and optimal price changes, the Galí and Gertler (1999) model. The economy is populated by a continuum of monopolistically competitive firms, normalized to the unit interval, each of which produces a differentiated product according to the Cobb-Douglas production technology: $y_t(i) = [e^{u_t} l_t(i)]^{\kappa} k_t(i)^{1-\kappa}$, $\kappa \in (0,1)$, where e^{u_t} is an aggregate labor-augmenting technology shock. Indexing firms by i, the output of the i'th firm and their labor and capital inputs are denoted $y_t(i)$, $l_t(i)$, and $k_t(i)$, respectively. It is assumed that the capital stock is owned by households and rented to firms in a perfectly competitive market, evolving over time according to $K_t = (1-\delta)K_{t-1} + I_t$, where I_t denotes aggregate investment and $\delta \in (0,1)$ is the depreciation rate. The final good, Y_t, is bought and sold in a perfectly competitive market and produced from the outputs of the individual firms according to the constant-returns-to-scale production technology $Y_t = \left[\int_0^1 y_t(i)^{\frac{\epsilon-1}{\epsilon}} di\right]^{\frac{\epsilon}{\epsilon-1}}$, $\epsilon \in (1,\infty)$. As is well-known, the demand schedule for the i'th firm's good, $y_t(i)$, takes the form $y_t(i) = Y_t \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon}$, where $P_t(i)$ is the price charged by the i'th firm and P_t is the aggregate price level.

Each period, a fixed proportion of firms, $1-\theta$, $\theta \in [0,1)$, are able to change prices. However, not all firms that change prices do so optimally. Within the share of firms that can change prices, a fixed proportion, $1-\omega$, $\omega \in [0,1)$, change their prices optimally, while the remaining proportion, ω, set their prices according to a simple rule of thumb. Unlike the Calvo

[7] To construct this estimate we have excluded six goods whose frequency of price adjustment did not satisfy $\xi \in (e^{-1}, 1)$. Five of these six goods fall into the food and energy category and would be excluded from core inflation; the sixth is airline fares. For these six goods, their frequency of price adjustment cannot be mapped into a continuous-time probability of changing prices.

model, where firms either set their price optimally or keep their price unchanged, here firms either set their prices optimally, change their prices using a rule of thumb, or keep their prices unchanged. Each period a measure equaling θ of firms do not change their prices, a measure equaling $\omega(1-\theta)$ of firms change their prices by rule of thumb, and a measure equaling $(1-\omega)(1-\theta)$ of firms set their prices to maximize profits, with firms falling randomly into one of these three categories, independently of their history of price changes.

The model has two key parameters, and each can be associated with a distinct cost impinging on a firm's pricing decision. The first set of costs, menu costs, are borne by firms when they changes prices, regardless of whether the price change is optimal or not; these costs are associated with θ. The second set of costs are those connected to the information gathering (Mankiw and Reis, 2002) and information processing (Sims, 2002) needed to determine the optimal price; these costs are associated with ω. Importantly, while obstensibly playing a role similar to ξ, θ represents a cost to changing prices, not a cost to making an optimal price change.

Unlike Galí and Gertler (1999), who assume a rule-of-thumb in which prices are set according to

$$P_t(i) = (1+\pi_{t-1}) P_{t-1}, \tag{13}$$

we introduce indexation and assume instead that the rule-of-thumb is

$$P_t(i) = (1+\pi_{t-1}) P_{t-1}(i), \tag{14}$$

a rule-of-thumb in which firms index their own prices to last period's aggregate inflation rate.[8] With this indexation-based rule-of-thumb, aggregate prices equal

$$\begin{aligned} P_t &\equiv \left[\int_0^1 P_t(i)^{1-\epsilon} di\right]^{\frac{1}{1-\epsilon}} \\ &= \left[(1-\theta)(1-\omega) P_t^{*1-\epsilon} + (1-\theta)\omega(1+\pi_{t-1})^{1-\epsilon} P_{t-1}^{1-\epsilon} + \theta P_{t-1}^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}, \end{aligned} \tag{15}$$

where P_t^* is the price chosen by the optimizing firms. Log-linearizing equation (15) about a zero-inflation nonstochastic steady state, and restricting θ and ω to satisfy $\omega + \theta > 0$, it follows that the quasi-difference in aggregate inflation is related to the optimal relative price according to

$$\widehat{\pi}_t = \frac{\omega(1-\theta)}{\theta+\omega(1-\theta)}\widehat{\pi}_{t-1} + \frac{(1-\omega)(1-\theta)}{\theta+\omega(1-\theta)}\widehat{p}_t^*. \tag{16}$$

[8] Eichenbaum and Fisher (2004) also analyze a version of the Galí-Gertler model in which firms index their own price changes to lagged inflation. Despite this similarity, their specification differs from ours with respect to the optimization problem solved by the price-setting firms, and gives rise to a quite different Phillips curve.

Now, in period $t+1$ a firm that cannot optimize its price between period t and period $t+1$ will expect to charge the price[9]

$$\begin{aligned} P_{t+1}(i) &= P_t(i)\left[\frac{\omega(1-\theta)}{\theta+\omega(1-\theta)}(1+\pi_t)+\frac{\theta}{\theta+\omega(1-\theta)}\right] \\ &\equiv P_t(i)\,S_{t+1}, \end{aligned}$$

from which it follows that the price the firm will expect to change in period $t+j$ is

$$P_{t+j}(i) = P_t(i)\prod\nolimits_{k=1}^{j} S_{t+k}. \tag{17}$$

Therefore, with $\mu \equiv \theta+\omega(1-\theta)$, representing the share of firms that cannot optimize their prices, those firms that can optimize will set $P_t(i)$ to maximize

$$F_t(i) = \mathrm{E}_t\sum_{j=0}^{\infty}(\beta\mu)^j\,Y_{t+j}\left[\left(\frac{P_t(i)\prod_{k=1}^{j}S_{t+k}}{P_{t+j}}\right)^{1-\epsilon} - mc_{t+j}(i)\left(\frac{P_t(i)\prod_{k=1}^{j}S_{t+k}}{P_{t+j}}\right)^{-\epsilon}\right]. \tag{18}$$

The resulting first-order condition can be written as

$$\mathrm{E}_t\sum_{j=0}^{\infty}(\beta\mu)^j\,y_{t+j}(i)\left[\frac{p_t^*(i)\left(\prod_{k=1}^{j}S_{t+k}\right)}{\prod_{k=1}^{j}(1+\pi_{t+k})} - \frac{\epsilon}{(\epsilon-1)}mc_{t+j}(i)\right] = 0. \tag{19}$$

Log-linearizing equation (19) and assuming symmetry yields

$$\widehat{p}_t^* = \beta\mu\mathrm{E}_t\widehat{p}_{t+1}^* + \beta\mu\left(\mathrm{E}_t\widehat{\pi}_{t+1} - \frac{\omega(1-\theta)}{\mu}\widehat{\pi}_t\right) + (1-\beta\mu)\,\widehat{mc}_t. \tag{20}$$

Combining equations (16) and (20), the resulting Phillips curve is

$$\begin{aligned} \widehat{\pi}_t &= \frac{\omega(1-\theta)}{\theta+\omega(1-\theta)(1+\beta)}\widehat{\pi}_{t-1} + \frac{\beta[\theta+\omega(1-\theta)]}{\theta+\omega(1-\theta)(1+\beta)}\mathrm{E}_t\widehat{\pi}_{t+1} \\ &\quad + \frac{(1-\omega)(1-\theta)(1-\beta\mu)}{\theta+\omega(1-\theta)(1+\beta)}\widehat{mc}_t. \end{aligned} \tag{21}$$

When $\omega = 0$ (and $\theta \neq 0$), i.e., when no firms index, the backward dynamics are zeroed out and equation (21) collapses to the New Keynesian Phillips curve, equation (1), but with θ representing ξ. Similarly, when $\theta = 0$ (and $\omega \neq 0$), i.e., when all firms change prices, equation (21) simplifies to

$$\pi_t = \frac{1}{1+\beta}\pi_{t-1} + \frac{\beta}{1+\beta}\mathrm{E}_t\pi_{t+1} + \frac{(1-\omega)(1-\beta\omega)}{(1+\beta)\,\omega}\widehat{mc}_t, \tag{22}$$

[9] This implementation of the Galí and Gertler (1999) model differs slightly, but importantly, from the original. We assume that the optimizing firms take into account that there is a non-zero probability that they will be indexing in subsequent periods in which they cannot not reoptimize; Galí and Gertler (1999) assume that optimizing firms behave like Calvo-pricing firms, assigning zero probability to indexation occuring in subsequent periods.

which is the Christiano, Eichenbaum, and Evans (2005) full-indexation model, but with ω playing the role of ξ. Clearly, the Galí-Gertler Phillips curve encompasses the Calvo Phillips curve and the full-indexation Phillips curve, but how does it relate to the partial-indexation Phillips curve employed by Smets and Wouters (2003)?

3.1 The partial-indexation model

Like the Galí-Gertler model, the partial-indexation model is related to the Calvo model, with a modification to the behavior of the firms that cannot optimize their prices. Where the Calvo model assumes that the share of non-optimizing firms keep their prices unchanged, the partial-indexation model assumes that they index their price changes as a proportion of last period's inflation rate. In this respect, the partial-indexation model can be viewed as a model in which it is costless to change prices, but costly to optimize the price change. But, although the partial-indexation model addresses the concerns in Estrella and Fuhrer (2002), because it assumes that all prices change every period, either optimally or through indexation, it is even more at odds with micro-data, and less economically plausible, than the Calvo model.

With $\eta \in [0,1]$ representing the indexation parameter, a firm that is unable to optimize its price between periods t and $t+j$ will sell its good in period $t+j$ at the price $P_{t+j}(i) = P_t(i) \prod_{k=1}^{j} (1 + \eta \pi_{t+k-1})$. Aside from the special situation where $\eta = 0$, in which case the Calvo model is restored, the inflation indexation augments the Calvo model by allowing lagged inflation to affect firms' pricing decisions.

Let $S_{t+j} \equiv \prod_{k=1}^{j} (1 + \eta \pi_{t+k-1})$, then the profit function for the optimizing firms becomes

$$F_t(i) = \mathrm{E}_t \sum_{j=0}^{\infty} (\beta\xi)^j Y_{t+j} \left[\left(\frac{S_{t+j} P_t(i)}{P_{t+j}} \right)^{1-\epsilon} - mc_{t+j}(i) \left(\frac{S_{t+j} P_t(i)}{P_{t+j}} \right)^{-\epsilon} \right], \tag{23}$$

for which the log-linear first-order condition for the optimal relative price is

$$\widehat{p}_t^* = (1 - \beta\xi) \mathrm{E}_t \sum_{j=0}^{\infty} (\beta\xi)^j \left[\widehat{mc}_{t+j} + \sum_{k=1}^{j} (\widehat{\pi}_{t+k} - \eta \widehat{\pi}_{t+k-1}) \right]. \tag{24}$$

Now, log-linearizing the Dixit-Stiglitz aggregator gives

$$\widehat{\pi}_t = \eta \widehat{\pi}_{t-1} + \frac{1-\xi}{\xi} \widehat{p}_t^*, \tag{25}$$

and combining equations (24) and (25) we arrive at

$$\widehat{\pi}_t = \frac{\eta}{1+\eta\beta} \widehat{\pi}_{t-1} + \frac{\beta}{1+\eta\beta} \mathrm{E}_t \widehat{\pi}_{t+1} + \frac{(1-\beta\xi)(1-\xi)}{(1+\eta\beta)\xi} \widehat{mc}_t, \tag{26}$$

in accordance with Smets and Wouters (2003). Notice that when $\eta = 0$, equation (26) collapses to equation (1) and that when $\eta = 1$, it is equivalent to the full-indexation Phillips curve, equation (22).

3.2 Model equivalence

Because the economic environments in which they are derived are similar, it is natural to ask whether there might be mathematical connections between the Galí-Gertler model and the partial-indexation model. For example, we showed earlier that by setting $\theta = 0$, the Galí-Gertler model is equivalent to the full-indexation model, which is a special case of the partial-indexation model in which $\eta = 1$, but with ω in the Galí-Gertler model playing the role of ξ.

Proposition 3: To a first-order log-linear approximation about a nonstochastic steady state with zero inflation, the partial-indexation Phillips curve and the Galí-Gertler Phillips curve are isomorphic.

Proof: Define $\eta \equiv \frac{\omega(1-\theta)}{\theta+\omega(1-\theta)}$ and $\xi \equiv \mu = \theta + \omega(1-\theta)$, then the partial-indexation Phillips curve can be written as

$$\begin{aligned} \widehat{\pi}_t &= \frac{\frac{\omega(1-\theta)}{\theta+\omega(1-\theta)}}{1+\beta\left(\frac{\omega(1-\theta)}{\theta+\omega(1-\theta)}\right)}\widehat{\pi}_{t-1} + \frac{\beta}{1+\beta\left(\frac{\omega(1-\theta)}{\theta+\omega(1-\theta)}\right)}\mathrm{E}_t\widehat{\pi}_{t+1} \\ &\quad + \frac{(1-\omega)(1-\theta)(1-\beta\mu)}{\left[1+\beta\left(\frac{\omega(1-\theta)}{\theta+\omega(1-\theta)}\right)\right][\theta+\omega(1-\theta)]}\widehat{mc}_t. \end{aligned} \tag{27}$$

After some simple cancellations equation (27) becomes

$$\begin{aligned} \widehat{\pi}_t &= \frac{\omega(1-\theta)}{\theta+\omega(1-\theta)(1+\beta)}\widehat{\pi}_{t-1} + \frac{\beta}{\theta+\omega(1-\theta)(1+\beta)}\mathrm{E}_t\widehat{\pi}_{t+1} \\ &\quad + \frac{(1-\omega)(1-\theta)(1-\beta\mu)}{\theta+\omega(1-\theta)(1+\beta)}\widehat{mc}_t, \end{aligned}$$

which has the same structure as the Galí-Gertler Phillips curve. Now, by inspection, for all $\omega \in [0,1)$ and $\theta \in [0,1)$ that satisfy $\omega + \theta > 0$, then $\eta \in [0,1]$ and $\xi \in (0,1)$, which establishes that the Galí-Gertler Phillips curve is a special case of the partial-indexation Phillips curve. Conversely, define $\theta \equiv \xi(1-\eta)$ and $\omega \equiv \frac{\xi\eta}{1-\xi(1-\eta)}$, which imply $\mu = \xi$, then the Galí-Gertler Phillips curve can be written as

$$\begin{aligned} \widehat{\pi}_t &= \frac{\xi\eta}{\xi(1-\eta)+\xi\eta(1+\beta)}\widehat{\pi}_{t-1} + \frac{\beta[\xi(1-\eta)+\xi\eta]}{\xi(1-\eta)+\xi\eta(1+\beta)}\mathrm{E}_t\widehat{\pi}_{t+1} \\ &\quad + \frac{(1-\xi)(1-\beta\xi)}{\xi(1-\eta)+\xi\eta(1+\beta)}\widehat{mc}_t, \end{aligned}$$

which in turn simplifies to

$$\widehat{\pi}_t = \frac{\eta}{1+\eta\beta}\widehat{\pi}_{t-1} + \frac{\beta}{1+\eta\beta}\mathrm{E}_t\widehat{\pi}_{t+1} + \frac{(1-\xi)(1-\beta\xi)}{(1+\eta\beta)\xi}\widehat{mc}_t. \tag{28}$$

Equation (28) has the same structure as the partial-indexation Phillips curve. With respect to the parameter spaces, again by inspection, for all $\eta \in [0,1]$ and $\xi \in (0,1)$ then $\omega \in [0,1)$ and $\theta \in [0,1)$ and $\theta + \omega > 0$, which establishes that the partial-indexation Phillips curve model is a special case of the Galí-Gertler Phillips curve. Since each specification is a special case of the other they must be isomorphic. ■

Proposition 3 establishes that the Galí-Gertler Phillips curve and the partial-indexation Phillips curve are mathematically equivalent, and this equivalence also has a strong intuition. The parameter η in the partial-indexation model has as its counterpart the convolution $\frac{\omega(1-\theta)}{\theta+\omega(1-\theta)}$ in the Galí-Gertler model. To appreciate why these two parameters play the same role, observe that the numerator of $\frac{\omega(1-\theta)}{\theta+\omega(1-\theta)}$ is the share of firms that index to lagged inflation and the denominator is the share of firms that are either indexing to lagged inflation or indexing to a zero inflation rate. In terms of the contribution to inflation being made by the non-optimizing firms, the convolution $\frac{\omega(1-\theta)}{\theta+\omega(1-\theta)}$ can be thought of as the weight on lagged inflation in a weighted average of lagged inflation and zero inflation, which is naturally equivalent to the weight on lagged inflation in a model with partial-indexation. Similarly, it should be clear that the term $\frac{(1-\omega)(1-\theta)}{\theta+\omega(1-\theta)}$ in equation (16) plays the same role as $\frac{(1-\xi)}{\xi}$ in equation (25) and that these two expressions are equal when $\xi = \theta + \omega(1-\theta)$, which is intuitive because ξ is the share of firms that do not optimize in the partial-indexation model and $\theta + \omega(1-\theta)$ is the share of firms that do not optimize in the Galí-Gertler model.

4 A New Keynesian business cycle model

The previous section showed that the Galí-Gertler Phillips curve can readily be compared to micro-estimates of the frequency of price adjustment and, further, that it encompasses the Calvo model, the full-indexation model, and the partial-indexation model. The latter result suggests that the Galí-Gertler Phillips curve has important advantages over these alternative specifications for macroeconometric analysis; the former result raises the question of whether the Galí-Gertler Phillips curve is in line with micro-evidence on the frequency of price adjustment? To investigate this question, in this section we develop a general equilibrium New Keynesian business cycle model suitable for estimation.

The model consists of three types of agent: households, firms, and a central bank. Firm behavior is described by the Galí-Gertler model, above. In the remainder of the section we outline the decision problems and the behavior of households and the central bank.

4.1 Households

Households choose consumption, c_t, investment, I_t their supply of labor, l_t, and their holdings of nominal money balances, m_t, and bonds, b_t, to maximize

$$\mathrm{E}_t \sum_{j=0}^{\infty} \beta^j \left[\frac{e^{g_t} \left(c_{t+j} - H_{t+j}\right)^{1-\sigma}}{1-\sigma} + \frac{\left(\frac{m_{t+j}}{P_{t+j}}\right)^{1-\alpha}}{1-\alpha} - \frac{l_{t+j}^{1+\chi}}{1+\chi} \right], \tag{29}$$

where $\{\sigma, \alpha, \chi\} \in (0, \infty)$, and g_t, $g_t \sim iid\left[0, \sigma_g^2\right]$, is an aggregate consumption-preference shock, subject to the budget constraint

$$c_t + \frac{m_t}{P_t} + \frac{b_t}{P_t} + I_t = w_t l_t + r_t K_t + \frac{(1+R_{t-1})}{P_t} b_{t-1} + \frac{m_{t-1}}{P_t} + \frac{\Pi_t}{P_t} \tag{30}$$

and the capital accumulation equation

$$K_{t+1} = (1-\delta) K_t + I_t, \tag{31}$$

where R_t denotes the nominal interest rate, w_t denotes the consumption real wage, r_t denotes the real rental payment on capital, Π_t denotes the lump-sum profits households earn from dividend payments from firms and the seigniorage revenues households receive from the government, and $K_t = \int_0^1 k_t(i)\, di$. Equation (29) allows for habit formation, positing that what matters for households is their consumption in relation to a habit stock, H_t. This habit stock is assumed to evolve according to

$$H_t = \gamma c_{t-1}^{\mathrm{D}} C_{t-1}^{1-\mathrm{D}}, \tag{32}$$

where $\gamma \in [0,1)$, $\mathrm{D} \in \{0,1\}$, and C_t represents aggregate, as opposed to household-level, consumption. The parameter D distinguishes between internal and external habits; when D = 1 the habit formation is internal and when D = 0 the habit formation is external (Abel, 1990). Since household consumption must always remain above the habit stock $(c_t - H_t) > 0$, additive habits are closely related to the notion that there is a subsistence level below which a household's consumption cannot fall. Carroll, Overland, and Weil (2000) and Boldrin, Christiano, and Fisher (2001) have shown that habit formation is important for explaining savings behavior and asset returns over the business cycle, respectively.

The first-order conditions for the Lagrangian, Λ, associated with the household's problem, include

$$\frac{\partial \Lambda}{\partial c_t} : e^{g_t}\left(c_t - H_t\right)^{-\sigma} - \beta\gamma \mathrm{D}\mathrm{E}_t\left[e^{g_{t+1}}\left(c_{t+1} - H_{t+1}\right)^{-\sigma}\right] - \lambda_t = 0, \tag{33}$$

$$\frac{\partial \Lambda}{\partial l_t} : \lambda_t w_t - l_t^{\chi} = 0, \tag{34}$$

$$\frac{\partial \Lambda}{\partial b_t} : \beta\left(1 + R_t\right)\mathrm{E}_t\left[\left(\frac{P_t}{P_{t+1}}\right)\lambda_{t+1}\right] - \lambda_t = 0, \tag{35}$$

$$\frac{\partial \Lambda}{\partial K_{t+1}} : \beta \mathrm{E}_t\left[\left(r_{t+1} + 1 - \delta\right)\lambda_{t+1}\right] - \lambda_t = 0. \tag{36}$$

Equation (33) simply defines λ_t, the shadow price of capital, to equal the marginal utility of consumption. Equation (34) implies that households supply labor up to the point where the marginal rate of substitution between consumption and leisure equals the consumption real wage, w_t. Equation (35) shows that the bond market clears at an aggregate stock of zero when the expected change in the shadow price of capital equals the ex ante real interest rate. Lastly, equations (36) and (35) imply that in equilibrium households are indifferent between owning bonds and capital.

Combining equations (33) and (35), the log-linear consumption Euler equation is

$$\begin{aligned}\mathrm{E}_t\Delta\widehat{c}_{t+1} &= \frac{\gamma}{\left(1+\gamma^2\beta\mathrm{D}\right)}\mathrm{E}_t\left[\Delta\widehat{c}_t + \beta\mathrm{D}\Delta\widehat{c}_{t+2}\right] \\ &\quad + \frac{\left(1-\gamma\right)\left(1-\gamma\beta\mathrm{D}\right)}{\sigma\left(1+\gamma^2\beta\mathrm{D}\right)}\left(R_t - \mathrm{E}_t\pi_{t+1} - \rho\right) - \frac{\left(1-\gamma\right)}{\sigma\left(1+\gamma^2\beta\mathrm{D}\right)}g_t.\end{aligned} \tag{37}$$

Note that the habit formation breaks the equality between the elasticity of intertemporal substitution and the (inverse) coefficient of relative risk aversion and that, with internal habits, expected consumption for two periods ahead affects current consumption.

4.2 Real marginal costs

As noted in Sections 2 and 3, theory establishes that the Phillips curve depends on real marginal costs. Appendix A shows that real marginal costs, that is the real resources firms must spend to produce an additional unit of aggregate output, are, to a log-linear approximation, given by

$$\widehat{mc}_t = \widehat{w}_t + \widehat{l}_t - \widehat{y}_t. \tag{38}$$

Equation (38), which says that real marginal costs are proportional to labor's share of income, explains the use of labor's share as a measure of real marginal costs when Phillips curves are estimated using single-equation GMM. However, the goal here is to estimate the Galí-Gertler

Phillips curve as part of a complete system, in order to study general equilibrium outcomes following shocks. To this end, we exploit labor market clearing to obtain a relationship between real marginal costs and the consumption gap.

Because the labor market clears when the consumption real wage equals the marginal rate of substitution between consumption and leisure, and habit formation affects the marginal rate at which households are prepared to substitute between consumption and leisure, it should come as no surprise that habit formation affects the dynamics of real wages and real marginal costs. In fact, Appendix A shows that the log-linear relationship between real marginal costs and consumption is given by

$$\begin{aligned}\widehat{mc}_t &= \left[\chi + \frac{\sigma\left(1+\gamma^2\beta \mathrm{D}\right)}{(1-\gamma)(1-\gamma\beta\mathrm{D})}\right]\widehat{c}_t - \frac{\sigma\gamma}{(1-\gamma)(1-\gamma\beta\mathrm{D})}\left(\widehat{c}_{t-1} + \beta\mathrm{D}\mathrm{E}_t\widehat{c}_{t+1}\right) \\ &\quad -(1+\chi)\,u_t - \frac{1}{(1-\gamma\beta\mathrm{D})}g_t. \qquad (39)\end{aligned}$$

The marginal utility of consumption is higher when consumption was high last period, which, for a given real wage, induces households to increase their labor supply in order to boost consumption and raise utility. At the macro-level, the labor supply increase lowers the market-clearing real wage and real marginal costs; hence the negative coefficient on lagged consumption in equation (39). More generally, real marginal costs rise when the consumption gap increases, but fall in response to either a positive technology shock or a positive consumption preference shock. It is not difficult to see why a positive technology shock lowers real marginal costs; positive technology shocks make it possible to produce more from given inputs. The negative coefficient on the consumption preference shock arises because a positive consumption preference shock raises the marginal utility of consumption, which, for a given real wage, induces households to substitute from leisure into consumption, and, in aggregate, the increased labor supply lowers the real wage and real marginal costs. The intuition for the positive coefficient on consumption is analogous to that for the consumption preference shock.

4.3 Central bank

The final actor in the model is the central bank. We do not develop a micro-founded model of central bank behavior. Instead, we posit a descriptive model that appears to characterize policy outcomes well over the business cycle. We assume that R_t is set according to

$$R_t = (1-\phi_3)\left[\rho + (1-\phi_1)\,\overline{\pi} + \phi_1\mathrm{E}_t\pi_{t+1} + \phi_2\widehat{c}_t\right] + \phi_3 R_{t-1} + \epsilon_t, \qquad (40)$$

which is a standard forward-looking Taylor-type rule, essentially the same as the specification studied by Clarida, Galí, and Gertler (1998, 2000). Equation (40) postulates that the central bank responds with inertia to future expected inflation and, through consumption, to the state of the business cycle. Expected future inflation rather than current or lagged inflation enters the rule to capture the fact that central banks consider the future evolution of the economy when conducting monetary policy.

5 Model estimation

The models that we seek to estimate are summarized by equations (17), (20), (39), (37), and (40), and each has a rational expectations solution of the form

$$\mathbf{z}_t = \mathbf{h} + \mathbf{H}\mathbf{z}_{t-1} + \mathbf{G}\mathbf{v}_t, \tag{41}$$

where $\mathbf{z}_t = \begin{bmatrix} \pi_t & \widehat{c}_t & R_t \end{bmatrix}'$ and $\mathbf{v}_t = \begin{bmatrix} u_t & g_t & \varepsilon_t \end{bmatrix}'$ and $\mathbf{h}$, $\mathbf{H}$, and $\mathbf{G}$ are each functions of $\mathbf{\Gamma}$, which denotes the vector of parameters to be estimated. By construction the eigenvalues of $\mathbf{H}$ are bounded by one in modulus.

The models are estimated using two likelihood-based estimators: maximum likelihood and Bayesian estimation. The maximum likelihood estimates reveal the model parameterizations that are most likely to have generated the data, but the axiom of correct specification that underlies maximum likelihood makes model comparison conceptually troublesome. By way of contrast, through the construction of posterior model probabilities and Bayesian posterior odds, the Bayesian approach readily accommodates model uncertainty and offers a coherent framework for comparing the different models; where possible, we present both sets of estimates, but rely on the Bayesian approach for model comparison.

5.1 Maximum likelihood estimation

Because the rational expectations equilibrium for each model takes the form of equation (41), the assumption that $\mathbf{v}_t \sim niid\,[\mathbf{0}, \mathbf{\Omega}]$ implies that the concentrated log-likelihood function for the model is

$$\log L_c\left(\Gamma; \{\mathbf{z}_{1t}\}_2^T | \mathbf{z}_{11}\right) \propto (T-1)\ln\left[\left|(\mathbf{G}_1)^{-1}\right|\right] - \frac{(T-1)}{2}\ln\left(\left|\widehat{\mathbf{\Omega}}\right|\right), \tag{42}$$

where

$$\widehat{\mathbf{\Omega}} = \sum_{t=2}^{T} \frac{\left[\mathbf{G}^{-1}\left(\mathbf{z}_{1t} - \mathbf{h}_1 - \mathbf{H}_{11}\mathbf{z}_{1t-1}\right)\right]\left[\mathbf{G}^{-1}\left(\mathbf{z}_{1t} - \mathbf{h}_1 - \mathbf{H}_{11}\mathbf{z}_{1t-1}\right)\right]'}{T-1}.$$

The parameters, $\mathbf{\Gamma}$, are obtained by maximizing equation (42), with their standard errors determined from the inverted-Hessian evaluated at the optimum. The maximum likelihood estimates were obtained using a two-step approach. In the first step the genetic algorithm described in Duffy and McNelis (2001) was used. Briefly, this genetic algorithm is a stochastic search method that facilitates searching for a maximum over a wide parameter space. To implement the algorithm a population of 1000 initial candidate solutions was used, where the candidate solutions were drawn from a multivariate uniform distribution whose bounds were chosen to ensure that the model had a unique stable equilibrium within the search area. The genetic algorithm was allowed to run for a maximum of 300 generations or until each of the candidate solutions was identical to 5 decimal places, producing a set of first-step parameter estimates.[10] In the second step, the parameter values that emerged from the genetic algorithm were used to initialize the BFGS optimization algorithm, which was iterated to convergence. The estimates that we report, and their standard errors, reflect the maximum obtained by the BFGS algorithm. This two-step approach allows us to search widely over the likelihood function, helping ensure that a global maximum rather than a local maximum is located.

5.2 Bayesian estimation

Let $\mathbf{M}$ denote the model space and $m_j \in \mathbf{M}$, $j \in \{1, 2, ..., M\}$, reference an arbitrary model. With the parameters of model m_j represented by $\mathbf{\Gamma}_{m_j}$, $p\left(\mathbf{\Gamma}_{m_j}|m_j\right)$, is the prior density for $\mathbf{\Gamma}_{m_j}$, $p\left(\{\mathbf{z}_{1t}\}_2^T|\mathbf{\Gamma}_{m_j}, m_j\right)$ is the conditional data density, and $p\left(\mathbf{\Gamma}_{m_j}|\{\mathbf{z}_{1t}\}_2^T, m_j\right)$ is the posterior density of the parameter density conditional on the data and the model. As always with Bayesian estimation, interest centers on the posterior density, which from Bayes theorem, is given by

$$p\left(\mathbf{\Gamma}_{m_j}|\{\mathbf{z}_{1t}\}_2^T, m_j\right) = \frac{p\left(\{\mathbf{z}_{1t}\}_2^T|\mathbf{\Gamma}_{m_j}, m_j\right) p\left(\mathbf{\Gamma}_{m_j}|m_j\right)}{p\left(\{\mathbf{z}_{1t}\}_2^T\right)}. \tag{43}$$

To draw from the posterior density, we use the random walk chain Metropolios-Hastings algorithm. Ten over dispersed chains of length $60,000$ were constructed from which the first $10,000$ "burn-in" draws were discarded, leaving a total of $500,000$ usable draws. Convergence

[10] Advantages to using a genetic algorithm are that it does not require taking numerical derivatives and, by sampling over the entire admissible parameter space, that it helps to ensure that a global maximum of the likelihood function is obtained. Other than the fact that "mutation" was not applied in the creation of "children," the genetic algorithm employed in this paper follows precisely that described in Appendix A of Duffy and McNelis (2001), to which interested readers are referred. Mutation was not applied since the solution obtained from the genetic algorithm was not the final estimate, but rather only the source of starting values for a quasi-Newton hill climber.

of the chains was determined using Gelman's (1995) diagnostic and Geweke's (1992) diagnostic.

For model comparison, we use Geweke's (1999) modification of the Gelfand and Dey (1994) method to calculate the marginal data density, or marginal likelihood,

$$p\left(\{z_t\}_2^T|m_j\right) = \int_{\mathbf{\Gamma}_{m_j}} p\left(\{z_t\}_2^T|\mathbf{\Gamma}_{m_j}, m_j\right) p\left(\mathbf{\Gamma}_{m_j}|m_j\right) d\mathbf{\Gamma}_{m_j}, \tag{44}$$

which is the probability of observing the data given model m_j. As equation (44) shows, the marginal likelihood is evaluated by averaging the conditional data density with respect to the prior density. After evaluating the marginal likelihood for each model, the posterior probability associated with model $m_k \in \mathbf{M}$ can be calculated according to

$$p\left(m_k|\{z_t\}_2^T\right) = \frac{p\left(\{z_t\}_2^T|m_k\right) p(m_k)}{\sum_{j=1}^M p\left(\{z_t\}_2^T|m_j\right) p(m_j)}, \tag{45}$$

where $p(m_j)$ is the prior probability associated with model m_j, $j \in \{1, 2, ..., M\}$.[11]

5.3 Data

To estimate the models, we use U.S. data spanning the period 1982.Q1 – 2002.Q4, which excludes the period of non-borrowed reserves targeting that occurred in the early 1980s, but otherwise reflects the time during which Volcker and Greenspan were Federal Reserve chairmen. We use the quarterly average of the federal funds rate to measure R_t, use $100 \times \ln\left(C_t/C_t^T\right)$ to measure the consumption gap, where C_t is real consumption and C_t^T is trend consumption,[12] and use $400 \times \ln(P_t/P_{t-1})$, where P_t is the PCE price index, to measure inflation.

5.4 Priors

Aside from the parameters describing the shock processes, the key model parameters are $\mathbf{\Gamma} = \{\chi, \rho, \gamma, \sigma, \theta, \omega, \overline{\pi}, \phi_1, \phi_2, \phi_3\}$. However, the data are sufficiently uninformative of the labor supply elasticity, χ, that precise estimates could not be obtained using maximum likelihood. As a consequence, and to enable comparison between the FIML and the Bayesian estimates, we set χ equal to 0.80 during estimation, with this elasticity value based on Smets and Wouters (2003). The priors for the remaining behavioral parameters are summarized in Table 2a.

[11] Schorfheide (2000) provides an interesting exposition of how Bayesian methods can be used to estimate dynamic optimization-based macro-models, like those studied in this paper.

[12] Trend consumption was constructed using the Hodrick-Prescott filter with $\lambda = 1600$.

Table 2a: Prior for Structural Parameters			
Parameter	Distribution	Mean	Standard Deviation
ρ	Normal	2.50	0.50
γ	Beta	0.75	0.10
σ	Gamma	4.00	2.00
θ	Beta	0.50	0.10
ω	Beta	0.80	0.10
$\overline{\pi}$	Normal	3.00	0.50
ϕ_1	Normal	1.50	0.20
ϕ_2	Normal	1.00	0.20
ϕ_3	Beta	0.75	0.10

Briefly, the priors for ρ and $\overline{\pi}$ have means equaling 2.50 and 3.00 percent, respectively, at annual rates. The priors for γ and ϕ_3 are each Beta distributions with means equaling 0.75, while that for the inflation indexation parameter, ω, also a Beta distribution, has mean equaling 0.80. Building in information from Bils and Klenow (2004), we use a prior for θ that has a Beta distribution with a mean of 0.50. The prior for the coefficient of relative risk aversion, σ, has a Gamma distribution with a mean equaling 4.00 and, to reflect the wide range of estimates in the literature, a relatively large standard deviation of 2.00. This prior distribution broadly reflects the range parameter values available in the literature.

The prior for the shock process was implemented as follows. First, the solution to the rational expectations model, equation (41), was written as

$$\mathbf{z}_t = \mathbf{h} + \mathbf{H}\mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_t, \tag{46}$$

where $\boldsymbol{\varepsilon}_t = \mathbf{G}\mathbf{v}_t$ are reduced form shocks. The priors for the elements in $\mathbf{\Sigma} = \text{E}\left[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t^{'}\right]$ are summarized in Table 2b.

Table 2b: Prior for Reduced Form Shocks			
Parameter	Distribution	Mean	Standard Deviation
σ_1	Inverted Gamma	1.00	0.20
σ_2	Inverted Gamma	0.50	0.20
σ_3	Inverted Gamma	0.70	0.20
$\text{cov}(\varepsilon_1\varepsilon_2)$	Normal	0.00	0.20
$\text{cov}(\varepsilon_1\varepsilon_3)$	Normal	0.00	0.20
$\text{cov}(\varepsilon_2\varepsilon_3)$	Normal	0.00	0.20

With respect to the model space, $\mathbf{M}$, we apply a discrete uniform prior to the model space, thus $p(m_j) = \frac{1}{M}$, $j \in \{1, ..., M\}$.

6 Model estimates

We estimate six specifications, interacting three Phillips curve specifications with, as a robustness check, external and internal habit formation. We begin by estimating specifications that employ the Galí-Gertler Phillips curve to examine whether it is consistent with micro-evidence on the frequency of price adjustment. Then, to determine whether the Galí-Gertler Phillips curve improves on the main alternatives,[13] and to assess how the competing models behave following shocks, we estimate specifications based on the Calvo Phillips curve and on the full-indexation Phillips curve, and evaluate their posterior model probabilities.[14]

6.1 Estimates with the Galí-Gertler Phillips curve

Table 3 presents our estimates of the specifications that employ the Galí-Gertler Phillips curve; Panels A and B report estimates for external habit formation and internal habit formation, respectively. The table displays FIML estimates, with standard errors in parentheses,[15] along with the posterior mean, median, and mode, and a 90 percent probability interval centered about the posterior median. Also shown are the maximized value of the (log-) likelihood function (log-L) and the (log-) marginal likelihood (log-ML), followed by the posterior model probability in parentheses. The FIML estimates are shown so that interested readers can assess the role the prior has in shaping the Bayesian estimates.

Note that the estimates in Table 3 reveal that the posterior mean, median, and mode are all very similar, indicating that the posterior distributions are all relatively symmetric.[16] Note, also, that, with one exception, the FIML estimates are very similar to the Bayesian estimates, indicating that the prior is not having a large influence on the Bayesian estimation. The one exception is the coefficient of relative risk aversion, σ. The imprecision associated with the FIML estimate indicates that the data contain relatively little information about σ, and it is this that allows the prior to influence the Bayesian estimation. Finally, note that

[13] Note that because the Gali-Gertler Phillips curve encompasses the Calvo Phillips curve and the full-indexation Phillips curve, it must have a higher maximized likelihood function than these specifications. However, due to the effect of the prior, the same logic does not apply to Bayesian estimation. With Bayesian estimation, a model can have a higher maximized likelihood than another, but have a smaller marginal likelihood and a smaller posterior model probability.

[14] For the Bayesian estimation of the models containing the Calvo Phillips curve, we set $\omega = 0$, with the priors for the remaining parameters given by Table 2. Similarly, to estimate the models containing the full-indexation Phillips curve, we set $\theta = 0$, with the remainder of the priors given by Table 2.

[15] These standard errors were calculated from the inverted Hessian matrix evaluated at the maximum of the likelihood function.

[16] This feature of the posterior distributions has been verified by plotting the marginal densities for each parameter. A figure showing these marginal densities is available upon request.

the estimates for the model with internal habits (Panel B) are essentially the same as those for the model with external habits (Panel A), thus the findings are robust to this modeling choice.

Table 3: FIML and Bayesian Estimates of the Galí-Gertler Model					
Panel A	External Habit Formation				
Parameter	FIML	Post. Mean	Post. Median	Post. Mode	90% Interval
ρ	2.517 (0.726)	2.576	2.577	2.576	$[2.182, 2.967]$
γ	0.821 (0.084)	0.821	0.821	0.823	$[0.768, 0.873]$
σ	5.665 (5.533)	3.957	3.916	3.808	$[2.462, 5.584]$
θ	0.303 (0.072)	0.361	0.360	0.361	$[0.301, 0.420]$
ω	0.966 (0.017)	0.951	0.951	0.951	$[0.937, 0.965]$
$\overline{\pi}$	3.339 (1.381)	3.245	3.243	3.253	$[2.902, 3.595]$
ϕ_1	1.634 (1.120)	1.563	1.561	1.561	$[1.377, 1.753]$
ϕ_2	1.282 (1.174)	1.001	1.001	1.001	$[0.793, 1.209]$
ϕ_3	0.867 (0.039)	0.849	0.849	0.849	$[0.822, 0.875]$
log-L	-262.321	–			
log-ML	–	-287.196 (0.509)			
Panel B	Internal Habit Formation				
Parameter	FIML	Post. Mean	Post. Median	Post. Mode	90% Interval
ρ	2.325 (0.918)	2.510	2.511	2.504	$[2.094, 2.923]$
γ	0.794 (0.070)	0.794	0.794	0.795	$[0.744, 0.842]$
σ	5.465 (5.530)	3.599	3.561	3.438	$[2.121, 5.217]$
θ	0.289 (0.075)	0.347	0.347	0.347	$[0.288, 0.405]$
ω	0.968 (0.017)	0.952	0.952	0.952	$[0.938, 0.965]$
$\overline{\pi}$	3.335 (1.592)	3.248	3.246	3.256	$[2.904, 3.600]$
ϕ_1	1.553 (1.185)	1.563	1.561	1.560	$[1.377, 1.753]$
ϕ_2	1.081 (1.192)	0.986	0.986	0.985	$[0.777, 1.194]$
ϕ_3	0.873 (0.042)	0.855	0.855	0.855	$[0.828, 0.881]$
log-L	-263.093	–			
log-ML	–	-287.998 (0.228)			

Turning to the coefficient values themselves, given the similarity between the estimates in Panels A and B, and the fact that the model with external habit formation has the greater posterior model probability, we focus our discussion on Panel A. The FIML estimate of the rate of time preference, ρ, is 2.517, while the Bayesian estimates are centered on 2.576. These values are consistent with estimates of the equilibrium real interest rate (Laubach and Williams, 2003) and place the quarterly discount factor at 0.994, in line with values widely used in calibration exercises, such as Rouwenhorst (1995).

Looking at the utility function parameters, the habit formation parameter, γ, is estimated to be about 0.821, implying that habit formation is important and that there is considerable inertia in consumption. Elsewhere, estimates of γ vary widely. Smets and Wouters (2003) estimate $\gamma = 0.54$, Altig, Christiano, Eichenbaum, and Linde (2004) estimate $\gamma = 0.65$, Giannoni and Woodford (2003) estimate $\gamma = 1.00$, while the results in Smets (2003) and Cho and Moreno (2005) imply that γ equals 0.79 and 1.00, respectively. With internal habit formation, Fuhrer (2000) estimates $\gamma = 0.8$, while Christiano, Eichenbaum, and Evans (2005) estimate $\gamma = 0.63$. Calibration exercises, based on either internal or external habit formation, often set γ to 0.80 (McCallum and Nelson, 1999). Turning to σ, estimates in the literature are also wide-ranging. The FIML estimate is imprecise, but places σ at about 5.67, while the Bayesian estimation returns a posterior mean for σ that is just under 4. Elsewhere, Fuhrer (2000) obtains $\sigma = 6.11$ while Kim (2000) obtains $\sigma = 14.22$. Using Bayesian methods, Smets and Wouters (2003) get $\sigma = 1.39$ for the posterior mean, while Levin, Onatski, Williams, and Williams (2005) get $\sigma = 2.19$ for the posterior mean.[17] At the other end of the spectrum, Rotemberg and Woodford (1997) estimate $\sigma = 0.16$, Amato and Laubach (2003) estimate $\sigma = 0.26$, and Giannoni and Woodford (2003) estimate $\sigma = 0.75$.

Regarding the policy-rule parameters, the FIML estimation places the implicit inflation target at around 3.3 percent, while the Bayesian estimation has $\overline{\pi}$ approximately equal to 3.25 percent. These estimates of $\overline{\pi}$ are very similar to those obtained by Clarida, Galí, and Gertler (2000), while being slightly higher than those obtained by Favero and Rovelli (2003), who estimate $\overline{\pi}$ to be 2.63 percent. We estimate the coefficient on expected future inflation to be about 1.6, the coefficient on the consumption gap to be about 1, and the coefficient on lagged interest rates to be about 0.85. These coefficients are all consistent with other estimated Taylor-type rules (see Clarida, Galí, and Gertler (2000) and Dennis (2006)), indicating an activist, but inertial, approach to monetary policy.

[17] As noted also in the context of Table 3, the Bayesian estimates of σ obtained by Smets and Wouters (2003) and Levin, Onatski, Williams, and Williams (2005) appear to be governed largely by their prior densities, indicating that the data are relatively uninformative for this elasticity. Smets and Wouters (2003) use a normal density, with a mean of 1 and a standard deviation of 0.375, for their prior and obtain a posterior mean of 1.39. Similarly, Levin, Onatski, Williams, and Williams (2005) use a normal density, with a mean of 2 and a standard deviation of 0.5, for their prior and obtain a posterior mean of 2.19.

With respect to pricing behavior, the two key parameters are θ and ω. The FIML estimate of θ is 0.303, while a two-standard-deviation confidence interval spans 0.159 to 0.447. The Bayesian estimation has the distribution for θ centered on about 0.361, with a 90 percent probability interval covering 0.301 to 0.420. These estimates place the discrete-time frequency of price adjustment somewhere around 0.70 (FIML) and 0.64 (Bayesian), representing relatively frequent price adjustment. Recall that the Bils-Klenow data-set, which shaped the prior for θ, suggested a discrete-time frequency of price adjustment of about 0.5. Thus, one of the main findings that emerges from the estimation of the Galí-Gertler Phillips curve is that macro-data are consistent with frequent price adjustment. In fact, if any criticism is to be leveled at the specification it is that the model implies that firms change prices too frequently, not too infrequently.

Finally, because the estimates of θ reveal that firms do, in fact, change prices quite frequently, they suggest that menu costs are not a huge impediment to a firm changing its price. At the same time, the estimates of ω are large, and they imply that most firms that change prices do so using a rule of thumb, a result that is consistent with the Zbaraki, Ritson, Levy, Dutta, and Bergin (2004) conclusion that information gathering/processing costs are more important for pricing behavior than menu costs.

6.2 Estimates with the Calvo Phillips curve

To estimate versions of the model with the Calvo Phillips curve, we set ω, the share of rule-of-thumb pricing firms, to zero. With this restriction, θ is equivalent to ξ in the Calvo model and it is labeled as such in Table 4 below. As earlier, we consider both internal and external habit formation; however, due to weak identification of some parameters,[18] we report only the Bayesian estimates.

[18] This weak identification arises because the absence of rule-of-thumb pricing means that information in lagged inflation is no longer available for econometric identification.

Table 4: Bayesian Estimates of the Calvo Model				
Panel A	External Habit Formation			
Parameter	Post. Mean	Post. Median	Post. Mode	90% Interval
ρ	2.593	2.594	2.591	$[2.228, 2.953]$
γ	0.802	0.803	0.804	$[0.745, 0.858]$
σ	2.632	2.598	2.496	$[1.581, 3.797]$
ξ	0.942	0.942	0.942	$[0.930, 0.953]$
$\overline{\pi}$	3.474	3.471	3.486	$[3.200, 3.759]$
ϕ_1	1.580	1.579	1.577	$[1.406, 1.758]$
ϕ_2	1.026	1.026	1.026	$[0.822, 1.232]$
ϕ_3	0.826	0.826	0.827	$[0.799, 0.852]$
log-ML	-338.516 $(2.6e^{-23})$			
Panel B	Internal Habit Formation			
Parameter	Post. Mean	Post. Median	Post. Mode	90% Interval
ρ	2.490	2.492	2.476	$[2.088, 2.890]$
γ	0.754	0.755	0.755	$[0.702, 0.806]$
σ	2.785	2.746	2.611	$[1.611, 4.091]$
ξ	0.944	0.944	0.944	$[0.933, 0.955]$
$\overline{\pi}$	3.467	3.464	3.482	$[3.196, 3.750]$
ϕ_1	1.579	1.577	1.576	$[1.404, 1.759]$
ϕ_2	0.986	0.986	0.984	$[0.779, 1.191]$
ϕ_3	0.829	0.830	0.831	$[0.801, 0.857]$
log-ML	-340.598 $(3.3e^{-24})$			

Looking at the estimates reported in Table 4, it is noteworthy that where these specifications have parameters in common with those estimated in Table 3, similar values are obtained, attesting to their structural nature. However, although similar estimates of common parameters are obtained, important differences between these specifications and those estimated earlier can be found in the (log-) marginal likelihoods and posterior model probabilities. Relative to the specifications estimated in Table 3, those in Table 4 have much lower marginal likelihoods and, given the discrete uniform prior over the model space, much lower posterior model probabilities. Essentially, having examined their fit to the data, we now attribute a probability of almost zero to the possibility that the Calvo model generated the data.

So, how do the estimates of the New Keynesian Phillips curve relate to those of the Galí-Gertler Phillips curve obtained earlier? Clearly, if ξ is interpreted naively as the share of firms that do not change prices each period, then the conclusions from the two models are very different. However, the Galí-Gertler model implies that the share of firms that do not make an optimal price change each period is given not by θ, but by $\theta + \omega(1-\theta)$, and now the two sets of estimates are easily reconciled. For example, using the posterior mean from Table 3 (Panel A), $\theta + \omega(1-\theta)$ equals 0.969, which is very similar to the posterior mean of ξ in Table

4 (Panel A), which equals 0.942. Overall, then, these indicate that the Calvo-share may be a relatively unbiased estimate of the share of firms that do not make an optimal price change, but it is a highly biased estimate of the discrete-time frequency of price adjustment, and, as a consequence, drastically overstates the implied average duration between price changes.

At the same time, the estimates of ξ shown in Table 4 are generally larger than those reported in Table 1. One possible reason for these larger estimates of ξ is that the estimates are shaped by the properties of a complete model, not just by the properties of the New Keynesian Phillips curve. Related to this, the likelihood-based estimators impose the discipline of "relevance" to the choice of econometric instruments, preventing possibly irrelevent variables from serving as instruments. A third possible reason for the larger estimates of ξ is that we do not use labor's share of income as our measure of real marginal costs. Instead, the measure of real marginal costs is derived within our models based on labor market clearing, essentially imposing equilibrium behavior on real wages, output, and hours worked.

6.3 Estimates with the full-indexation Phillips curve

The final two models that we estimate employ the full-indexation Phillips curve. As established earlier, the full-indexation Phillips curve can be obtained as a special case of the Galí-Gertler Phillips curve by setting θ equal to zero. With this restriction the parameter ω is equivalent to ξ, and it is labeled this way in Table 5. As earlier, the parameters in the policy rule and those that govern household behavior are essentially the same as those shown in Tables 3 and 4; clearly these parameters are robust to changes in how firms' pricing behavior is modeled. In addition, the coefficient values shown in Panel B, which relate to the model with internal habit formation, reinforce those shown in Panel A, which relate to the model with external habit formation, from which it follows that the findings with respect to pricing behavior are robust to how the habit formation is modeled.

Table 5: FIML and Bayesian Estimates of the Full-Indexation Model					
Panel A	External Habit Formation				
Parameter	FIML	Post. Mean	Post. Median	Post. Mode	90% Interval
ρ	2.643 (0.508)	2.651	2.653	2.651	$[2.281, 3.017]$
γ	0.817 (0.077)	0.835	0.835	0.837	$[0.780, 0.888]$
σ	7.583 (7.344)	4.281	4.231	4.098	$[2.581, 6.157]$
ξ	0.945 (0.021)	0.936	0.936	0.936	$[0.918, 0.953]$
$\overline{\pi}$	2.855 (0.674)	2.973	2.974	2.974	$[2.536, 3.411]$
ϕ_1	1.700 (0.620)	1.529	1.528	1.530	$[1.321, 1.737]$
ϕ_2	1.211 (1.082)	1.016	1.016	1.016	$[0.799, 1.234]$
ϕ_3	0.851 (0.035)	0.836	0.836	0.837	$[0.806, 0.865]$
log-L	-267.611	–			
log-ML	–	-288.430 (0.148)			
Panel B	Internal Habit Formation				
Parameter	FIML	Post. Mean	Post. Median	Post. Mode	90% Interval
ρ	2.616 (0.522)	2.618	2.620	2.618	$[2.242, 2.991]$
γ	0.804 (0.072)	0.816	0.817	0.818	$[0.765, 0.867]$
σ	6.426 (7.330)	3.747	3.697	3.549	$[2.108, 5.558]$
ξ	0.948 (0.020)	0.941	0.941	0.940	$[0.924, 0.957]$
$\overline{\pi}$	2.828 (0.688)	2.958	2.958	2.958	$[2.520, 3.391]$
ϕ_1	1.696 (0.620)	1.532	1.532	1.532	$[1.326, 1.738]$
ϕ_2	1.111 (1.065)	1.006	1.006	1.005	$[0.790, 1.222]$
ϕ_3	0.852 (0.035)	0.840	0.841	0.841	$[0.810, 0.870]$
log-L	-268.057	–			
log-ML	–	-288.688 (0.114)			

So what are the main results in Table 5? First, the estimates of ξ are very much in line with those obtained for the specifications using the Calvo Phillips curve, so although inflation is not endogenously persistent in the Calvo model, this does not seem to be distorting its estimates of ξ. Second, the estimates of ξ place the share of firms that optimize their price changes at just over 5 percent per quarter.

However, the similarities in the parameter estimates between Table 5 and Table 3 cloak important differences between the models. From an economic perspective, in the full-indexation model, all firms change their prices every period, which, of course, is inconsistent with micro-data. By way of contrast, in the Galí-Gertler model some firms change their prices optimally, some use a rule-of-thumb, and some keep their price unchanged. From a statistical perspective, the models employing the Galí-Gertler Phillips curve have higher maximized likelihoods, higher marginal likelihoods, and higher posterior model probabilities than those containing the full-indexation Phillips curve, signaling that the Galí-Gertler model receives greater support from the data. Importantly, among the six models we estimate, only in those models employing the Galí-Gertler Phillips curve did the posterior model probabilities rise relative to their prior model probabilities.

7 Pricing and New Keynesian business cycle dynamics

In this section we demonstrate that the differences between the three pricing models are not just statistical, nor are they just theoretical, rather they are economically important. We study how the models respond to shocks, considering consumption preference shocks, technology shocks, and monetary policy shocks. Furthermore, we use the Bayesian estimates to construct predictive densities for each model and for each shock, and, exploiting the posterior model probabilities, we use Bayesian model averaging to examine the (weighted) average response to each shock and to construct 90-percent probability intervals about these responses. The results for one-standard-deviation shocks[19] are shown in Figure 1, which plots the median of the predictive densities for three of the six models, together with the results from the Bayesian model averaging exercise. Although all six models are used for the Bayesian model averaging, to avoid clutter, we report the individual responses for the three Phillips curves, but only for the specifications with external habit formation.

To understand these shock responses it is useful to focus on the Bayesian model averaging exercise. Panels A to C correspond to the consumption preference shock, Panels D to F correspond to the technology shock, and Panels G to I correspond to the monetary policy shock. Following a positive consumption preference shock, households take advantage of the fact that higher utility can be achieved by consuming more now and increase their labor supply in order to raise their income to facilitate greater consumption expenditures (Panel A). The increase in labor supply is partly offset by a rise in labor demand, as firms increase production to meet rising demand, but, on balance, the market-clearing real wage declines, lowering real marginal costs and causing a small and gradual decline in aggregate inflation (Panel B).

[19] We measure one standard deviation using the mean of the posterior distributions for the shock standard errors in each model.

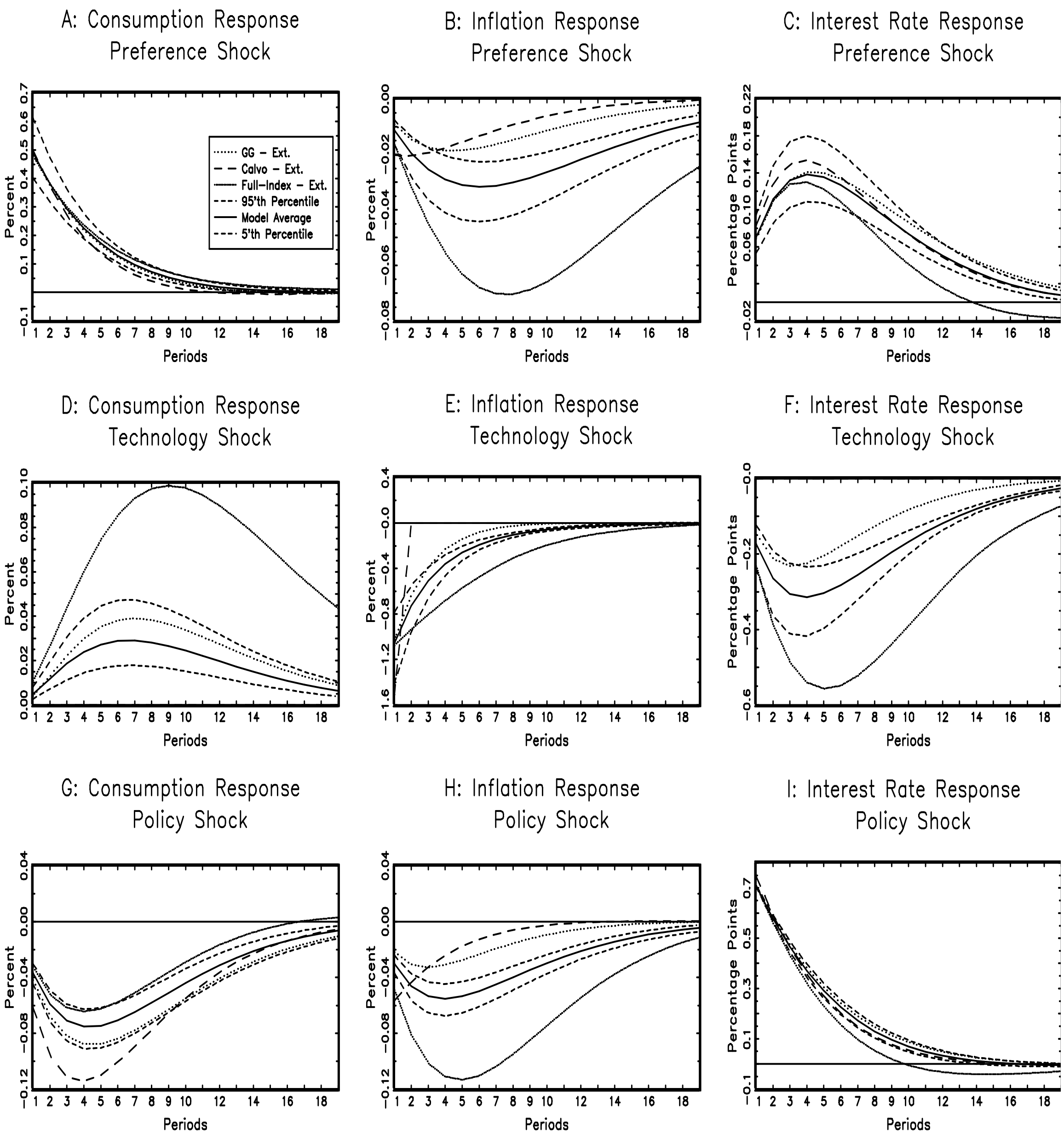

Figure 1: Predictive Densities Following One-Standard-Deviation Shocks

Faced with stronger consumer demand, and only slightly lower inflation, the central bank raises interest rates (Panel C).

Following a positive technology shock, the marginal product of labor increases pushing up the demand for labor and raising the market-clearing real wage. Household income rises due to the higher real wage, and from households increasing their hours worked, which pushes up consumption (Panel D). At the same time, the increase in the marginal product of labor has the effect of lowering real marginal costs, which puts downward pressure on inflation (Panel E). In this case, the decline in inflation is substantial and the policy response is to lower nominal, and hence also real, interest rates (Panel F). Finally, following a positive monetary policy shock, real interest rates rise (Panel I), which induces households to defer consumption. To offset the fall in consumption (Panel G), households increase their labor supply, which puts downward pressure on real wages. Although firms respond to declining demand by reducing their demand for labor, the market-clearing real wage falls, lowering real marginal costs and inflation (Panel H).

Relative to the responses of the Galí-Gertler model, the poor performance of the Calvo model is clear. Following a technology shock (Panel E) inflation falls, but then immediately returns to baseline, without any effect on consumption or interest rates. More generally, the Calvo model's behavior following all three shocks differ importantly from the Galí-Gertler model in that they are not "hump-shaped," underscoring the Estrella and Fuhrer (2002) criticism of the New Keynesian Phillips curve. With regard to the full-indexation model, although its responses are hump-shaped, they are also generally much larger than those of either the Galí-Gertler model or the Bayesian model average. These large responses are particularly evident in how the model behaves following the technology shock (Panels D to F), but are also evident in Panels B and H. The source of these large responses is the fact that many firms index to lagged inflation and no firms keep their prices unchanged following shocks. Inflation's large responses give rise to large interest rate responses by the central bank, which, in turn, generate relatively large consumption responses by households. It is clear from the posterior model probabilities and the behavior of the Bayesian model average that the data provide considerably less support for the behavior of the full-indexation model than they do for the behavior of the Galí-Gertler model.

8 Conclusion

The New Keynesian Phillips curve, generally derived from the Calvo model, has been widely criticized for being economically implausible, for being inconsistent with micro-data on the frequency of price adjustment, and for being unable to account for the persistence in inflation. Popular alternatives to the Calvo model, such the full-indexation model and the partial-indexation model are much better able to explain the persistence in inflation, but, because they assume that all prices change every period, they too are economically implausible and are unable to match micro-evidence on the frequency with which actual prices change. These criticisms are important because New Keynesian business cycle models are increasingly used to study issues such as how monetary policy should be conducted to maximize welfare, and the nature of these policies hinge critically on precisely how and why prices are rigid. More generally, they challenge whether the leading New Keynesian models of price adjustment provide a useful and economically sensible description of inflation dynamics. Against this background, the main contribution of this paper is to demonstrate that the Galí and Gertler 1999) pricing model can successfully address these criticisms.

We begin by presenting estimates of Calvo-share obtained when the New Keynesian Phillips curve is estimated in isolation, outlining the implications of these estimates for the average duration between price changes. Next, we emphasize that issues such as heterogeneity in the frequency of price adjustment across firms, continuous-time price setting, as well as the conceptual difference between a price change and an optimal price change mean that a meaningful comparison of the average durations between price changes obtained from micro-data to those implied by estimates of the Calvo-share is not straightforward. However, to the extent that these issues can be addressed, they confirm that estimates of the Calvo-share are inconsistent with Bureau of Labor Statistics price data, and to the extent that they cannot be addressed, they call for a micro-founded model of price setting that can distinguish conceptually between price changes and optimal price changes.

Next, we introduce the rule-of-thumb pricing model developed by Galí and Gertler (1999), a model in which each period a share of firms get to change their prices and within this share a proportion change their prices optimally while the remaining proportion change their prices by a rule of thumb. We highlight that when it comes to reconciling macro- and micro-evidence on the frequency of price adjustment, and to accounting for the persistence in inflation, the Galí-Gertler pricing environment holds many attractions. First, the model is one in which not all prices change every period and in which when prices change they do not necessarily change optimally. Second, the model's share parameters can be interpreted easily in light of the costs firms face when changing prices. Traditional menu cost factors, which affect all price changes not just optimal price changes, are readily associated with the share of firms that change their prices. Similarly, information gathering/processing costs are readily associated with the share of price-changing firms that resort to rule-of-thumb price-setting. Third, because the rule of thumb is one in which firms index their prices to lagged inflation, the model has a mechanism for generating intrinsic inflation persistence.

After outlining the Galí-Gertler pricing model, we derive its Phillips curve and relate it to other specifications in the literature. Specifically, we prove that the Galí-Gertler Phillips curve encompasses the Calvo Phillips curve, the full-indexation Phillips curve, and the partial-indexation Phillips curve, from which it follows that the Galí-Gertler Phillips curve can explain inflation at least as well as these more widely used alternatives. This encompassing result, together with the fact that the full- and the partial-indexation models counterfactually force all firms to change their price every period, makes the Galí-Gertler Phillips curve particularly attractive for empirical applications. Taking this as motivation, we build a small-scale New Keynesian business cycle model and estimate versions of it on U.S. macroeconomic data.

The main empirical results are as follows. First, broadly in line with the Bils and Klenow study of Bureau of Labor Statistics price data, which suggest a quarterly frequency of price adjustment of about 0.5, our estimates of the Galí-Gertler model place the quarterly frequency of price adjustment at just over 0.6. In this respect, the Galí-Gertler model is a considerable improvement on the Calvo model, and to the extent that it is at odds with the Bils-Klenow study it is because the model implies too little price rigidity rather than too much. Second, with around 60 percent of firms changing their prices each quarter and with 95 percent of them resorting to rule-of-thumb price-setting, our estimates are consistent with the view that menu costs are a much less important factor for price setting than information gathering/processing costs. These findings are robust to whether households have internal or external habit formation and to whether the model is estimated using FIML or Bayesian methods.

Third, reflecting the model's empirical advantages, the Bayesian estimation assigns much higher posterior model probabilities to the models containing the Galí-Gertler Phillips curve, particularly the specification with external habit formation, than it does to the models containing either the Calvo Phillips curve or the full-indexation Phillips curve. Our finding that a posterior model probability of almost zero is attributed to the Calvo Phillips curve is consistent with the fact that, unlike the Galí-Gertler model, the Calvo model cannot generate hump-shaped responses to shocks.

Clearly the Galí and Gertler (1999) model offers important advantages over other popular pricing models and deserves greater empirical attention as a consequence. At the same time, the model's micro-foundations could be made more rigorous in as much as its share parameters are standing in for a more complicated optimization problem confronting firms, a problem with a state-contingent aspect. Although it is well-known that state-contingent and time-consistent pricing models behave similarly when inflation is low and stable, the obvious next step would be to build the menu costs and the information gathering/processing costs formally and directly into the firm's pricing problem. We leave this interesting and important exercise for future work.

A Appendix: Aggregate real marginal costs

Cost minimization implies that firms rent capital and labor such that

$$\frac{W_t}{p_t(i)} = mc_t(i)\,\frac{\kappa y_t(i)}{l_t(i)},$$

implying that a firms' real marginal costs depend on the ratio of its *production* real wage to its marginal productivity of labor, i.e.,

$$\begin{aligned} mc_t(i) &= \frac{1}{\kappa}\frac{W_t l_t(i)}{p_t(i)\,y_t(i)} \\ &= \frac{1}{\kappa}\frac{w_t l_t(i)}{y_t^{\frac{1}{\epsilon}}\,y_t(i)^{\frac{\epsilon-1}{\epsilon}}}, \end{aligned} \tag{A1}$$

where equation (A1) has been expressed in terms of the *consumption* real wage. Of course, since all firms face the same rental prices for capital and labor and are subject to the same aggregate technology shock, they employ capital and labor in the same ratio and share the same real marginal costs.

Although all firms face the same real marginal cost, it is still convenient to define aggregate real marginal costs through the aggregator

$$mc_t \equiv \left[\int_0^1 mc_t(i)^{1-\epsilon}\,di\right]^{\frac{1}{1-\epsilon}}. \tag{A2}$$

Combining equations A1 and A2 results in

$$\begin{aligned} mc_t &= \frac{w_t}{\kappa y_t^{\frac{1}{\epsilon}}}\left[\int_0^1 \left(\frac{l_t(i)}{y_t(i)^{\frac{\epsilon-1}{\epsilon}}}\right)^{1-\epsilon} di\right]^{\frac{1}{1-\epsilon}} \\ (1+\widehat{mc}_t)\,\overline{mc} &= \frac{(1+\widehat{w}_t)\,\overline{w}}{\kappa(1+\widehat{y}_t)^{\frac{1}{\epsilon}}\,\overline{y}^{\frac{1}{\epsilon}}}\left[\int_0^1 \left(\frac{\left(1+\widehat{l}_t(i)\right)\overline{l(i)}}{(1+\widehat{y}_t(i))^{\frac{\epsilon-1}{\epsilon}}\,\overline{y(i)}^{\frac{\epsilon-1}{\epsilon}}}\right)^{1-\epsilon} di\right]^{\frac{1}{1-\epsilon}}. \end{aligned}$$

Focusing on a symmetric equilibrium,

$$\begin{aligned}\widehat{mc}_t &\simeq \widehat{w}_t - \frac{1}{\epsilon}\widehat{y}_t + \int_0^1 \widehat{l}_t(i)\,di + \frac{1-\epsilon}{\epsilon}\int_0^1 \widehat{y}_t(i)\,di \\ \widehat{mc}_t &\simeq \widehat{w}_t + \widehat{l}_t - \widehat{y}_t, \end{aligned} \tag{A3}$$

which is equation (38) in the text. Equation (A3) establishes that to a first-order log-linear approximation aggregate real marginal costs depend on the *consumption* real wage and the aggregate marginal productivity of labor.

Turning to the firm-level production function,

$$y_t \equiv \left[\int_0^1 y_t(i)^{\frac{\epsilon-1}{\epsilon}}\,di\right]^{\frac{\epsilon}{\epsilon-1}} = \left[\int_0^1 \left([e^{u_t} l_t(i)]^{\kappa} k_t(i)^{1-\kappa}\right)^{\frac{\epsilon-1}{\epsilon}} di\right]^{\frac{\epsilon}{\epsilon-1}},$$

and log-linearizing

$$\widehat{y}_t \simeq u_t + \widehat{l}_t + (1-\kappa)\left(\widehat{k}_t - \widehat{u}_t - \widehat{l}_t\right). \tag{A4}$$

Similarly, the log-linearized resource constraint is

$$\widehat{y}_t = \frac{\overline{c}}{\overline{y}}\widehat{c}_t + \left(1 - \frac{\overline{c}}{\overline{y}}\right)\widehat{i}_t. \tag{A5}$$

To proceed further we make two simplifying assumptions. The first assumption is that capital per effective worker is constant over time, which implies that $\left(\widehat{k}_t - \widehat{u}_t - \widehat{l}_t\right) = 0$. The second assumption is that investment is driven solely by an accelerator mechanism, i.e., that $\widehat{i}_t = \widehat{y}_t$. Combining equation (A3) with (a log-linearized) equation (34) and exploiting these two simplifying assumptions, the expression for real marginal costs becomes

$$\widehat{mc}_t = \chi\widehat{y}_t - (1+\chi)\,u_t - \widehat{\lambda}_t. \tag{A6}$$

Now log-linearizing equation (33) gives

$$\widehat{\lambda}_t = -\frac{\sigma\left(1+\gamma^2\beta\mathrm{D}\right)}{(1-\gamma)(1-\gamma\beta\mathrm{D})}\widehat{c}_t + \frac{\sigma\gamma}{(1-\gamma)(1-\gamma\beta\mathrm{D})}\left(\widehat{c}_{t-1} + \beta\mathrm{D}\mathrm{E}_t\widehat{c}_{t+1}\right) + \frac{1}{(1-\gamma\beta\mathrm{D})}g_t, \tag{A7}$$

implying that real marginal costs equal

$$\begin{aligned}\widehat{mc}_t \;=\; &\left[\chi + \frac{\sigma\left(1+\gamma^2\beta\mathrm{D}\right)}{(1-\gamma)(1-\gamma\beta\mathrm{D})}\right]\widehat{c}_t - \frac{\sigma\gamma}{(1-\gamma)(1-\gamma\beta\mathrm{D})}\left(\widehat{c}_{t-1} + \beta\mathrm{D}\mathrm{E}_t\widehat{c}_{t+1}\right) \\ &-(1+\chi)\,u_t - \frac{1}{(1-\gamma\beta\mathrm{D})}g_t, \end{aligned} \tag{A8}$$

which is equation (39) in the text.

References

[1] Altig, D., Christiano, L., Eichenbaum, M., and J. Linde, (2004), "Firm-Specific Capital, Nominal Rigidities, and the Business Cycle," Northwestern University, manuscript.

[2] Abel, A., (1990), "Asset Prices under Habit Formation and Catching up with the Joneses," *American Economic Review* (Papers and Proceedings), 80, 2, pp38-42.

[3] Amato, J., and T. Laubach, (2003), "Estimation and Control of an Optimization-Based Model with Sticky Prices and Wages," *Journal of Economic Dynamics and Control*, 27, pp1181-1215.

[4] Bils, M., and P. Klenow, (2004), "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112, pp947-985.

[5] Blinder, A., (1994), "On Sticky Prices: Academic Theories Meet the Real World," in Mankiw, N. G. (ed), *Monetary Policy*, University of Chicago Press, Chicago, pp117-150.

[6] Boldrin, M., Christiano, L., and J. Fisher, (2001), "Habit Persistence, Asset Returns, and the Business Cycle," *American Economic Review*, 91, 1, pp149-166.

[7] Calvo, G., (1983), "Staggered Contracts in a Utility-Maximising Framework," *Journal of Monetary Economics*, 12, pp383-398.

[8] Carlton, D., (1986), "The Rigidity of Prices," *American Economic Review*, 76, 4, pp637-658.

[9] Carroll, C., Overland, J., and D. Weil, (2000), "Saving and Growth with Habit Formation," *American Economic Review*, 90, 3, pp341-354.

[10] Cecchetti, S., (1986), "The Frequency of Price Adjustment: A Study of the Newsstand Prices of Magazines," *Journal of Econometrics*, 31, pp255-274.

[11] Cho, S., and A. Moreno, (2005), "A Small-Sample Study of the New Keynesian Macro Model," *Journal of Money, Credit, and Banking*, forthcoming.

[12] Christiano, L., Eichenbaum, M., and C. Evans, (2005), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113, pp1-45.

[13] Clarida, R., Galí, J., and M. Gertler, (1999), "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 37, 4, pp1661-1707.

[14] Clarida, R., Galí, J., and M. Gertler, (2000), "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, February, pp147-180.

[15] Dennis, R., (2006), "The Policy Preferences of the U.S. Federal Reserve," *Journal of Applied Econometrics*, 21, pp55-77.

[16] Duffy, J., and P. McNelis, (2001), "Approximating and Simulating the Stochastic Growth Model: Parameterized Expectations, Neural Networks, and the Genetic Algorithm," *Journal of Economic Dynamics and Control*, 25, pp1273-1303.

[17] Eichenbaum, M., and J. Fisher, (2004), "Evaluating the Calvo Model of Sticky Prices," National Bureau of Economic Research Working Paper #10617.

[18] Erceg, C., Henderson, D., and A. Levin, (2000), "Optimal Monetary Policy with Staggered Wage and Price Contracts," *Journal of Monetary Economics*, 46, pp281-313.

[19] Estrella, A., and J. Fuhrer, (2002), "Dynamic Inconsistencies: Counterfactual Implications of a Class of Rational-Expectations Models," *American Economic Review*, 92, 4, pp1013-1028.

[20] Favero, C., and R. Rovelli, (2003), "Macroeconomic Stability and the Preferences of the Fed. A Formal Analysis, 1961-98," *Journal of Money, Credit, and Banking*, 35, pp545–556.

[21] Fuhrer, J., (2000), "Optimal Monetary Policy in a Model with Habit Formation," *American Economic Review*, 90, 3, pp367-390.

[22] Galí, J., and M. Gertler, (1999), "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics*, 44, pp195-222.

[23] Galí, J., Gertler, M., and D. López-Salido, (2001), "European Inflation Dynamics," *European Economic Review*, 45, pp1237-1270.

[24] Gelfand, A., and D. Dey, (1994), "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society Series B*, 56, pp501-514.

[25] Gelman, A., (1995), "Inference and Monitoring Convergence," in Gilks, W., Richardson, S., and D. Spiegelhalter, (eds), *Practical Markov Chain Monte Carlo*, Chapman and Hall Press, London, pp131-143.

[26] Geweke, G., (1992), "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," in Bernardo, J., Dawid, A., and A. Smith (eds), *Bayesian Statistics*, Clarendon Press, Oxford, 4, pp641-649.

[27] Geweke, G., (1999), "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication," *Econometric Reviews*, 18, pp1-126.

[28] Giannoni, M., and M. Woodford, (2003), "How Forward-Looking is Optimal Monetary Policy?" *Journal of Money, Credit, and Banking*, 35, 6, pp1425-1469.

[29] Jung, Y., and T. Yun, (2005), "Implications of Inventory Models for the New Keynesian Phillips Curve," Kyunghee University, manuscript.

[30] Kashyap, A., (1995), "Sticky Prices: New Evidence from Retail Catalogs," *Quarterly Journal of Economics*, February, pp245-274.

[31] Kim, J., (2000), "Constructing and Estimating a Realistic Optimizing Model of Monetary Policy," *Journal of Monetary Economics*, 45, pp329-359.

[32] Laubach, T., and J. Williams, (2003), "Measuring the Natural Rate of Interest," *Review of Economics and Statistics*, 85, 4, pp1063-1070.

[33] Levin, A., Onatski, A., Williams, J., and N. Williams, (2005), "Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models," National Bureau of Economic Research Working Paper #11523.

[34] Liu, Z., and L. Phaneuf, (2005), "Technology Shocks and labor Market Dynamics: Some Evidence and Theory," Emory University, manuscript.

[35] Lucas, R., (1976), "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy*, 1, pp19-46.

[36] Mankiw, N. G., (1985), "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly," *Quarterly Journal of Economics*, May, pp529-537.

[37] Mankiw, N. G., and R. Reis, (2002), "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117, 4, pp1295–1328.

[38] McCallum, B., and E. Nelson, (1999), "Nominal Income Targeting in an Open-Economy Optimizing Model," *Journal of Monetary Economics*, 43, pp553-578.

[39] Ravenna, F., and C. Walsh, (2006), "Optimal Monetary Policy and the Cost Channel," *Journal of Monetary Economics*, 53, 2, pp199-216.

[40] Roberts, J., (1995), "New Keynesian Economics and the Phillips Curve" *Journal of Money, Credit, and Banking*, 27, 4, pp975-984.

[41] Rotemberg, J., (1982), "Sticky Prices in the United States," *Journal of Political Economy*, 60, pp1187-1211.

[42] Rotemberg, J., and M. Woodford, (1997), "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," in Bernanke, B., and J. Rotemberg, (eds) *NBER Macroeconomics Annual 1997*, MIT Press, Cambridge.

[43] Rouwenhorst, K. G., (1995), "Asset Pricing Implications of Equilibrium Business Cycle Models," in Cooley, T., (ed), *Frontiers of Business Cycle Research*, Princeton University Press, Princeton, New Jersey.

[44] Rudd, J., and C. Whelan, (2006), "Can Rational Expectations Sticky-Price Models Explain Inflation Dynamics?" *American Economic Review*, 96, 1, pp303-320.

[45] Sbordone, A., (2002), "Prices and Unit Labor Costs: A New Test of Price Stickiness," *Journal of Monetary Economics*, 49, pp265-292.

[46] Schorfheide, F., (2000), "Loss Function-Based Evaluation of DSGE Models," *Journal of Applied Econometrics*, 15, pp645-670.

[47] Sims, C., (2002), "Implications of Rational Inattention," Princeton University, manuscript.

[48] Smets, F., (2003), "Maintaining Price Stability: How Long is the Medium Term," *Journal of Monetary Economics*, 50, pp1293-1309.

[49] Smets, F., and R. Wouters, (2003), "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1, 5, pp1123-1175.

[50] Taylor, J., (1980), "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy*, 88, pp1-24.

[51] Walsh, C., (2003), "Speed Limit Policies: The Output Gap and Optimal Monetary Policies," *American Economic Review*, 93, 1, pp265-278.

[52] Woodford, M., (2003), *Interest and Prices*, Princeton University Press, Princeton, New Jersey.

[53] Zbaraki, M., Ritson, M., Levy, D., Dutta, S., and M. Bergin, (2004), "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets," *Review of Economics and Statistics*, 86, 2, pp514-533.