Testing the Calvo model of sticky prices

Martin Eichenbaum and Jonas D. M. Fisher

Introduction and summary

A classic question in macroeconomics is: Why do changes in monetary policy affect aggregate economic activity? The answer to this question is of central importance to monetary policymakers. This is because policymakers require a convincing model of the monetary transmission mechanism in order to evaluate the consequences of alternative choices.

A key assumption in many models used to assess the effects of monetary policy is that the nominal prices of goods are "sticky." By this we mean that firms do not change their prices each period in response to the different shocks impacting on their environment. Models embodying this assumption typically have the property that policy actions that raise the money supply and lower short-term interest rates lead to expansions in aggregate economic activity. These types of models are increasingly being used by central banks around the world to help guide policymakers in setting monetary policy.

Different approaches to modeling sticky prices have been adopted in the literature. In one class of models, referred to as *time-dependent models*, the number of firms that change prices in any given period is specified exogenously. Classic models of this sort were developed by Taylor (1980) and Calvo (1983). Modern variants are now central elements of a large class of models.[1] A key feature of Calvo–Taylor pricing models is that forward-looking firms understand they will only periodically reoptimize prices. So, firms front load higher future expected real marginal costs into their current prices. They do this because they may not be able to raise prices when the higher marginal costs materialize. Similarly, to avoid declines in their relative prices, firms front load future inflation into the prices that they set. As typically formulated, these models often imply that deviations of economy-wide inflation from its long-run value depend primarily on current and expected changes in firms' real marginal costs.

In a different class of models, often referred to as *state-dependent pricing models*, the number of firms changing prices in any given period is determined endogenously. Dotsey, King, and Wolman (1999) model this phenomenon by assuming that firms pay a fixed cost when they change their price. In contrast, Burstein (2002) assumes that firms pay a fixed cost for changing price plans. Once they pay this cost, firms can choose not only their current price, but also a plan specifying an entire sequence of future prices. A key property of these models is that small and large changes in monetary policy have qualitatively different effects on aggregate economic activity.

While state-dependent models seem very promising (at least to us), they are substantially more difficult to work with than time-dependent models. In addition, the two classes of models generate similar results for many policy experiments that are relevant in moderate inflation economies such as the U.S.[2] For these reasons, modern variants of Taylor and Calvo models continue to play a central role in the analysis of monetary policy.

This article addresses the question: Are time-dependent models good models in an empirical sense? For concreteness, we focus on the Calvo sticky pricing model. In principle, there are a variety of ways to test this model. For example, one could embed it in a fully specified general equilibrium model of the economy. This would involve, among other things, modeling household labor and consumption decisions, credit markets, fiscal policy, and monetary policy. If in addition, one specified the nature of all the shocks impacting on the

Martin Eichenbaum is a professor of economics at Northwestern University, a research fellow at the National Bureau of Economic Research, and a consultant to the Federal Reserve Bank of Chicago. Jonas D. M. Fisher is a senior economist and economic advisor at the Federal Reserve Bank of Chicago.

economy, one could estimate and test the model using a variety of statistical methods like maximum likelihood.[3] Another strategy would be to assess the model's predictions for a particular shock, such as a disturbance to monetary policy or a shock to technology.[4]

Here, we focus on tests of the model using the econometric strategy pioneered by Hansen (1982) and Hansen and Singleton (1982) and applied to the Calvo model by Gali and Gertler (1999) and Eichenbaum and Fisher (2003). The idea is to exploit the fact that in any model incorporating Calvo pricing, certain restrictions must hold. One can test these restrictions, without making assumptions about other aspects of the economy. Of course, in the end, we need a fully specified model of the economy within which to assess the consequences of alternative policy. The approach that we discuss here has the advantage of focusing on the empirical plausibility of one key building block that could be an element of many models.

Our analysis proceeds as follows. In the next section, we summarize the Calvo model. Standard versions of the model assume that when firms reoptimize their price plans, the new plan takes effect immediately. As in Eichenbaum and Fisher (2003), we allow for the possibility that when firms reoptimize their price plans at time t, the new plan only goes into effect at time $t + \tau$, where $\tau \geq 0$ and a time period corresponds to a quarter. The standard Calvo model corresponds to the assumption that τ is equal to zero. By varying τ, we can vary the information set that firms have at their disposal when making new price decisions.

The following section discusses an econometric strategy for estimating and testing the model, taking into account the possibility of measurement error in the variables of interest, particularly inflation. Then, we discuss the four measures of inflation that we use in our empirical analysis, as well as our measure of marginal cost. Finally, we present our results, drawing heavily from Eichenbaum and Fisher (2003).

Our main findings can be summarized as follows. First, using postwar U.S. time-series data, we find strong evidence against the standard Calvo model ($\tau = 0$). This is true regardless of whether we allow for a structural break in monetary policy in the early 1980s, which a number of researchers argue occurred with the onset of the Volker–Greenspan era. Second, once we allow for a lag between the time that firms reoptimize and the time that they implement their new plans ($\tau > 0$), the model is no longer rejected. Third, allowing for measurement error in inflation also overturns the rejection of the standard Calvo model ($\tau = 0$). For reasons we discuss below, we are more comfortable with the second of the two resolutions.

Consider first the possibility that τ exceeds zero. If we use the full post-1959 sample period, we require that $\tau = 2$, that is, firms set prices two quarters in advance, to avoid rejecting the model. Frankly, we are skeptical that there is a six-month delay between when firms reoptimize their price plans and when they actually implement the new plan. So we do not view this as a plausible way of overturning the evidence against the standard Calvo model. Fortunately, once we allow for a break in monetary policy, the required value of τ arguably drops to one, which seems more reasonable on a priori grounds.

Turning to the other resolution, we find that even with independently and identically distributed (iid) classical measurement error, there is only marginal evidence against the standard Calvo model using the whole sample period. Once we allow for a break in monetary policy, we find virtually no evidence against the model. In addition, we cannot reject the null hypothesis that firms reoptimize prices, on average, once a year. This seems reasonable in light of the assumptions usually made in the literature. Of course a key question is: How large is the measurement error required to overturn the rejection of the model? We quantify the size of the measurement error using a variety of statistics. Our own view is that for the gross domestic product (GDP), Consumer Price Index (CPI), and personal consumption expenditures (PCE) deflator-based measures of inflation, the size of the required measurement error is reasonable, according to a variety of metrics documented in the article. We tentatively conclude that there is little evidence against the restrictions implied by the Calvo sticky price model.

The Calvo model of sticky prices

As discussed in the introduction, there are a variety of ways to model nominal rigidities in goods prices. Here we discuss the model of price setting associated with Calvo (1983). Since our objective is to derive the testable implications of this model per se, we do not embed it within a general equilibrium framework.

At time t, a final good, Y_t, is produced by a perfectly competitive firm. It does so by combining a continuum of intermediate goods, indexed by $j \in [0,1]$ using a constant returns to scale technology. We let P_t and P_{jt} denote the time t price of the final and intermediate good j, respectively. Profit maximization implies that the demand for intermediate good j is a decreasing function of the relative price of that good and an increasing function of aggregate output, Y_t.

The intermediate good $j \in [0,1]$ is produced by a monopolist that uses the following technology:

1) $Y_{jt} = A_t k_{jt}^{\alpha} L_{jt}^{1-\alpha}$,

where $0 < \alpha < 1$. Here, L_{jt} and k_{jt} denote time t labor and capital services used to produce the jth intermediate good, respectively. Intermediate firms rent capital and labor in perfectly competitive factor markets. The variable A_t denotes possible stochastic disturbances to technology.

Profits are distributed to the firms' owners at the end of each period. Let s_t denote the representative firm's real marginal cost, that is, the change in an optimizing firm's real total cost associated with increasing output by one unit.[5] Given our assumptions, marginal costs depend on the parameter α and factor prices, which the firm takes as given. The firm's time t profits are:

$$\left[\frac{P_{jt}}{P_t} - s_t\right] P_t Y_{jt},$$

where P_{jt} is firm j's price.

We assume that firms set prices according to a variant of the mechanism spelled out in Calvo (1983). In each period, a firm faces a constant probability, $1 - \theta$, of being able to reoptimize its nominal price. So, on average, a firm reoptimizes its price every $(1 - \theta)^{-1}$ periods. For example, if a period is one quarter and θ is 0.75, the firm reoptimizes on average once a year. We assume for simplicity that the firm's ability to reoptimize its price is independent across firms and time. For now we leave open the issue of what information set the firm has when it resets its price.

A standard assumption in the literature is that if the firm does not reoptimize its price, it updates its price according to the rule:

2) $P_{jt} = \bar{\pi} P_{jt-1}$,

where $\bar{\pi}$ is the long-run average gross rate of inflation (see, for example, Erceg, Henderson, and Levin, 2000, and Yun, 1996).[6]

As in Christiano, Eichenbaum, and Evans (2001), we interpret the Calvo price-setting mechanism as capturing firms' response to various costs of changing prices. The basic idea is that in the presence of these costs, firms fully optimize prices only periodically and follow simple rules for changing their prices at other times.

Let $\tilde{P}_t$ denote the value of P_{jt} set by a firm that can reoptimize at time t. Our notation does not allow $\tilde{P}_t$ to depend on j. We do this because, in models like ours, all firms that can reoptimize their price at time t choose the same price (see Woodford, 1996, and Yun, 1996). The firm chooses $\tilde{P}_t$ to maximize the expected present value of profits. We suppose that the firm sets $\tilde{P}_t$ on the basis of the information that it has at time $t - \tau$. When $\tau = 0$, the firm sees the realization of all time t variables when resetting its price. We refer to this version of the model as the standard Calvo model. The assumption that $\tau > 0$ is similar in spirit to the model in Mankiw and Reiss (2002), who think of firms as having flexible prices but "sticky" information sets. Alternatively one can imagine that resetting prices is a costly time consuming event for managers, so that prices must be set τ periods in advance. Given our assumptions, if the firm can reset its prices every period, then it would set its price, $\tilde{P}_t$, equal to a markup over the expected marginal cost conditional on information at $t - \tau$.

Log linearizing the first-order condition of the firm around the relevant steady state values, we obtain:

3) $$\hat{\tilde{p}}_t = E_{t-\tau}\left[\hat{s}_t + \sum_{l=1}^{\infty}(\beta\theta)^l\left(\hat{s}_{t+l} - \hat{s}_{t+l-1}\right) + \sum_{l=1}^{\infty}(\beta\theta)^l \hat{\pi}_{t+l}\right].$$

Here, $E_{t-\tau}$ denotes the conditional expectations operator. For example, $E_{t-\tau}\hat{s}_t$ denotes agents' expectations of $\hat{s}_t$ conditional on the information that they have at time $t - \tau$. In addition, $\tilde{p}_t = \tilde{P}_t / P_t$, and a hat over a variable indicates the percent deviation from its steady state value.

As noted by Christiano, Eichenbaum, and Evans (2001), several features of equation 3 are worth emphasizing. First, if inflation is expected to be at its steady state level and real marginal costs are expected to remain constant after time t, then the firm sets $\hat{\tilde{p}}_t = E_{t-\tau}\hat{s}_t$. Second, suppose the firm expects real marginal costs to be higher in the future than at time t. Anticipating those future marginal costs, the firm sets $\hat{\tilde{p}}_t$ higher than $E_{t-\tau}\hat{s}_t$. It does so because it understands that it may not be able to raise its price when those higher marginal costs materialize. Third, suppose firms expect inflation in the future to exceed its steady state level. To avoid a decline in its relative price, the firm incorporates expected future changes in the inflation rate into $\hat{\tilde{p}}_t$.

It follows from well-known results in the literature that the aggregate price level can be expressed as:[7]

4) $$P_t = \left[(1-\theta)\left(\tilde{P}_t\right)^{\frac{1}{1-\lambda}} + \theta\left(\bar{\pi}P_{t-1}\right)^{\frac{1}{1-\lambda}}\right]^{1-\lambda},$$

where $\lambda \in [1,\infty)$ is a parameter that controls the degree of substitutability of intermediate goods in the production of the final good. Log linearizing this relation and using it in conjunction with equation 3 implies that inflation satisfies:[8]

$$5)\quad \theta\hat{\pi}_t = \beta\theta E_{t-\tau}\hat{\pi}_{t+1} + (1-\beta\theta)(1-\theta)E_{t-\tau}\hat{s}_t.$$

While equation 5 is the focus of our empirical analysis, it is useful to note that it implies:

$$6)\quad \hat{\pi}_t = \frac{(1-\beta\theta)(1-\theta)}{\theta}E_{t-\tau}\sum_{j=0}^{\infty}\beta^j\hat{s}_{t+j}.$$

Relation 6 makes clear a central prediction of the model: Deviations of inflation from its steady state depend only on firms' expectations of current and future deviations of real marginal cost from its steady state value. So for example, in the short run, the growth rate of money, interest rates, or technology shocks affects inflation only by its effect on real marginal costs. In the long run, the rate of inflation depends on the average growth rate of money.

Assessing the empirical plausibility of the model

Here, we discuss the limited information strategy for testing the Calvo sticky price model pursued by Gali and Gertler (1999) and Eichenbaum and Fisher (2003), among others. The basic idea is to focus on the testable restrictions of the Calvo pricing model, while leaving unspecified other aspects of the economy.

To derive the testable implications of the Calvo model, it is convenient to define the random variable.

$$\psi_{t+1} = \left[\theta\hat{\pi}_t - \beta\theta\hat{\pi}_{t+1} - (1-\beta\theta)(1-\theta)\hat{s}_t\right].$$

Note that agents that reoptimize their price do so on the basis of their time $t-\tau$ information. The other prices that affect the time t inflation rate were already set on the basis of information before time $t-\tau$. This means that inflation is predetermined at time $t-\tau$. In principle, there are a variety of ways to test this assumption. For example, we could test whether any variable dated between time $t-\tau$ and t has explanatory power for time t inflation.

Here, we test this implication indirectly. Since $\hat{\pi}_t$ is in agents' time $t-\tau$ information set, equation 5 can be written as:

$$7)\quad E_{t-\tau}\psi_{t+1} = 0.$$

Relation 7 implies that the error agents make in forecasting the value of ψ_{t+1} when they reoptimize prices at time $t-\tau$ is uncorrelated with the information that they have at their disposal. Suppose that the $k \times 1$ vector of variables $X_{t-\tau}$ is in agents' time $t-\tau$ information set. Below, we refer to these variables as *instruments*. Then relation 7 implies the system of k equations:

$$8)\quad E_{t-\tau}\psi_{t+1}X_{t-\tau} = 0.$$

This in turn implies that the unconditional covariance between ψ_{t+1} and $X_{t-\tau}$ is equal to zero:

$$9)\quad E\psi_{t+1}X_{t-\tau} = 0.$$

Relation 9 provides us with a way to estimate the parameters of the model. Moreover, if the dimension of $X_{t-\tau}$ is greater than the number of parameters to be estimated, we can use these restrictions to test the model. To discuss our procedures for doing this, it is useful to recognize the dependence of ψ_{t+1} on the unknown value of (θ,β) by writing equation 9 as

$$10)\quad E\,[\psi_{t+1}(\theta,\beta)\,X_{t-\tau}] = 0.$$

Hansen (1982) provides conditions under which equation 10 can be used to consistently and efficiently estimate (θ,β) using generalized method of moments (GMM).[9] To discuss his procedure in our context, we define the vector

$$g_T(\theta,\beta) = \left(\frac{1}{T}\right)\sum_{t=1}^{T}\left[\psi_{t+1}(\theta,\beta)\,X_{t-\tau}\right].$$

Here T denotes the size of our sample. We also denote the true value of (θ,β) by (θ_0,β_0). The vector $g_T(\theta,\beta)$ is a consistent estimator of $E\,[\psi_{t+1}(\theta,\beta)\,X_{t-\tau}]$. The value of $E\,[\psi_{t+1}(\theta,\beta)\,X_{t-\tau}]$ is in general not equal to zero except at (θ_0,β_0). We estimate the parameter vector (θ_0,β_0) by choosing (θ,β) to make $g_T(\theta,\beta)$ as close to zero as possible in the sense of minimizing

$$11)\quad J_T = g_T(\theta,\beta)'W_T g_T(\theta,\beta).$$

Here, W_T is a symmetric positive definite matrix that can depend on sample information. Also the prime symbol (′) denotes the transpose operator. A given choice of W_T implies that we are choosing (θ,β) to minimize the sum of squares of k linear combinations of the elements of $g_T(\theta,\beta)$.

Hansen (1982) shows that the choice of W_T that minimizes the asymptotic covariance matrix of our estimator depends on the serial correlation properties of the error term $\psi_{t+1}(\theta,\beta)$. In Eichenbaum and Fisher (2003), we show that the exact serial correlation properties of this error term depend on the value of τ. For example, if $\tau = 0$, then our model implies that $\psi_{t+1}(\theta,\beta)$ is serially uncorrelated. For $\tau \geq 1$, then $\psi_{t+1}(\theta,\beta)$ has a moving average representation of order 1. One does not have to impose this restriction in constructing an $\tau - 1$ estimate of W_T.[10] However, as we describe below, whether one does so has an important impact, in practice, on inference.

Hansen proves that the minimized value of the GMM criterion function, J_T, is asymptotically distributed as a χ^2 random variable with degrees of freedom equal to the difference between the number of unconditional moment restrictions imposed (k) and the number of parameters being estimated. We use this fact to test the restrictions imposed by the Calvo model.

Allowing for measurement error

The previous discussion assumes that inflation and real marginal costs are measured without error. We conclude this section by reviewing the results in Eichenbaum and Fisher (2003) about how measurement error affects the analysis. The possibility of measurement error in inflation is of particular interest to us. This is because a number of authors have noted that when they include a lagged inflation term in objects like ψ_{t+1}, it enters with a significant coefficient (see, for example, Gali and Gertler, 1999, and Fuhrer and Moore, 1995). These authors have interpreted this lagged term as evidence of firms that do not have rational expectations. Measurement error can provide an alternative interpretation of these findings.

There are well-known problems involved in measuring inflation. For example, it is widely believed that official CPI-based measures of inflation are biased due to changes in product quality and the benchmark basket of goods over time (see Shapiro and Wilcox, 1996). These problems are particularly severe when measuring rates of inflation over long periods. To the extent that this bias is constant, it does not affect our analysis. However, we must modify our econometric procedures to allow for time varying measurement error. Here we discuss the implications of classical measurement error.

Suppose that the econometrician has a measure of inflation π_t^m that is related to true inflation (π_t) via the relationship

$$\pi_t^m = \pi_t + u_t.$$

We suppose that u_t has a moving average representation of order q, denoted $MA(q)$, and that u_t is uncorrelated with π_t and the other variables in agents' information set at all leads and lags. This latter assumption defines what we mean when we say that u_t is classical measurement error. We continue to assume that agents see actual inflation.

To see how these assumptions impact our econometric procedures, consider the case in which τ is equal to zero and u_t is iid ($q = 0$). The econometrician now sees ϕ_{t+1}, which is a "polluted" version of the error term, ψ_{t+1}, that is the basis of the estimation procedure. The random variables ϕ_{t+1} and ψ_{t+1} are related as follows:

$$\phi_{t+1} = \psi_{t+1} + \theta\varepsilon_t - \theta\beta\varepsilon_{t+1}.$$

While ψ_{t+1} is uncorrelated with the elements of agents' time t information set, ϕ_{t+1} is correlated with π_t^m. Accordingly, measured time t inflation is not a valid instrument, that is, it cannot be included in X_t.

The presence of iid measurement error in inflation also means that ϕ_{t+1} has an $MA(1)$ representation. This affects the nature of the restrictions that the model imposes on the weighting matrix W_T. In Eichenbaum and Fisher (2003), we show how to estimate the volatility of the error term relative to the volatility of ϕ_{t+1}, as well as the contribution of measurement error to the volatility of measured inflation. This provides us with two metrics for assessing the size of the measurement error.

We refer the reader to Eichenbaum and Fisher (2003) for a discussion of the more general case in which u_t has a higher order MA representation. For our purposes, the key result is that when u_t has an $MA(q)$ structure, then ϕ_{t+1} has an $MA(q + 1)$ representation so that one must exclude $q + 1$ lags of inflation from the list of instruments, X_t. This structure also affects the restrictions that we can impose on the weighting matrix W_T.

We conclude this section by considering the possibility that real marginal costs are measured with a classical measurement error term that has an $MA(q)$ representation. If $\tau = 0$, then ϕ_{t+1} will have an $MA(q)$ representation, and one must exclude q lags of real marginal costs from the list of instruments, X_t. Below, we abstract from this source of measurement error and refer the reader to Eichenbaum and Fisher (2003) for an analysis of this case.

Measuring inflation and real marginal cost

Next, we discuss the measures of inflation and real marginal costs that we can use in our empirical analysis.

Inflation

Many different measures of inflation are of interest to economists and policymakers. Given the abstract nature of the Calvo model, there is no obviously *right* measure to use in our empirical analysis. In light of this, we considered four measures of inflation based on four measures of the aggregate price level: 1) the GDP deflator, 2) the price deflator for the nonfarm business sector (NFB), 3) the Consumer Price Index (CPI), and 4) the price deflator for personal consumption expenditures (PCE).[11] For each price measure, we constructed a measure of quarterly inflation over the period 1959–2001. In our empirical work, we measure $\hat{\pi}_T$ as the difference between actual time *t* inflation and the sample average of inflation.

Our different inflation measures are displayed in figure 1. As we can see, they behave in a similar manner over long periods of time. Inflation was low in the decade of the 1960s, then began a rapid rise with one peak in the early 1970s and another in the late 1970s. Thereafter, the different measures begin a long decline to very low levels by 2001. However, there are important differences between them over shorter periods. Since the Calvo model purports to account for movements in inflation over short periods of time, it is important to assess the robustness of our results using the different measures of inflation.

Real marginal costs

In our model, real marginal costs are given by the real product wage divided by the marginal product of labor. Given the production function we assumed in equation 1, this implies that real marginal cost is proportional to labor's share in national income, $W_tL_t/(P_tY_t)$, where W_t is the nominal wage. In practice, we measure W_tL_t as nominal labor compensation in the nonfarm business sector and we measure P_tY_t as nominal output of the nonfarm business sector. The variable $\hat{s}_t$ is then measured as the difference between the log of our measure of labor's share and its mean. This is a standard measure of $\hat{s}_t$, which has been used by Gali and Gertler (1999) and Sbordone (2001).

Rotemberg and Woodford (1999) discuss possible corrections to this measure that are appropriate for different assumptions about technology. These include corrections to take into account a non-constant elasticity of factor substitution between capital and labor and the presence of overhead costs and labor adjustment costs. In Eichenbaum and Fisher (2003), we discuss results for these alternative measures of marginal costs. In addition we allow for the possibility that firms require working capital to finance payments to variable factors of production. We argue in that paper that these corrections do not affect the qualitative nature of the results discussed below.

Panel A of figure 2 displays the log of our measure of real marginal cost and inflation measured using the GDP deflator for the sample 1959–2001. Notice that from the mid-1960s on, these two time series co-move positively. The bottom panel of figure 2 displays the dynamic correlations between real marginal cost at date *t* and inflation at date $t - k$, $k = -4, -3, \ldots, 4$. Clearly, inflation is positively correlated with past and future marginal costs.

Empirical results

Now, we present our empirical results.

The standard Calvo model

We begin by analyzing results based on the standard Calvo model, by which we mean the model described above with $\tau = 0$. In addition, we initially abstract from measurement error in inflation. We consider two specifications of the instrument vector X_t. Let Z_t denote the six dimensional vector consisting of the time *t* value of real marginal cost, quadratically detrended real GDP, inflation, the growth rate of an index of commodity prices, the spread between the annual interest rate on the ten-year Treasury bond and three-month Treasury bill, and the growth rate of nominal wages in the nonfarm business sector. This corresponds to the basic set of instruments used in Gali and Gertler (1999).

In our first specification, X_t is given by

$$X_t^1 = \{1, Z_{t-j}, j = 0,1,2,3\}'.$$

As we discuss below, there are reasons to think that such a large set of instruments leads to misleading inference about the plausibility of the overidentifying restrictions implied by the model. With this in mind, we consider a second set of instruments given by

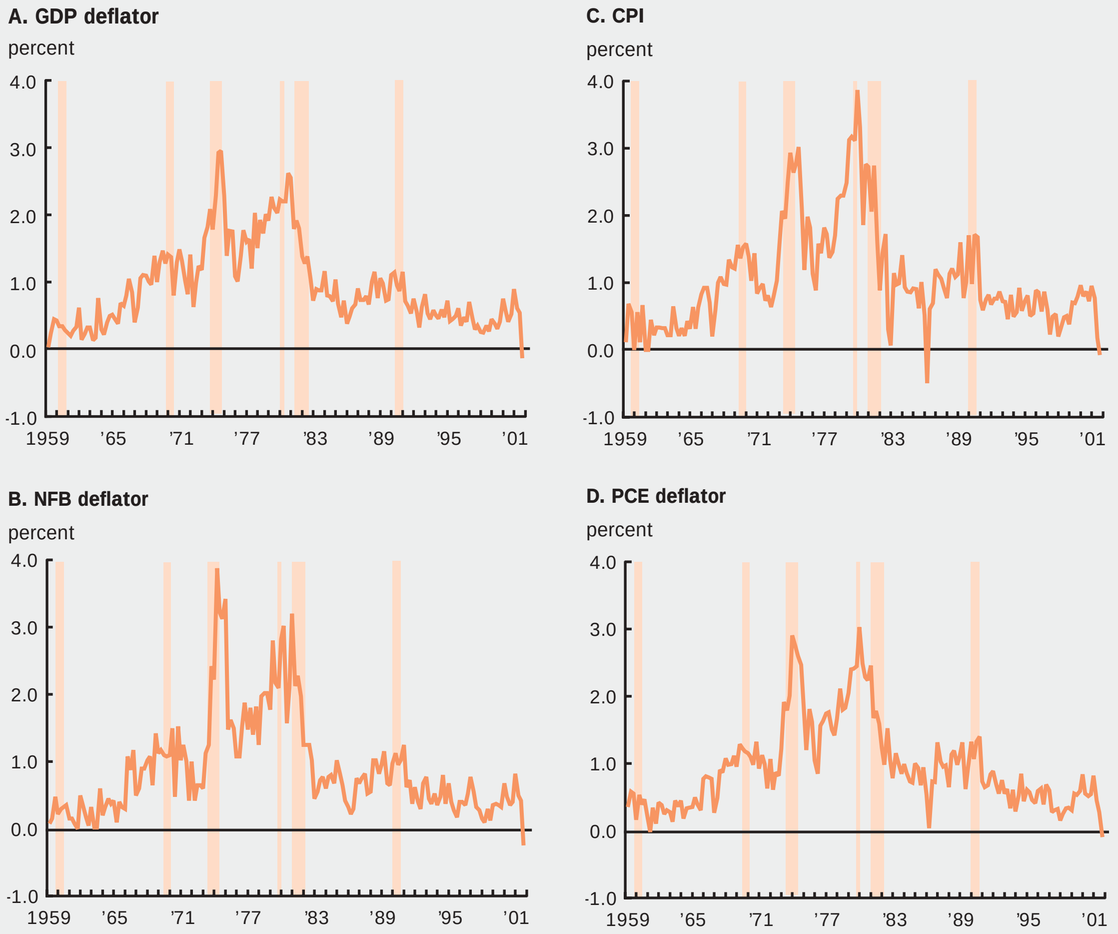
$$X_t^2 = \{1, Z_t, \psi_t\}'.$$

In Eichenbaum and Fisher (2003), we report that it is difficult to estimate β with great precision across the different specifications considered. Here, we summarize the results based on the assumption that $\beta = 0.99$.

Panel A of table 1, based on Eichenbaum and Fisher (2003), summarizes results when the standard Calvo model is estimated using the instrument vector X_t^1. We report our estimates of the parameter θ (standard

FIGURE 1

Four measures of inflation



Note: Shaded areas indicate official National Bureau of Economic Research recession periods.
Source: Authors' calculations based upon data from Haver Analytics.

error in parentheses) and the J_T statistic (p-value in brackets). The label L refers to the maximal degree of serial correlation that we allow for when estimating the weighting matrix W_T. We consider two values for L: 1) $L = 0$, which corresponds to the degree of serial correlation in ψ_{t+1} implied by this model, and 2) $L = 12$, the value used by Gali and Gertler (1999). Both values of L are admissible. But, by setting L to zero, we are imposing all of the restrictions implied by the model. This may lead to greater efficiency of our estimator and more power in our test of the overidentifying restrictions.

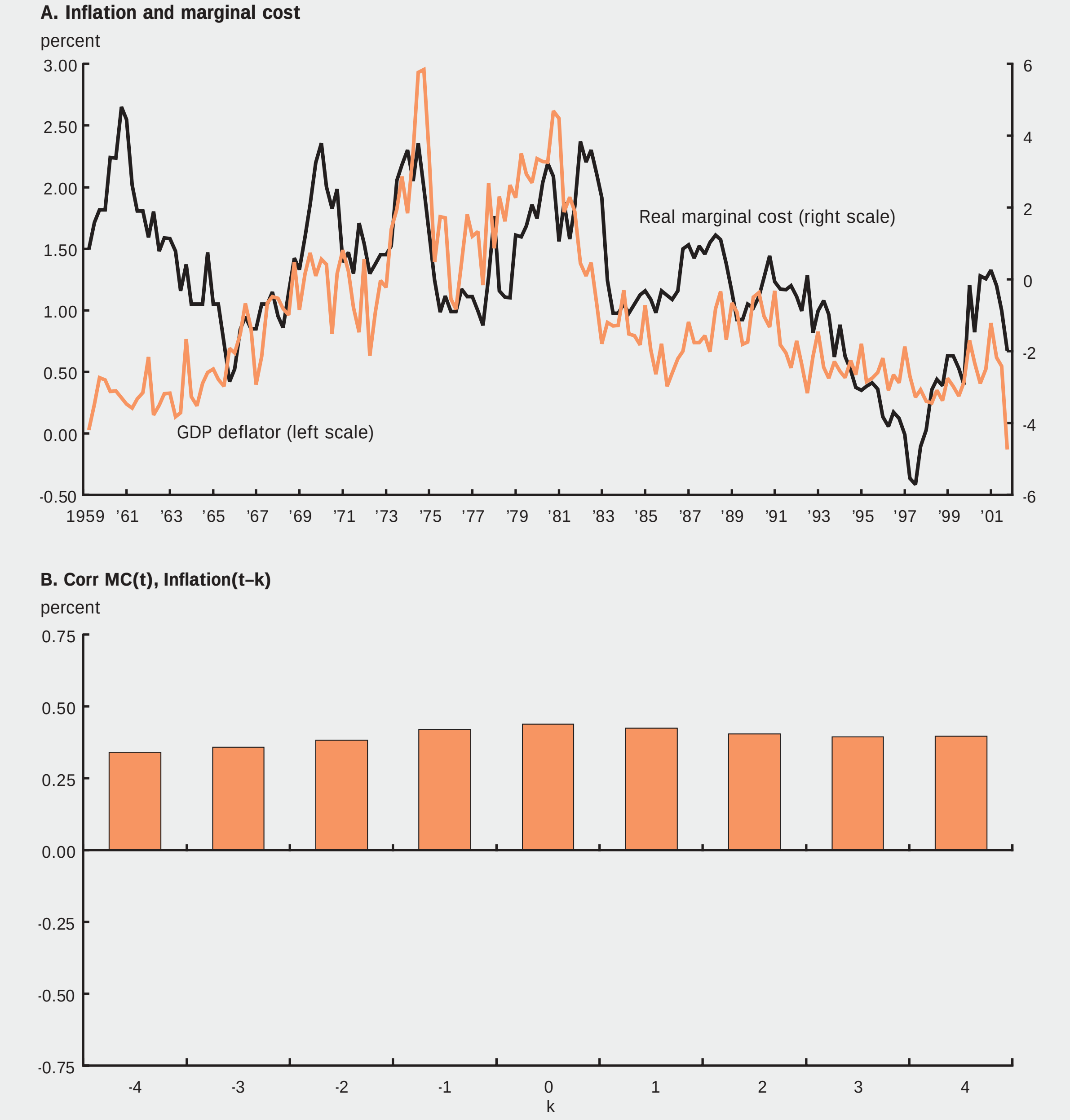
From table 1 we see that the parameter θ is estimated with relatively small standard errors. In addition, the point estimate itself is reasonably robust across the different inflation measures and the two values of L. The point estimates range from a low of 0.84 to a high of 0.91. This implies that, on average, firms wait between six and 11 quarters before reoptimizing their prices.

We hesitate to attribute too much importance to these point estimates. It is true that when $L = 12$ there is virtually no evidence against the model, at least based on the J_T statistic. This is consistent with results from Gali and Gertler (1999). However, when we set $L = 0$, the model is strongly rejected for three of the four inflation measures. In particular, the p-values for the non-CPI based inflation measures are well below 1 percent. Even in the CPI case, the p-value is 2 percent. Evidently, imposing all of the relevant restrictions implied by the model on the weighting matrix has an important impact on inference.

Panel B reports results based on the instrument vector X_t^2. A number of results are worth noting. First, our point estimates of θ are similar to those in panel A. Second, comparing the J_T statistics for $L = 12$ across

FIGURE 2

GDP deflator inflation and real marginal cost



Source: Authors' calculations based upon data from Haver Analytics.

the two instrument vectors, we find that there is more evidence against the model with the smaller list of instruments. However, with the instrument list X_t^2, the model is still not rejected at conventional significance levels for any inflation measure. Third, the model is decisively rejected when we set $L = 0$. Regardless of which inflation measure we use, the p-value of the J_T statistics is virtually zero. In light of these results, in Eichenbaum and Fisher (2003), we stress the importance of working with the smaller instrument set and imposing all of the model relevant restrictions on the weighting matrix W_T. For the rest of this article, we confine ourselves to results generated in this way.

An important maintained assumption of the previous results is that it is appropriate to use the entire sample period to estimate the model. In fact numerous observers

TABLE 1

Estimates of the standard model, 1959:Q1–2001:Q4

Inflation measure	$L = 0$		$L = 12$	
	θ	J_T	θ	J_T
A. Instruments: $\{1, Z_t, \ldots, Z_{t-3}\}'$				
GDP deflator	0.89 (0.03)	49.4 [0.001]	0.91 (0.02)	13.2 [0.95]
NFB deflator	0.86 (0.03)	41.1 [0.01]	0.86 (0.02)	12.8 [0.96]
CPI	0.88 (0.05)	38.9 [0.02]	0.86 (0.02)	12.6 [0.96]
PCE deflator	0.87 (0.02)	44.8 [0.004]	0.88 (0.02)	12.8 [0.96]
B. Instruments: $\{1, Z_t, \psi_{t-1}\}'$				
GDP deflator	0.90 (0.05)	28.2 [9e–5]	0.91 (0.03)	10.3 [0.11]
NFB deflator	0.84 (0.03)	30.6 [3e–5]	0.85 (0.03)	8.8 [0.18]
CPI	0.88 (0.06)	30.1 [4e–5]	0.87 (0.03)	10.1 [0.12]
PCE deflator	0.87 (0.04)	36.9 [2e–6]	0.89 (0.03)	11.5 [0.07]

Notes: The J_T statistics are distributed as χ^2 random variables with six and 23 degrees of freedom in panels A and B, respectively. Standard errors are in parentheses. P-values are in brackets. GDP is gross domestic product; NFB is nonfarm business; CPI is Consumer Price Index; and PCE is personal consumption expenditures.
Source: Authors' calculations based upon data from Haver Analytics.

have argued that there was an important change in the nature of monetary policy with the advent of the Volker disinflation in the early 1980s. Moreover, it is often argued that the Fed's operating procedures were different in the early 1980s than in the post-1982 period. Accordingly, we reestimated the standard Calvo model over two distinct subsamples: 1959:Q1–79:Q2 and 1982:Q3–2001:Q4.

Table 2 reports the subsample results (here $L = 0$ and the instrument vector is X_t^2). For the first sample period, there is strong evidence against the model for at least two measures of inflation. In particular, the p-values of the J_T statistic obtained using the NFB and PCE deflators are virtually zero. There is somewhat less evidence against the model when we use the GDP and CPI deflator based measures of inflation. Here the p-values are 0.04 and 0.02, respectively. In these cases, the point estimate of θ is 0.84. Taking sampling uncertainty into account, we would not reject the null hypothesis that, on average, firms wait about a year before reoptimizing their prices.

Turning to the second subsample, we see that there is substantially less evidence against the model. Here, the p-values associated with the J_T statistics obtained using the NFB, CPI, and PCE deflators are 0.06, 0.22, and 0.10, respectively. The only case in which we can reject the model at the 1 percent level of significance is when we use the GDP deflator to measure inflation. Interestingly, our point estimates of θ for specifications that are not strongly rejected are similar across subsamples. Again, taking sampling uncertainty into account, we would not reject the null that, on average, firms wait about a year before reoptimizing their prices.

Alternative timing assumptions

We now consider the results of estimating the model assuming $\tau = 1$ or $\tau = 2$. For these cases, our instrument list is given by X_{t-1}^2 and X_{t-2}^2, respectively. Panels A and B of table 3, based on Eichenbaum and Fisher (2003), summarize results for these cases. We begin by considering the full sample results. Two results are worth noting here. First, the point estimates of θ are similar across the different values of τ considered, including $\tau = 0$ discussed above. Second, when $\tau = 1$, the model's overidentifying restrictions are decisively

TABLE 2

Subsample estimates of standard model

Inflation measure	1959:Q1–79:Q2		1982:Q3–2001:Q4	
	θ	J_T	θ	J_T
GDP deflator	0.84 (0.04)	13.4 [0.04]	0.86 (0.04)	17.0 [0.009]
NFB deflator	0.74 (0.03)	21.7 [0.001]	0.86 (0.04)	12.2 [0.06]
CPI	0.84 (0.06)	14.8 [0.02]	0.86 (0.05)	8.21 [0.22]
PCE deflator	0.83 (0.03)	22.6 [9e–4]	0.85 (0.04)	10.4 [0.10]

Notes: The J_T statistics are distributed as χ^2 random variables with six degrees of freedom. Standard errors are in parentheses. P-values are in brackets. GDP is gross domestic product; NFB is nonfarm business; CPI is Consumer Price Index; and PCE is personal consumption expenditures.
Source: Authors' calculations based upon data from Haver Analytics.

TABLE 3

Alternative timing assumptions

Inflation measure	Full sample θ	Full sample J_T	1959:Q1–79:Q2 θ	1959:Q1–79:Q2 J_T	1982:Q3–2001:Q4 θ	1982:Q3–2001:Q4 J_T
A. Prices chosen one period in advance						
GDP deflator	0.87 (0.03)	22.8 [8e–4]	0.78 (0.04)	13.3 [0.04]	0.82 (0.02)	15.1 [0.02]
NFB deflator	0.82 (0.04)	28.0 [9e–5]	0.66 (0.03)	22.1 [0.001]	0.81 (0.03)	12.7 [0.04]
CPI	0.86 (0.04)	18.4 [0.005]	0.72 (0.03)	15.7 [0.02]	0.83 (0.04)	6.37 [0.38]
PCE deflator	0.83 (0.02)	22.9 [8e–4]	0.75 (0.03)	12.9 [0.04]	0.82 (0.03)	8.51 [0.20]
B. Prices chosen two periods in advance						
GDP deflator	0.90 (0.05)	9.46 [0.15]	0.91 (0.04)	4.27 [0.64]	0.86 (0.04)	7.63 [0.27]
NFB deflator	0.87 (0.06)	3.20 [0.78]	0.78 (0.06)	6.07 [0.42]	0.86 (0.04)	7.35 [0.29]
CPI	0.86 (0.05)	10.8 [0.09]	0.81 (0.04)	5.02 [0.54]	0.85 (0.05)	4.08 [0.67]
PCE deflator	0.88 (0.04)	7.72 [0.26]	0.84 (0.03)	5.28 [0.51]	0.88 (0.06)	3.52 [0.74]

Notes: The J_T statistics are distributed as χ^2 random variables with six degrees of freedom. Standard errors are in parentheses. P-values are in brackets. GDP is gross domestic product; NFB is nonfarm business; CPI is Consumer Price Index; and PCE is personal consumption expenditures.
Source: Authors' calculations based upon data from Haver Analytics.

rejected. The p-value associated with the J_T statistic corresponding to every measure of inflation is very small. However, when $\tau = 2$, there is very little evidence against the model. In no case is the p-value less than 0.08. Our view is that this is somewhat of a Pyrrhic victory for the Calvo model. It is entirely possible that there is some delay between when firms reoptimize their price plans and when they actually implement the new plan. But, it is not clear that a six-month delay is a credible assumption.

Consider next the results obtained over the sample period 1959:Q1–72:Q2. A number of interesting findings emerge. First, when $\tau = 1$ the point estimates of θ are substantially smaller than the corresponding estimates obtained over the full sample. For example, with CPI inflation, the point estimate of θ falls from 0.86 to 0.72. Second, with the exception of NFB inflation, there is only marginal evidence against the model when $\tau = 1$. Third, there is virtually no evidence against the model when $\tau = 2$.

Finally, consider the results obtained over the sample period 1982:Q3–2001:Q4. Notice that the point estimates of θ are larger than the corresponding estimates obtained for the first subsample for all values of τ. Perhaps more importantly, there is relatively little evidence against the model with $\tau = 1$ and virtually no evidence against the model when $\tau = 2$.

Viewed as a whole, these results indicate that the Calvo model performs reasonably well if we allow for a split in the sample period and for a lag of roughly one quarter between when firms reoptimize their price plan and when they actually implement the new plan.

Impact of measurement error in inflation

We now consider the results of estimating the model allowing for the possibility that inflation is measured with error of the form

$$12)\quad u_t = \varepsilon_t + \gamma_1\varepsilon_{t-1} + \gamma_2\varepsilon_{t-2} + \ldots + \gamma_q\varepsilon_{t-q}.$$

The model is otherwise the standard Calvo model ($\tau = 0$). For each measure of inflation, we report results for the minimal level of q, such that the overidentifying restrictions of the model are not rejected at the 1 percent level.

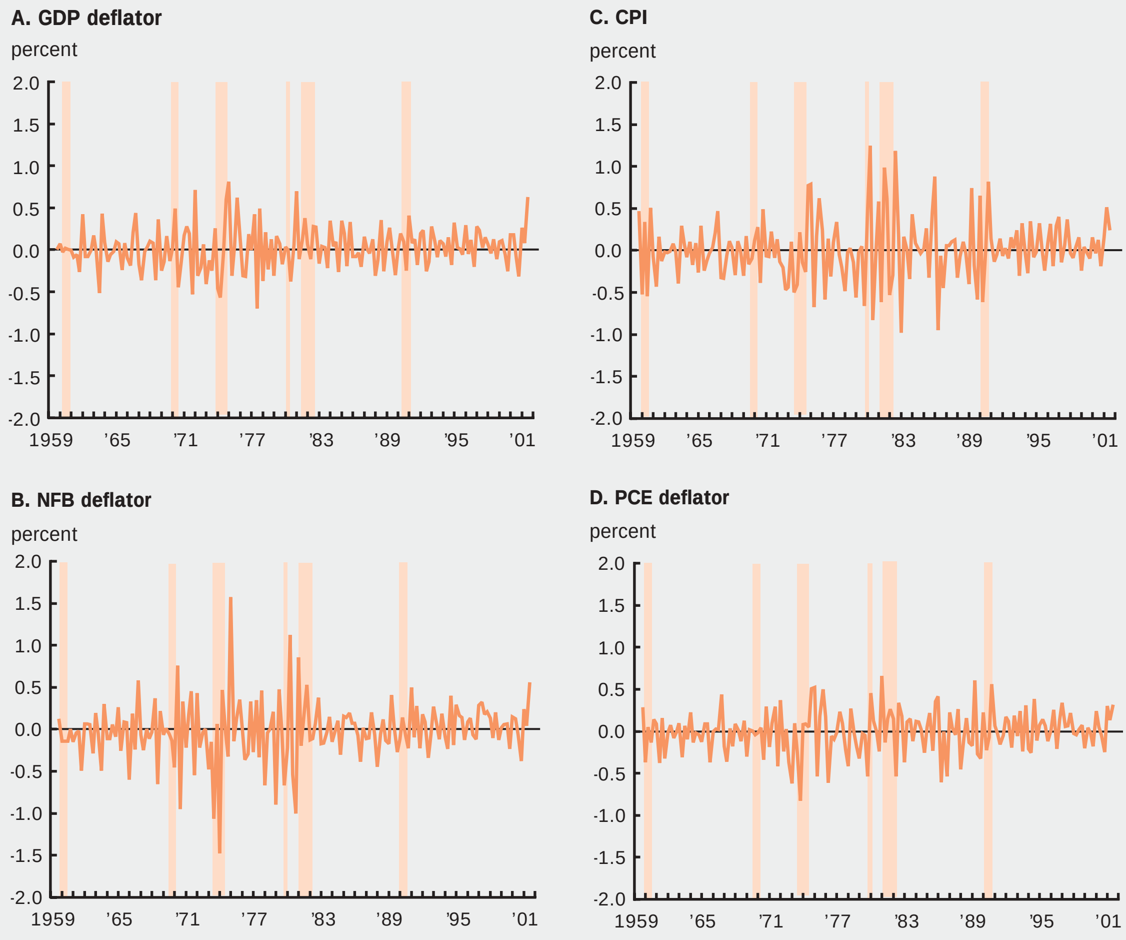
To motivate why this model of the measurement could improve the model's performance, figure 3 displays the basic Euler equation errors emerging from the standard Calvo model estimated over the full sample period. These errors are negatively serially correlated (in all cases the first order correlation coefficient is about –0.25) and are plausibly modeled with a low-order moving average representation. As relation 12 indicates, even iid measurement error can generate a time series for ϕ_t that is negatively serially correlated.

Panel A of table 4 reports the results of estimating the model, allowing for measurement error, based on the full sample period. To help assess the magnitude of the measurement error, the column labeled Γ_1 reports our estimate of the ratio of the variance of true inflation to the variance of the measurement error. This is one measure of the extent of measurement error in the inflation data. Below we refer to Γ_1 as the signal to noise ratio in the inflation data. The column labeled Γ_2 reports our estimate of the percentage of the variance of the composite error term ϕ_t due to classical measurement error that is observed by the econometrician.

A number of key results are worth noting here. First, for all measures of inflation, allowing for iid measurement error overturns the strong rejection of the standard model reported in table 1. Indeed, for the GDP and NFB measures, the overidentifying restrictions cannot be rejected at even the 10 percent level. For the PCE deflator, these restrictions are not rejected at the 4 percent significance level. Second, taking sampling uncertainty into account, our point estimates of the parameter θ are reasonably similar to those reported in table 1. Third, measurement error appears to be more important for the NFB deflator. For the GDP, CPI, and PCE deflators, the ratio of the variance of true inflation to the variance of the measurement error (Γ_1) is 18.4, 13.7, and 19.5, respectively. Evidently, the signal to noise ratio in these inflation measures is high. In the case of the NFB deflator, this ratio is

FIGURE 3

Euler equation residuals



Note: Shaded areas indicate official National Bureau of Economic Research recession periods.
Source: Authors' calculations based upon data from Haver Analytics.

roughly 7.43, so the signal to noise ratio is lower. The percentage of the variance of the composite error term observed by the econometrician due to classical measurement error (Γ_2) is 0.48, 0.45, and 0.46, for the GDP, CPI, and PCE deflators, respectively. But for the NFB deflator, this ratio is roughly 0.73. So based on either the Γ_1 or the Γ_2 statistic, there appears to be more noise associated with the NFB deflator.

Panels B and C in table 4 report our subsample results. Note that for every measure of inflation, there is virtually no evidence against the model in either sample period, once we allow for even iid measurement error. Our point estimates of θ are higher in the second sample period, implausibly so for the NFB deflator. But taking sampling uncertainty into account, one cannot reject the hypothesis, for any measure of inflation or in either subsample period, that firms reoptimize prices, on average, once a year (θ = 0.75).

Turning to our measures of the importance of classical measurement error, a number of results are worth noting. First, in the pre-1979 sample period, the importance of measurement error, assessed using either the Γ_1 or Γ_2 statistic, is highest for the NFB measure of inflation. Indeed, the value of the Γ_2 statistic is so high (0.80) that we are led to conclude that either 1) the NFB is a relatively unreliable measure of true inflation in the first period, or 2) our model of measurement error is implausible. Second, in the post-1982 sample period, NFB inflation has estimated measurement error properties that are quite similar to those of the GDP and PCE deflators. Third, there is a substantial decline in the signal to noise ratio for all three of the inflation measures in the second subsample period.

Viewed as a whole, these results indicate that allowing for classical measurement error results is a large improvement in the model's performance.

TABLE 4

Measurement error in inflation

Inflation measure	θ	J_T	Γ_1	Γ_2
A. Full sample				
GDP deflator	0.91 (0.04)	10.3 [0.11]	18.4	0.48
NFB deflator	0.90 (0.05)	9.54 [0.15]	7.43	0.73
CPI	0.90 (0.06)	16.6 [0.01]	13.7	0.45
PCE deflator	0.91 (0.04)	12.9 [0.04]	19.5	0.46
B. 1959:Q1–79:Q2				
GDP deflator	0.86 (0.04)	4.97 [0.55]	14.8	0.48
NFB deflator	0.82 (0.05)	6.42 [0.38]	5.64	0.80
CPI	0.85 (0.06)	5.87 [0.44]	24.1	0.38
PCE deflator	0.87 (0.05)	5.94 [0.43]	31.8	0.32
C. 1982:Q3–2001:Q4				
GDP deflator	0.92 (0.07)	6.39 [0.38]	3.10	0.59
NFB deflator	0.93 (0.08)	6.14 [0.41]	4.27	0.47
CPI	0.88 (0.06)	5.48 [0.48]	2.12	0.59
PCE deflator	0.92 (0.09)	4.47 [0.61]	3.07	0.65

Note: Γ_1 is the ratio of the variance of the true inflation rate to the variance of the measurement error component; Γ_2 is the fraction of the variance of ϕ_{t+1} due to measurement error. The J_T statistics are distributed as χ^2 random variables with six degrees of freedom. Standard errors are in parentheses. P-values are in brackets. GDP is gross domestic product; NFB is nonfarm business; CPI is Consumer Price Index; and PCE is personal consumption expenditures.
Source: Authors' calculations based upon data from Haver Analytics.

Conclusion

This article discussed the empirical performance of the Calvo model of sticky goods prices. We argued there is overwhelming evidence against this model. But this evidence was generated under three key maintained assumptions. First, there is no lag between the time firms reoptimize their price plans and the time they implement those plans. Second, there is no measurement error in inflation. Finally, monetary policy was the same in the pre-1979 period and the post-1982 period.

Drawing heavily from results in Eichenbaum and Fisher (2003), we discussed the impact of relaxing each of these assumptions. Relaxing the first and third assumptions overturns the evidence against the model, if we are willing to assume that firms wait roughly one quarter before implementing new price plans. Relaxing just the second assumption by allowing for iid classical measurement error is sufficient by itself to render the evidence against the standard Calvo model marginal. If we relax both the second and third assumptions, we find virtually no evidence against the model. Moreover, we find little evidence against the view that firms reoptimize their prices, on average, once a year.

NOTES

[1]See for example, Chari, Kehoe, and McGrattan (2000), Christiano, Eichenbaum, and Evans (2001), Erceg, Henderson, and Levin (2000), Gali and Gertler (1999), Rotemberg and Woodford (1997), and Yun (1996).

[2]For example, Burstein (2002) shows that for moderate changes in the growth rate of money (less than or equal to 5 percent on a quarterly basis), traditional time-dependent models are a good approximation of state-dependent models.

[3]See, for example, Ireland (1997) and Cho and Moreno (2002).

[4]See, for example, Christiano, Eichenbaum, and Evans (2001) and Altig, Christiano, Eichenbaum, and Linde (2003), respectively.

[5]We do not index s_t by j, because all firms have identical marginal costs.

[6]Others, like Dotsey, King, and Wolman (1999), and Woodford (1996), assume $P_{jt} = P_{jt-1}$. Christiano, Eichenbaum, and Evans (2001) also consider a dynamic indexing scheme in which $P_{jt} = \pi_{t-1} P_{jt-1}$. In Eichenbaum and Fisher (2003), we evaluate the performance of the Calvo model under these alternative specifications.

[7]See, for example, Calvo (1983).

[8]For a proof of this, see Woodford (1996) or Yun (1996).

[9]The key assumption is that $\{\hat{\pi}_t, \hat{s}_t, X_t\}$ is a stationary and ergodic process. We also require that $k \geq 2$.

[10]That is, when constructing an estimate of W_T, one could allow for higher order serial correlation in the error term than the theory implies.

[11]Detailed data sources are discussed in the appendix.

APPENDIX

Our data are from the Haver Analytics database. For each data series below, we provide a brief description and, in parentheses, the Haver codes for the series used.

- Price measures: GDP deflator is the ratio of nominal GDP (GDP) and real chain-weighted GDP (GDPH); nonfarm business deflator (LXNFI); Consumer Price Index (PCU); and personal consumption expenditures deflator (JCBM2).

- Real marginal costs: Share of labor income in nominal output for the nonfarm business sector, which is proportional to the U.S. Bureau of Labor Statistics' measure of nominal unit labor costs divided by the nonfarm business deflator (LXNFU/LXNFI).

- Instruments: Quadratically detrended real GDP is the residual of a linear regression of real GDP (GDPH) on a constant, t and t^2; inflation is the first difference of the log of the price measures; the index of commodity prices is the Commodity Research Bureau's index of prices of all commodities (PZALL); the interest rate spread is the difference between a composite of yields on interest rates on Treasury bonds of maturity ten years and greater (FLGT) and the interest rate on three-month Treasury bills (FTBS3); and growth rate of nominal wages is the first difference of the log of nominal compensation in the nonfarm business sector.

REFERENCES

Altig, David, Lawrence Christiano, Martin Eichenbaum, and Jesper Linde, 2003, "Monetary policy and the diffusion of technology shocks," Northwestern University, manuscript.

Burstein, A., 2002, "Inflation and output dynamics with state dependent pricing decisions," University of Michigan, manuscript.

Calvo, Guillermo, 1983, "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics*, Vol. 12, No. 3, pp. 383–398.

Chari, V. V., Patrick Kehoe, and Ellen McGrattan, 2000, "Sticky price models of the business cycle: Can the contract multiplier solve the persistence problem?," *Econometrica*, Vol. 68, No. 5, September, pp. 1151–1179.

Cho, S., and A. Moreno, 2002, "A structural estimation and interpretation of the new Keynesian macro model," Columbia University, manuscript.

Christiano, Lawrence, Martin, Eichenbaum, and Charles Evans, 2001, "Nominal rigidities and the dynamic effects of a shock to monetary policy," Northwestern University, manuscript.

Dotsey, Michael, Robert, G. King, and Alexander L. Wolman, 1999, "State-dependent pricing and the general equilibrium dynamics of money and output," *Quarterly Journal of Economics*, Vol. 114, No. 2, May, pp. 655–690.

Eichenbaum, M., and J. Fisher, 2003, "Testing models of sticky prices," Northwestern University, manuscript.

Erceg, Christopher, J., Dale W. Henderson, and Andrew T. Levin, 2000, "Optimal monetary policy with staggered wage and price contracts," *Journal of Monetary Economics*, Vol. 46, No. 2, October, pp. 281–313.

Fuhrer, Jeffrey, and Jeffery Moore, 1995, "Inflation persistence," *Quarterly Journal of Economics*, Vol. 110, No. 1, pp. 127–159.

Gali, Jordi, and Mark Gertler, 1999, "Inflation dynamics: A structural econometric analysis," *Journal of Monetary Economics*, Vol. 44, pp. 195–222.

Hansen, Lars P., 1982, "Large sample properties of generalized method of moments estimators," *Econometrica*, Vol. 50, pp. 1029–1054.

Hansen, Lars P., and Kenneth J. Singleton, 1982, "Generalized instrumental variables estimation of nonlinear rational expectations models," *Econometrica*, Vol. 50, pp. 1269–1286.

Ireland, Peter, 1997, "A small, structural, quarterly model for monetary policy evaluation," *Carnegie Rochester Conference Series on Public Policy*, Vol. 47, pp. 83–108.

Kim, Jinill, 2000, "Constructing and estimating a realistic optimizing model of monetary policy," *Journal of Monetary Economics*, Vol. 45, No. 2, April, pp. 329–359.

Mankiw, N. G., and R. Reiss, 2002, "Sticky information: A proposal to replace the new Keynesian Phillips Curve," *Quarterly Journal of Economics*, Vol. 117, November, pp. 1295–1328.

Rotemberg, Julio J., and Michael Woodford, 1999, "The cyclical behavior of prices and costs," *Handbook of Monetary Economics*, Vol. 1B, Michael Woodford and John Taylor (eds.), Amsterdam; New York, and Oxford: Elsevier Science, North-Holland.

__________, 1997, "An optimization-based econometric framework for the evaluation of monetary policy," *Macroeconomics Annual*, National Bureau of Economic Research, Vol. 12.

Sbordone, Argia, 2001, "An optimizing model of U.S. wage and price dynamics," Rutgers University, manuscript.

Shapiro, Matthew, and David W. Wilcox, 1996, "Mismeasurement in the Consumer Price Index: An evaluation," *Macroeconomics Annual*, National Bureau of Economic Research, Vol. 11, pp. 93–142.

Taylor, John, 1980, "Aggregate dynamics and staggered contracts," *Journal of Political Economy*, Vol. 88, February.

Woodford, Michael, 1996, "Control of the public debt: A requirement for price stability," National Bureau of Economic Research, working paper, No. 5684.

Yun, Tack, 1996, "Nominal price rigidity, money supply endogeneity, and business cycles," *Journal of Monetary Economics*, Vol. 37, No. 2, pp. 345–370.