May, 2012

Do We Know If Prices Are Forward Looking?

Qiang Zhang

Do We Know Whether Prices Are Forward Looking?[1]

Qiang Zhang
Department of Economics
University of Leicester
qz14@le.ac.uk

Taylor Smith
Department of Economics
University of Mississippi
jtsmith5@olemiss.edu

John R. Conlon
Department of Economics
University of Mississippi
jrconlon@olemiss.edu

May, 2012 (In Revision. Do not quote without permission.)

[1]We are grateful to Brent Meyer for sharing with us the CPI Detailed Report data used in Bryan and Meyer (2010).

Abstract

This paper reexamines the debate on the relative importance of forward versus backward looking price-setting behavior in a hybrid Phillips curve model. We first discuss the challenges in identifying the separate effects of expected future inflation as opposed to lagged inflation. We then develop a model where short run fluctuations in flexible-price-sector inflation resemble measurement errors relative to the long-term-contract/core component of inflation modelled in standard New Keynesian Phillips curves. The measurement error perspective suggests that one should be careful to lag instruments, in order to avoid correlations between the measurement errors in lagged, current, and lead overall inflation and the measurement errors in the instruments. When we do this, the estimated hybrid Phillips curve tends to become significantly more backwards looking. To offset the decline in instrument strength caused by lagging instruments for additional periods, we find that expanding the instrument set by including alternative inflation measures, such as the producer price index for finished goods, improves identification of model parameters. The resulting estimates still point to significantly more backward-looking behavior, implying that the share of backward-looking firms may be close to one.

Key Words: identification, firm price-setting behavior, New Keynesian Phillips Curve, sticky price inflation.

1 Introduction

One of the central issues in Macroeconomics is the degree to which price setting behavior is forward looking or backward looking. While Keynesian-style price stickiness is possible in models with long-term contracts whether agents are forward looking or backward looking, the issue of monetary policy credibility becomes significantly more important if prices are forward, rather than backward looking. For example, if price setters are forward looking, then fairly rapid disinflation is possible, but only if the public trusts the central bank's commitment to price stability. See e.g. Ball (1994, 1995). By contrast, if price setters are backward looking, then rapid disinflation will tend to be costly regardless of the central bank's ability to commit itself, so central bank attempts to commit itself to rapid, credible disinflation may significantly increase the costs of the disinflation.

More generally, Rotemberg and Woodford (1997), Levin, Wieland and Williams (1999) and McCallum and Nelson (1999) have all shown that the performance of different monetary policy proposals in stablizing output and inflation depends on whether forward-looking behavior dominates backward looking behavior in price setting.

This paper reexamines the debate between Gali and Gertler, Rudd and Whelan, and others on the relative importance of forward versus backward looking price setting behavior. In an extremely influential paper, Gali and Gertler (1999) constructed a hybrid New Keynesian Plillips curve (NKPC) model, incorporating both forward and backward looking agents, and presented estimates which suggested that actual price setting behavior is significantly more forward than backward looking. In a very compelling response, Rudd and Whelan (2005, 2006) argued that there was actually no evidence of forward-looking behavior in the U.S. data used by Gali and Gertler. Gali, Gertler and Lopez-Salido (2005) responded that, when one imposed the restrictions of the structural model on the estimates, the results were again forward looking. However, it was not clear why those structural restrictions would have such dramatic effects on their estimates, and, unfortunately, Rudd and Whelan never seem to have addressed this issue.

The Rudd-Whelan criticism is also related to the idea that it is difficult to disentangle the separate effects of lagged and expected future inflation on current inflation using lagged variables as instruments. That is, weak identification problems are likely to arise in NKPC models, as argued by Mavroeidis (2005). This problem has also been studied by Ma (2002), Nason and Smith (2008), Kleibergen and Mavroeidis (2009), and others.

This paper returns to this debate. We first discuss the challenges in independently instrumenting expected future inflation as opposed to lagged inflation, suggesting that instruments may well be, in this sense, quite weak. We then develop a model where short run fluctuations in flexible-price-sector inflation resemble measurement errors relative to the long-term-contract/core component of inflation modelled in the standard New Keynesian Phillips curve (NKPC). We argue that these measurment errors may tend to bias estimates towards appearing spuriously forward-looking. The measurement error perspective also suggests that one should be careful to lag instruments, to avoid correlation between the measurement error in lagged inflation and the measurement error in the instruments. When we do this, the estimated hybrid Phillips curve tends to become significantly more backward looking.

Thus, consider the Gali-Gertler hybrid NKPC:

$$\pi_t = \lambda mc_t + \gamma_f E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \epsilon_t \tag{1}$$

where π_t is inflation between periods $t-1$ and t, $E_t\pi_{t+1}$ is expected future inflation, π_{t-1} is lagged inflation, mc_t is real marginal cost, and Gali, Gertler, Lopez-Salido (2005, p. 1108) assume ϵ_t is *i.i.d.* The Gali-Gertler approach uses GMM to estimate (1) without the expectations on π_{t+1}:

$$\pi_t = \lambda mc_t + \gamma_f \pi_{t+1} + \gamma_b \pi_{t-1} + \hat{\epsilon}_t \tag{2}$$

and with the orthogonality conditions being that $\hat{\epsilon}_t$ is uncorrelated with their instruments, which are essentially lagged values of a set of macroeconomic variables. Here

$$\hat{\epsilon}_t = \epsilon_t - \gamma_f[\pi_{t+1} - E_t\pi_{t+1}] \tag{3}$$

i.e, the error term in (1) minus γ_f times the unanticipated part of π_{t+1}. If agents are rational then the second term in (3) will be uncorrelated with the lagged values of the macroeconomic variables. Thus, if these lagged variables are also uncorrelated with ϵ_t then they will be valid instruments, i.e., they will not introduce any particular bias into estimates of (1).

Unfortunately, standard NKPC models do not generally have a good theory of the ϵ_t term in (1), so it is unclear whether ϵ_t will be uncorrelated with the lagged macroeconomic variables. Furthermore, it is not enough for ϵ_t to be uncorrelated with these instruments. The instruments must also be able to separately identify the effects of π_{t-1} and $E_t\pi_{t+1}$ in (1). That is, there must be information contained in the set of instruments which is able to have a separate influence on $E_t\pi_{t+1}$, independent of its influence on π_{t-1}, and without that influence working through ϵ_t. In other words, the instruments must not be weak or irrelevant.

For example, arguing along the lines of Mavroeidis (2005), suppose mc_t follows an AR(1) process, so $mc_t = \rho mc_{t-1} + \eta_t$, and suppose $\epsilon_t \equiv 0$ for simplicity. Suppose we estimate (2) in this case, using period $t-1$ information as instruments. In large samples, this becomes, essentially, regressing $E_{t-1}\pi_t$ against $E_{t-1}mc_t$, $E_{t-1}\pi_{t+1}$ and π_{t-1}. However, if mc_t follows the above AR(1) process, these latter three variables will be perfectly multicolinear, so (2) is not identified. Thus, any hope of identifying (1) or (2) depends on the details of the time-series structure of the variables mc_t and π_t. Of course, this makes it extremely likely that (1) and (2) will turn out to be either unidentified, or at best, very weakly identified.

On some level the Rudd-Whelan (2005, 2006) argument is similar to the argument in Mavroeidis (2005), but more direct. Like Mavroeidis they focus on the theoretical prediction that the mc_t variable should drive the inflation process. However, rather than arguing indirectly that the dynamics of mc_t are insufficiently rich to allow identification of the separate effects of mc_t, π_{t-1} and $E_t\pi_{t+1}$, they argue that, according to theory, expected discounted future values of mc_t should influence π_t. Since they are unable to find any such influence in their data, they conclude that price setting behavior is not forward looking.

This paper follows Mavroeidis (2005) and Rudd and Whelan (2005) by asking whether there is enough usable independent variation in the right hand side variables in (1) or (2) to estimate these equations reliably. However, unlike those previous papers, we focus less on mc_t and more on the variables $E_t\pi_{t+1}$ and π_{t-1}. We also look more carefully at the error term ϵ_t in order to determine whether this error term might lead to variation in our right hand side variables which would lead to *misleading* estimates of (1) and (2).

We therefore start by writing down a simple model of a NKPC that yields a theory of the error term in (1). Essentially, we simply assume a flexible-price sector as well as a sticky price sector as in Aoki (2001), and the error term arises from marginal cost shocks to the flexible price sector. We then ask about the implications of this for estimation of (1) or (2). We argue that the flexible-price sector's aggregate supply shocks resemble measurement errors in π_t and π_{t-1}. While measurement errors in the dependent variable, π_t itself, do not necessarily cause any biases, measurement errors in the right hand side variable, π_{t-1}, may bias estimates in the usual way. That is, they lead to variations in π_{t-1} which do not necessarily yield movements in π_t (or may even cause movements in the wrong direction). Thus, careless estimates, which do not account for measurement errors may tend to bias γ_b downward. This, in turn, can lead to an upward bias in γ_f, making the data seem spuriously forward-looking. On the other hand, if the measurment error in π_{t+1} is correlated with measurment errors in the instruments, this may partly counteract the above effect.

2 A Three-Sector Model with Core Inflation

This section sets up and solves a modified NKPC with three sectors – a flexible price sector, a sticky price sector with forward-looking rational firms, and a sticky price sector with backward-looking firms. Of course, the two sticky price sectors correspond to the two sectors in the Gali-Gertler (1999) model. The flexible price sector, on the other hand, resembles that in Aoki (2001), and gives us a model of the error term ϵ_t in (1). It also gives us the distinction between overall inflation and core inflation which will be the basis for our measurement error story.

We begin by expressing the overall price level p_t as a weighted average of the aggregate price level among flexible price and sticky price firms, p_t^x and p_t^s, respectively:

$$p_t = \alpha p_t^x + (1 - \alpha) p_t^s, \tag{4}$$

where here and henceforth all price level and cost variables are in natural logs. This gives the highly simplified Aoki (2001) component of our model.

Next, we assume that the price level in the flexible price sector is simply real marginal cost in that sector, $\tilde{mc}_t$, plus a measure of the price level, which for simplicity we assume is p_t^s:

$$p_t^x = \tilde{mc}_t + p_t^s \tag{5}$$

(though see below for a brief discussion of the effect of replacing (5) with $p_t^x = \tilde{mc}_t + p_t$). In the sticky price sector, a fraction $1 - \theta$ of firms are assumed to be allowed to update their prices each period, so

$$p_t^s = \theta p_{t-1}^s + (1 - \theta) \bar{p}_t^* \tag{6}$$

where $\bar{p}_t^*$ is the average new price chosen in period t by the sticky price firms. Next we assume that a fraction ω of sticky-price firms are backward looking, and a fraction $1-\omega$ is forward looking, so

$$\bar{p}_t^* = (1-\omega)p_t^f + \omega p_t^b \tag{7}$$

where p_t^f is the price chosen in period t by forward-looking sticky-price firms, and p_t^b is the price chosen in period t by backward looking sticky-price firms. The forward looking firms choose their prices as a weighted average of expected future nominal marginal costs,

$$p_t^f = (1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k E_t mc_{t+k}^n \tag{8}$$

where β is their discount factor and $mc_{t+k}^n = mc_{t+k} + p_{t+n}$ is nominal marginal cost of the sticky price firms in period $t+k$. The backward looking firms, by contrast, choose p_t^b as in Gali-Gertler,

$$p_t^b = \bar{p}_{t-1}^* + \pi_{t-1}^s \tag{9}$$

except that we've replaced π_{t-1} from Gali and Gertler with inflation among sticky-price firms, i.e., $\pi_t^s = p_t^s - p_{t-1}^s$. Note that this assumes that backward looking agents focus on sticky price inflation as their best predictor of future macroeconomic conditions. This assumes (a) that forward-looking long-term contracts provide a better forcast for future conditions than prices in the flexible-price sector, but (b) backward-looking agents cannot distinguish between forward-looking long-term-contract price setters, and backward-looking price setters like themselves. We do not believe that anything below depends on assumption (b), so (9) could very easily be replaced by $p_t^b = \bar{p}_{t-1}^* + \pi_{t-1}^f$ without changing anything substantive below.

We solve the model in Appendix A. The solution is

$$\pi_t^s = \omega\phi^{-1}\pi_{t-1}^s + \beta\theta\phi^{-1}E_t\pi_{t+1}^s + (1-\theta)(1-\omega)(1-\beta\theta)\phi^{-1}[mc_t + \alpha\tilde{m}c_t] \tag{10}$$

where

$$\phi = \theta + \omega \left[1 - \theta(1 - \beta)\right]. \tag{11}$$

To make sense of (10), note that (4) and (5) imply

$$p_t = p_t^s + \alpha \tilde{mc}_t = p_t^s + \nu_t \tag{12}$$

where we abreviate $\alpha \tilde{mc}_t$ as ν_t for simplicity. If we let

$$\begin{aligned} \gamma_b &= \omega \phi^{-1}, \\ \gamma_f &= \beta \theta \phi^{-1}, \\ \lambda &= (1-\theta)(1-\omega)(1-\beta\theta)\phi^{-1} \end{aligned} \tag{13}$$

above, then (10) becomes

$$\pi_t^s = \gamma_b \pi_{t-1}^s + \gamma_f E_t \pi_{t+1}^s + \lambda mc_t + \lambda \nu_t \tag{14}$$

A few points are in order regarding (14). First note that (14) describes the behavior of *sticky* price inflation, π_t^s, not overall inflation. Also, (14) has an error term, which derives from the aggregate supply shocks caused by fluctuations in the flexible-price sector's real marginal cost variable, $\tilde{mc}_t$. These are the two ways that (14) deviates from the standard hybrid NKPC.

Of course, (14) resembles the standard hybrid NKPC in other ways. In particular, note that, even if the true model is purely backward looking, so $\omega = 1$, the forward-looking coefficient, γ_f, will still be nonzero, and in fact, $\gamma_f / \gamma_b = \beta\theta$ in this case. Since $\beta \approx 1$ and θ also tends to be quite large (for example, most of the Gali-Gertler estimates tend to be around 0.8 or larger), this suggests that, even if behavior is *entirely* backward looking, we would expect γ_f to be almost as large as γ_b.

Finally, note that variations in the model would lead to equations very similar to (10). For example, if we replaced (5) with $p_t^x = \tilde{mc}_t + p_t$ (instead of $\tilde{mc}_t + p^s_t$), then the coefficient α on $\tilde{mc}_t$ would be replaced by $\alpha/(1-\alpha)$ (see the discussion after (A.9) in the appendix). On the other hand, if we replaced $mc^n_{t+k} = mc_{t+k} + p_{t+n}$ with $mc^n_{t+k} = mc_{t+k} + p^s_{t+n}$, then no $\tilde{mc}_t$ term would appear at all in (10).

3 Econometric Issues in the Three-Sector Model

The results in (14) suggest that, if one wants to estimate a hybrid NKPC model, one should use some sort of measure of core inflation rather than total inflation in one's estimates, and we shall do this below. However it is, presumably, impossible to perfectly disentangle sticky from flexible price sectors. In this section we therefore briefly consider the implications of the model when we estimate (1), if the true model is (14).

3.1 Inflation Measures and Instrument Lags

The implications of our model depend on the behavior of the real marginal cost variable, $\tilde{mc}_t$, in the flexible price sector, through the disturbance $\nu_t = \alpha\tilde{mc}_t$.

Specifically, (12) and (14) imply that

$$\pi_t = \gamma_b \pi_{t-1} + \gamma_f E_t \pi_{t+1} + \lambda mc_t + \eta_t \tag{15}$$

where

$$\begin{aligned} \eta_t &= -\gamma_f E_t \nu_{t+1} + (1 + \lambda + \gamma_f)\nu_t - (1 + \lambda + \gamma_b)\nu_{t-1} + \gamma_b \nu_{t-2} \\ &= -\gamma_f (\nu_{t+1} - \nu_t) + (1 + \lambda)(\nu_t - \nu_{t-1}) - \gamma_b (\nu_{t-1} - \nu_{t-2}) + \varepsilon_{t+1}. \end{aligned} \tag{16}$$

Here $\varepsilon_{t+1} \equiv \nu_{t+1} - E_t \nu_{t+1}$ is the expectations error in ν_{t+1}.

Thus, (16) suggests that the error term, η_t, has a moving average structure. This structure implies that, if ν_t is *i.i.d.*, it will be negatively correlated with the right-hand side variables π_{t-1} and $E_t\pi_{t+1}$ under the assumption of non-zero ω and θ,[1] because both inflation terms share a common term with η_t, i.e.

$$\pi_{t-1} = \pi^s_{t-1} + \nu_{t-1} - \nu_{t-2} \tag{17}$$

by first-differencing eq. (12), which also implies

$$E_t\pi_{t+1} = E_t\pi^s_{t+1} + E_t\left(\nu_{t+1} - \nu_t\right) \tag{18}$$

In addition if ν_t is $i.i.d$, (17) implies that the covariance of π_{t-2} and η_t is $\gamma_b\sigma^2_\nu$, so π_{t-2} cannot be a valid instrument, either. To the extent that π_{t-2} may also be correlated with other variables chosen as instruments at period $t-2$,[2] correcting for the endogeneity in π_{t-1} requires instruments to be lagged at least *three* periods (relative to period t).

Thus, we would expect estimates of (15) which include π_{t-1} in the instrument set to suffer from the classic measurement errors problem, and so, lead to downward biases in γ_b. In addiction, these problems may persist if we remove π_{t-1} but keep π_{t-2} in the instrument set. Thus, since π_{t-1} and π_{t+1} tend to be positively correlated (0.62 for our sample), we would expect this to lead to an upward bias in γ_f. Finally, to the extent that our estimates suffer from a weak instruments problem, we would expect these estimates to be very sensitive to this measurement error problem, so the resulting biases can potentially be quite large.

Of course, if ν_t is autocorrelated and stationary, then lagging inflation three periods may not be enough, so we also consider estimates with our choice of instruments lagged four or more periods. Some researchers have intuively argued for lagging instruments by two periods to avoid the en-

[1] A non-zero ω implies $\gamma_b \neq 0$, and a non-zero θ ensures $\gamma_f \neq 0$.

[2] For example, the output gap of the sticky price sector may be negatively affected by the real marginal cost in the flexible price sector. So a rise in ν_{t-2} may induce a fall in the output gap in the sticky price sector. If the output gap in the flexible price sector is zero (due to flexible prices), then the aggregate output gap at $t-2$ will be negatively correlated with ν_{t-2}. This implies that output gap in $t-2$ also cannot be a valid instrument.

dogeneity caused by an autoregressive component in the instruments. Our model indicates that that intuition, although it is qualitatively correct, may miss the moving average structure above and may so lead to the use of invalid instruments such as π_{t-2}, as well as output gap and real marginal cost from period $t-2$.

On the other hand, suppose ν_t is a random walk, rather than an *i.i.d.* process, so $\nu_t = \nu_{t-1} + \mu_t$, say, where $E_t\mu_{t+1} = 0$. Then (19) implies that

$$\eta_t = -\gamma_f \mu_{t+1} + (1+\lambda)\mu_t - \gamma_b \mu_{t-1} + \varepsilon_{t+1},$$

and (12) implies

$$\pi_{t-1} = \pi^s_{t-1} + \mu_{t-1}.$$

Therefore π_{t-1}, again, is correlated with the error term η_t and so, is not a valid instrument. However, $\pi_{t-2} = \pi^s_{t-2} + \mu_{t-2}$, so the presence of the measurement error term μ_{t-2} does not imply correlation with the error term in (15) and π_{t-2} *is* a valid instrument in this case. In practice, the first difference of ν_t inferred from the data on the overall and sticky price inflations seems stationary and serially uncorrelated (see below for details), π_{t-2} may therefore be a valid instrument. However, if we imagine the error term η_t reflecting not just the influence of sectors with very flexible prices, but also sectors with moderately flexible prices, then π_{t-2} may not be a valid instrument. More generally, it is conceivable that η_t may include specification and measurement errors from other sources than considered in the present paper. Such errors still invalidate π_{t-2} as instrument.

By (18), even if we had a perfect measure of $E_t\pi_{t+1}$, it would still be an endogenous regressor in (15) because of its correlation with η_t due to the second term in (18). This is a novel and surprising econometric implication that is missing in the influential one- or two-sector NKPC models such as those developed by Roberts (1995, 1997) and Gali and Gertler (1999). The reason is, again, because those NKPC's are not derived with a theory on what their error terms should include. Very often, in the estimation of those NKPC's, researchers correctly speculate that supply shocks should enter the error term. However, without a more complete model such as the one presented above, it is impossible to speculate on what the error structure may be.

The estimation of (14) will also shed new insight on firm pricing behavior, because it is sticky price inflation that the Calvo (1983) model and its extension in GG (1999) describes. Furthermore, (14) is free of the above measurement error problem. The error term of this equation is driven by the flexible price sector's real marginal cost $\tilde{mc}_t$. Hence, to estimate (14) properly requires knowing the statistical properties of $\tilde{mc}_t$. Since $\tilde{mc}_t$ is not directly observable, we will use the difference between flexible and sticky price inflations that we construct from micro price duration estimates to infer its autocorrelation and heteroscedasticity. In addition, we will also exploit the cross-correlations between this difference and sticky price inflation to infer the endogeneity of the three regressors in (14).

A few words on the choice of estimators are in order. The fact that ν_t and η_t may be autocorrelated implies that estimators that perform better than 2-step GMM in the presence of weak instruments, but are sensitive to violations of the *i.i.d.* assumption, should not be chosen for the estimation of (14) and (15). This rules out k-class estimators such as LIML, bias-adjusted 2SLS and Fuller's modified LIML. See Baum, Schaffer and Stillman (2007). However, the continuously updated GMM of Hansen, Heaton and Yaron (1996) does not require the *i.i.d.* assumption and may be less biased in an finite sample than the 2-step GMM and seems more robust to weak instruments. See e.g. Hahn, Hausman and Kuersteiner (2004). Hence, in our empirical work, we will report both 2-step and continuously updated GMM estimates.

3.2 Weak Identification and Under Identification

Of course, when instruments are lagged three or more periods, their correlations with the endogenous regressors π_{t-1}, $E_t\pi_{t+1}$ and mc_t tend to decline, possibly leading to irrelevant and weak instruments issues. Irrelevant instruments lead to under-identification, whereas weak instruments tend to produce weak identification of slope parameters.

To keep instrument strength under check when we lag instruments for three or more periods, we exploit two tests of weak identification. The Angrist-Pischke (2009) F test tests weak identification of each slope coefficient in a regression with multiple endogenous regressors. The idea of this test is to first isolate the variation in the fitted value of an endogenous regressor that cannot be explained by the fitted values of other endogenous regressors, where all fitted values are from first-stage regressions of endogenous regressors onto the instruments. Then how much of this variation is jointly explained by the excluded instruments is measured by an F test in an OLS regression. The Kleibergen and Paap (2006) Wald *rk* F test, in contrast, tests if the NKPC equation as a whole is weakly identified. We do not report the usual first-stage F statistic, because it is valid when there is only one endogenous regressor. Nor do we report the Cragg and Donald (1993) Wald F statistic: it requires the error term in the NKPC equation to be *i.i.d.*

If tests of weak identification do not reject it, estimation and inference based on the standard asymptotic theory becomes invalid for small to moderate sample sizes. It is therefore important to conduct weak-identification robust tests of the three-sector NKPC model. For this purpose, we report the Anderson and Rubin (1949) χ^2 test of the joint hypothesis that the orthogonality conditions are valid and the slope coefficients are equal zero. Since by lagging instruments by three or more periods, we increase the chance of the orthogonality conditions being true, we interpret a rejection by this test as a sign that the slopes γ_b, γ_f, and λ are different from zero.

Furthermore, we also test for the possibility of under-identification. Under identification occurs when the full rank condition for identification fails. This failure may happen for each endogenous variable when there is no new information in one instrument given the other instruments, so that it is just redundant, or when all three canonical correlations between linear combinations of three endogenous regressors and linear combinations of all K instruments are not significantly different from zero. We will adopt a χ^2 test of under-identification suggested by Angrist-Pischke (2009) that can be applied to each endogenous variable in a linear regression equation. The Kleibergen and Paap (2006) rk LM statistic is another option for testing under-identification at the equation level. For our goal of sorting out the relative importance of forward- and backward-looking behavior, the AP (2009) χ^2 test is more informative. This is because there are cases when one behavior is estimated to clearly dominate the other by the 2-step GMM, yet this test may reveal that the dominating behavior is in fact unidentified. In such a scenario, the point estimates alone may be misleading.

If we reject under-identification, then the question is whether we reject weak identification. If we do not reject weak identification, then we ask what the weak-identification robust test tells us about the slope coefficients in the NKPC equation. If we reject weak identification, we use the point estimates as well as the 95% confidence intervals for them to infer the importance of forward-looking behavior.

Last but not least, we explore the possibility of finding new instruments that help us to protect our estimation against instruments weakening when instruments are lagged three or more periods. The tests of under or weak identification that we have just described are useful for understanding if new instruments improve our estimation. In particular, if these tests show that under- or weak identification tends to disappear, or be a less severe problem after these new instruments are added to the instruments list in the estimation, then we conclude that the resulting point estimates should carry more weight than those based on standard instruments in this literature.

3.3 Other Econometric Issues

Our economic model provides a useful platform for addressing the econometric issues that surrounds the GMM estimation of (1). Rudd and Whelan (2005) pointed out that the difficulty with the single-equation instrumental variables approach is that variables chosen as instruments may in fact directly cause current inflation in addition to indirectly affecting it through the expected future inflation term in the hybrid NKPC equation. In such a scenario, the GMM estimation of (1) may be biased towards finding a significant role for expected future inflation, regardless of whether such a role exists in the data.

On the other hand, RW's (2005) point can be extended to other instruments for real marginal cost. To demonstrate that they only indirectly affect current inflation through real marginal cost, it is necessary to include them in the right-hand side of (1) as well.[3]

In a single-equation setting, one natural response to this concern is to estimate an expanded version of (1) by including the instruments in question in the right-hand side. In fact, this is what GG (1999) and GGLS (2001, 2003) have tried to do when they include additional lags of inflation in the right hand side of (1). However, since these additional lags are also used as instruments in these two studies, the flexible inflation lags appear in both instruments and the disturbance term. Consequently, the orthogonality conditions that underlie the GMM estimation should most likely fail to hold due to the non-zero correlations between lags of flexible inflation. Hence, even though they have shown that the additional lags of inflation are not statistically significant, to the extent that it may be caused by inconsistent estimation due to the failure of orthogonality conditions, such exercises still

[3]Of course, all instruments cannot be included in the right-hand side of an expanded version of eq. (1) at the same time.

cannot rule out that additional inflation lags move not only expected future inflation but also current inflation.

Our economic model outlined above suggests that if overall inflation is used to estimate (1), sticky price inflation lagged by suitable numbers of periods will be better instruments than lags of overall inflation in GMM estimation. This is because now to address RW's concern, such lags of sticky price inflation can be included in (1) without causing the orthogonality conditions to fail, provided that sticky price inflation is precisely measured. In our empirical work, we plan to experiment with alternative measures of sticky price inflation based on different cutoffs of median price durations to check the robustness of our results.

The micro price durations estimated by Bils and Klenow (2004, BK henceforth) and Nakamura and Steinsson (2008, NS henceforth) allow us to construct alternative sticky and flexible price inflations based on the items included in the U.S. consumer price index. The BK study does not remove sales price in estimating price durations, while the NS paper presents results that include or exclude sales.

Furthermore, with sticky and flexible price inflations explicitly measured, we can compute cross correlations between these two measure of inflation over alternative leads and lags, so that we rule out lags of sticky price inflation that are correlated with period $t-1$, t and $t+1$ flexible price inflations that appear in the error term of the NKPC equation.

4 Data

To estimate the three-sector model, we need measures of core or sticky price inflation and flexible price inflation, overall price inflation, real marginal cost, and output gap.

Empirical studies of micro price stickiness referenced above allow flexible and sticky price inflations to be constructed. One set of such measures is provided by Bryan and Meyer (2010). They first classify the 350 U.S. CPI entry level items (ELIs) studied by BK (2004) into the 45 major components of CPI inflation. They then use BK's (2004) estimates of ELI price durations to compute the average price duration for each of the 45 CPI components, and label those with average price duration estimates shorter than 4.3 months as the flexible price group. The other of the 45 components, with average price durations longer than 4.3 months, constitute the sticky price group. The cutoff 4.3 months is the median of the estimated price durations of the 350 CPI entry level items taken from BK (2004). The sticky price inflation is the weighted average inflation of the CPI components with average price durations above the 4.3 month cutoff. The flexible price inflation is the weighted average inflation of the remaining CPI components.

Also using data on CPI ELIs, NS (2008) have estimated generally longer price durations, mainly because of their removal of sales (temporary price cuts) from the data. Depending on how product substitution is treated, the median of their estimated price durations is 8 or 9 months, about twice as long as the 4.3 months referenced above.[4]

Since sales do not fit into the Calvo model underlying our three-sector NKPC, we have constructed our own measures of flexible and sticky price inflations by combining NS's (2008) micro price duration estimates with the inflation rates of the same 45 CPI components as in BM (2010), based on the CPI Detailed Report data. We use a 3-month cutoff for the average price duration to classify these 45 CPI components into flexible and sticky price groups. This choice of cutoff is dictated by our use of quarterly data: at quarterly frequency, flexible prices are those that are adjusted at least once a quarter. In Tables 1 and 2, we list the flexible and sticky price categories along with each category's average price duration and weight in the CPI.

[4]They also find that the price duration estimates for the components of producer price index (PPI) seem to be less sensitive to the inclusion or removal of sales. However, since the PPI categories that they investigate do not seem to include any service, it is not clear how robust this finding is.

As Table 1 indicates, the flexible price goods and services are used cars, car and truck rental/lease, three categories of energy goods, and lodging away from home. Notably, no food items enter this table, because the two most flexible food price categories, fresh fruits and vegetables and dairy and related products, have average durations of 3.6 and 3.7 months. They are therefore included in Table 2 as sticky price categories in the CPI. This may seem surprising, because food and energy goods are routinely considered the most volatile CPI components, and are therefore netted out in the BLS's core CPI inflation measure. However, recall that our classification is based on NS's (2008) estimates which is based on removing sales. To the extent that most foods are perishable and may be subject to more frequent sales than non-perishable goods, our classification of food categories is reasonable.

In Figure 1, we display the different properties of our measures of sticky and flexible price inflations for the period 1967-2011, which is the sample period of our data. The sticky price inflation is very persistent and stable, whereas the flexible price inflation is transient and volatile. Since we will estimate two versions of hybrid NKPC, one based on overall price inflation and the other on sticky price inflation, it is useful to know the cross-correlation between these two inflation series as well.

Whether we use BK or NS measures of price durations (i.e. whether sales are removed from micro price data or not), our two measures of flexible price inflation created with a 3-month price duration cutoff displays no autocorrelation at the 5% significance level. Both are heteroscedastic. This suggests that when overall price inflation is used in estimation, conditional heteroscedasticty must be accounted for. It also suggests that the autocorrelation of the vector $\mathbf{Z}_t\eta_t$ in the orthogonality conditions is likely to be fully determined by the autocorrelations of the instruments $\mathbf{Z}_t$. Such information pins down the bandwidth parameter in the HAC covariance estimator—a luxury not affordable without an economic model on what the error term in (1) should be.

Our measure of real marginal cost of the sticky price sector (relative to its steady state value) follows GG's (1999) practice of using the percentage deviation of the labor share of private sector output from its steady-state value. However, instead of using the non-farm business sector labor share as in their paper, we must include the agricultural sector to calculate the overall real marginal cost for the entire private sector, because agricultural products prices are now classified as sticky at quarterly frequency as just explained. On the other hand, we must ideally remove the energy sector as well as three transportation categories involving used cars and car/truck rental/lease, in addition to lodging away from home, in computing the labor share of the sticky price sector. Given that the National Income and Product Accounts tables with expanded details do not report income data at such disaggregate levels, we use the labor share of national income for the private sector to construct the real marginal cost for both the flexible and sticky price sectors. Hence, our alternative measure of real marginal cost is the percentage deviation from the sample average of labor share of private sector national income calculated from the Table 1.12, *National Income by Type of Income*.[5]

In addition, the output gap measure that appears in our instruments list is defined as the percentage deviation of real GDP from its potential level as determined by the U.S. Congressional Budget Office (CBO). The cyclical unemployment, another possible instrument, is measured as the difference between the civilian unemployment rate and the long-term natural rate of unemployment, also given by the CBO.

[5] Proprietor's income is assumed to have the same sample average labor share as national income. More precisely, we solve the labor share of proprietor's income that satisfies this assumption, given that the labor share of national income is a function of the assumed labor share of proprietor's income.

5 Empirical Results

To test the three-sector NKPC model, we run three sets of estimations with standard instruments, and two additional sets of estimation with additional new instruments. Our first three sets of estimation are for the overall CPI inflation instrumented by its lags [i.e. eq. (15)], the overall CPI inflation instrumented by lags of sticky price CPI inflation, and then the sticky price CPI inflation instrumented by its own lags [i.e. eq. (14)]. The motivation for this design is to test if we will discover the same pattern of estimation and testing results across these three specifications. All three are testable implications of our model. If we obtain different patterns of estimates across them, our model must have missed something important in the real world. However, if our estimates display similar patterns, these three specifications will serve as robustness checks for each other.

In addition, our estimation differs from GG (1999) in some other important ways. First, we focus on estimating the linear versions of the hybrid NKPC, because of our desire to understand how well different sets of instruments perform in identifying this simpler form than the structural, nonlinear, form. The econometric theory for instrumental variables estimators and associated tests for linear regressions is currently better developed than that for nonlinear models. Figuring out what instruments work out better in the linear NKPC should then help improve the identification and estimation of the fully structural version of the NKPC. Second, since we have constructed model-consistent sticky price inflation, and our measure of it is based on recent advances in estimating micro price stickiness, our estimates of the role of forward- and backward-looking behavior in inflation dynamics are likely to be more accurate.

5.1 Results Based on Standard Instruments

Tables 3-5 report the estimation and test results when instruments are similar to those used by GGLS (2006), except here we experiment with eight more lags for them one by one. More specifically, each set of instruments consists of four lags of inflation and two lags of output gap and real marginal cost, i.e. eight instruments.

Table 3 is based on using the overall CPI inflation as the inflation measure. Hence the NKPC equation estimated here is (15). The top two rows of this table report the estimates of γ_b and γ_f. Although forward-looking dominates backward-looking when instruments are lagged once or twice, this dominance is not robust to additional lagging of instruments. For the last seven sets of estimation, γ_b dominates γ_f twice, γ_f dominates γ_b twice, and the remaining three sets of γ_f and γ_b estimates are not significantly different from each other.

Not surprisingly, when instruments are lagged more and more quarters, the AP (2009) χ^2 test displays an increasing proportion of non-rejection of the null hypothesis of under-identification: the p-values in the row headed by AP χ^2 change from virtually 0 to close to 1. In addition, the AP (2009) F statistic for weak identification in general declines with the number of lags in instruments, suggesting that weak-identification increasingly may not be rejected. The critical values of this test statistic is currently unknown, but Baum et al. (2007) recommend the use of Stock and Yogo (2005, p. 100) critical values for the Cragg and Donald (1993) F statistic for the case of a single endogenous regressor instead. These critical values change with the number of instruments used. For eight instruments and at the 5% level, they are

- 5% maximal IV relative bias: 20.25
- 10% maximal IV relative bias: 11.39

- 20% maximal IV relative bias: 6.69
- 30% maximal IV relative bias: 4.99

These critical values are based on the researcher's desired limit on the maximal bias in the 2SLS estimator relative to that of the OLS estimator. Hence, to achieve the 5% maximal relative bias in the IV estimation, the AP F statistic has to be at least 20.25. On the other hand, an AP F statistic of 5.2 implies that the maximal IV relative bias is somewhere between 20% and 30%. Some authors use the cutoff of 30% maximal IV relative bias for an instrument to be considered weak. However, it should be noted that these critical values are derived under the assumption of *i.i.d.* errors, and therefore can only be seen as suggestive.

On the other hand, the KP *rk* F test for weak identification at the equation level also declines in general with the number of lags in instruments. The suggestive benchmarks for understanding what these statistics imply are Stock and Yogo's (2005) critical values for the case of three endogenous regressors and eight instruments at the 5% level of significance

- 5% maximal IV relative bias: 15.18
- 10% maximal IV relative bias: 9.01
- 20% maximal IV relative bias: 5.69
- 30% maximal IV relative bias: 4.46

In the bottom row of Tables 3, all but one KP *rk* F statistics are below 4.46. Hence, this test seems to suggest that weak identification becomes a serious problem once instruments are lagged more than once. This is in agreement with the AP F test results aforementioned.

Tables 4 and 5 report similar results to those in Table 3. That is, as instruments are lagged more periods, the dominance of γ_f over γ_b tends to weaken. However, the weak instruments problem also tends to worsen. Therefore, it does not seem clear if our challenge to GG (1999) and GGLS (2006) is econometrically sound up to this point, although in Table 5, when instruments are lagged twice, it is clear that γ_b clearly dominates γ_f, and this set of results pass the KP *rk* F test (6.40) at the 5% level for the 20% maximal IV relative bias. This dilemma leads us to consider two measures: new instruments, and different estimators such as the continuously updated GMM, that are known to be less susceptible to instrument weakness.

5.2 Results Based on Adding New Instruments

In Tables 4 and 5, we report two sets of results with four lagged values of the inflation in the Producer Price Index (PPI) for Finished Goods added to the instruments list. The intuition for considering a PPI inflation measure as instruments lies in the possibility that what producers receive for selling their products (which is what PPI measures) in the recent past may be correlated with CPI inflation of the last period and the expected inflation for the next period in opposite directions, thereby helping to tell apart the two coefficients γ_f over γ_b. For example, the PPI inflation of recent quarters may pass through to CPI inflation, but at the same time alert the Federal Reserve to the prospect of higher future inflation if no policy action should take place. If this prospect leads the Fed to raise interest rates according to the Taylor rule, it may help lower inflation expectation.

Before we describe the empirical results in Tables 4 and 5, an update of the suggestive 5% critical values for the AP F and KP rk F tests are in order. For twelve instruments, the former test with a single endogenous regressor has the following benchmarks

- 5% maximal IV relative bias: 21.01
- 10% maximal IV relative bias: 11.52

- 20% maximal IV relative bias: 6.53
- 30% maximal IV relative bias: 4.75

And for the KP rk F test, the following critical values may be useful for inference on weak instruments at the equation level

- 5% maximal IV relative bias: 17.8
- 10% maximal IV relative bias: 10.01
- 20% maximal IV relative bias: 5.90
- 30% maximal IV relative bias: 4.42

When eq. (14) is estimated with these new instruments, the dominance of γ_f over γ_b nearly disappears: only in two of the nine sets, does this dominance occur. The first incidence of these two cases is when we treat the first lag of sticky price inflation as exogenous, which itself may be legitimate if sticky price inflation is accurately measured. However, the second incidence (see column 3) is in fact associated with rather weakly identified γ_f: though the AP F statistic for this slope is 5.05, above the critical value of 4.75 for 30% maximal IV bias, the KP *rk* F statistic for this case is merely 2.94, one of the three lowest in this table and is below the 4.42 cutoff for the same level of bias at the equation level. In fact, the only set of results beyond first lagging of instruments that pass the weak instruments test is

On the other hand, the AP (2009) χ^2 test displays an over-whelming rejection of the null hypothesis of under-identification: the p-values in the row headed by AP χ^2 are all virtually 0. At the same time, the AP F test statistic becomes larger than those in Table 5 across all instrument lags. Hence not only there is now more evidence for backward-looking, but also under-identification is no longer a problem, and weak identification has become far less severe when compared with Tables 4 and 5. For example, in column 5, which is for the case of the instruments lagged by five to eight quarters, the three AP F statistics are 7.95, 7.73, and 8.73. Checked against the suggestive critical values 6.53 reported above, these numbers imply that the relative biases in GMM estimates for the three slopes in the NKPC are all within 10-20% of those in the OLS estimates, though the joint AP rk F test indicates that the biases in slope estimates falling between 20% to 30% cannot be rejected, either. This is still less than ideal, but practically speaking, it is also reassuring because we have added only one new series to bear on the weak instruments issue. For this case, it is notable that γ_b dominates γ_f in statistically significant terms. In addition, the KP *rk* F statistic becomes larger at the far end of lag numbers for instruments than those in Table 5.

Table 7 present some results that are similar to those in Table 6: under-identification is uniformly rejected across all nine sets of instruments (the AP χ^2 statistics all have zero p-values), and weak identification is less of a problem (as indicated by the increase of AP F and KP *rk* F statistics relative to those in Table 4). γ_b is dominated by γ_f in the first three cases and they all pass both weak instruments tests if the 30% maximal relative bias is used as the cutoff. However, for the next three cases, this is reversed, with γ_b not always significantly larger than γ_f. In addition, they each pass a weak identification test at the 5% level for the 30% maximal relative bias cutoff. The remaining three cases are evenly distributed in terms of the strength of forward- v.s. backward-looking behavior.

In particular, since in this model, when all firms are backward-looking, $\gamma_b = \gamma_f = 0.5$, it is clear that more than half of the results in these two tables do not reject this equality under the standard asymptotic theory, implying that the majority of firms may in fact be backward-looking in price setting.

6 Conclusions

We have derived a simple three-sector model of firm pricing that leads to a hybrid NKPC for the sticky price inflation that differs slightly from the influential Gali and Gertler (1999) formulation. Our formulation implies that the error term to the NKPC should include the relative price of the flexible price sector, in the spirit of Aoki (2001), whereas their formulation does not have an intrinsic error term due to their focus on a homogeneous sticky price sector.

Alternatively, our model implies a hybrid NKPC for overall price inflation with an error term that is characterized by a moving average structure in the relative price of the flexible price sector. A major econometric implication of this formulation is that valid instrumenting requires inflation to be lagged by two or more periods when overall price inflation is used as the inflation measure.

In order to estimate these two specifications of hybrid NKPC implied by our model, we have used the micro price duration estimates reported by recent empirical studies to construct sticky price inflation and flexible price inflation. Using these data along with the overall CPI inflation in our estimation, we find that the dominant role of forward-looking behavior in inflation dynamics is mainly driven by once or twice lagged instruments and by treating the first lag of inflation as exogenous in the NKPC equation. Once we go beyond two lags in instrumenting as suggested by our economic model, and once we introduce new information into the instrument list, the evidence for forward-looking behavior become much weaker. It seems more often to observe roughly equal forward- and backward-looking strength in inflation dynamics. This implies that the share of backward-looking firms in the sticky price sector must be sufficiently close to one. In addition, there is improvement in identification as new instruments are brought in to strengthen estimation: under-identification is decisively rejected, and weak identification is less of a problem.

We plan to experiment with alternative measures of expected inflation as well as other new instruments in future research. Survey and forecast based measures of expected inflation should be good instruments for the CPI expected inflation in the NKPC equation that we estimate. If they are positively correlated with future inflation, but not strongly correlated with past inflation, they may help improve identification as well.

7 References

Anderson, Theodore, and Herman Rubin (1949), "Estimation of the parameters of a single equation in a complete system of stochastic equations," *Annals of Mathematical Statistics* 20, 46–63.

Angrist, Joshua and Pischke, Jorn-Stephen (2009), *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press.

Aoki, Kosuke (2001), "Optimal monetary responses to relative price changes, "*Journal of Monetary Economics* 48, 55-80

Ball, Laurence (1994), "Credible disinflation with staggered price-setting," *American Economic Review* 84, 282-289.

Ball, Laurence (1995), "Disinflation with imperfect credibility," *Journal of Monetary Economics* 35, 5-23.

Baum, Christopher, Mark E. Schaffer, and Steven Stillman (2007), "Enhanced routines for instrumental variables/generalized method of moments estimation and testing," *Stata Journal,* 7, 465-505.

Bils, Mark and Peter Klenow (2004), "Some evidence on the importance of sticky prices," *Journal of Political Economy* 112, 947-985.

Bryan, Michael and Brent Meyer (2010), "Are some prices in the CPI more forward looking than others? We think so." *Economic Commentary*, Federal Reserve Bank of Cleveland.

Cragg, John, and Stephen Donald (1993), "Testing identifiability and specification in instrumental variables models," *Econometric Theory* 9, 222–240.

Cumby, Robert and John Huizinga (1992), "Testing the autocorrelation structure of disturbances in ordinary least squares and instrumental variables regressions,"*Econometrica* 60, 185-95.

Gali, Jordi, and Mark Gertler (1999), "Inflation dynamics: a structural econometric analysis," *Journal of Monetary Economics* 44, 195-222.

Gali, Jordi, Mark Gertler and J. David Lopez-Salido (2005), "Robustness of the estimates of the hybrid New Keynesian Phillips curve," *Journal of*

Monetary Economics 52, 1107-1118.

Hahn, Jinyong., Jerry Hausman, and Guido Kuersteiner (2004), "Estimation with weak instruments: Accuracy of higher-order bias and MSE approximations, "*Econometrics Journal* 7, 272-306.

Kleibergen, Frank, and Richard Paap (2006), "Generalized reduced rank tests using the singularvalue decomposition," *Journal of Econometrics* 127, 97-126.

Kleibergen, Frank and Sophocles Mavroeidis (2009), "Weak instrument robust tests in GMM and the New-Keynesian Phillips curve," *Journal of Business & Economic Statistics*, 27, 293-311.

Levin, Andrew, Volker Wieland, and John C. Williams (1997), "Robustness of simple monetary policy rules under model uncertainty," in *Monetary Policy Rules*, 263-318, National Bureau of Economic Research, Inc.

Ma, Adrian (2002), "GMM estimation of the new Phillips curve," *Economics Letters* 76, 411-417.

Mavroeidis, Sophocles (2005), "Identification issues in forward-looking models estimated by GMM, with an application to the Phillips curve," *Journal of Money, Credit and Banking* 37, 421-448.

McCallum, Bennet and Edward Nelson (1999), "Nominal income targeting in an open-economy optimizing model,"*Journal of Monetary Economics* 43, 553-.578.

Nakamura, Emi and Jon Steinsson (2008), "Five facts about prices: a reevaluation of menu cost models," *Quarterly Journal of Economics* 123, 1415-1464.

Nason, James M. and Gregor W. Smith (2008), "Identifying the New Keynesian Phillips curve, " *Journal of Applied Econometrics* 23, 525-551.

Roberts, John M. (1997), "Is inflation sticky?" *Journal of Monetary Economics*, 39, 173-196.

Rotemberg, Julio and Michael Woodford (1997), "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," *NBER*

Macroeconomics Annual 1997, 297-361.

Rudd, Jeremy, and Karl Whelan (2005), "New tests of the New Keynesian Phillips curve," *Journal of Monetary Economics* 52, 1167-1181.

Rudd, Jeremy, and Karl Whelan (2006), "Can rational expectations sticky-price models explain inflation dynamics?" *American Economic Review* 96, 303-320.

A Solving the Three-Sector Model

To solve this model, start from (7) and substitute in (8) and (9) to get

$$\overline{p}_t^* = (1-\omega)(1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k E_t mc_{t+k}^n + \omega(\overline{p}_{t-1}^* + \pi_{t-1}^s) \tag{A.1}$$

Pulling the period t terms from the infinite series

$$\overline{p}_t^* = (1-\omega)(1-\beta\theta)(mc_t + p_t) + (1-\omega)(\beta\theta)E_t p_{t+1}^f + \omega(\overline{p}_{t-1}^* + \pi_{t-1}^s) \tag{A.2}$$

where we have also used $mc_t^n = mc_t + p_t$. Now adding and subtracting $\omega(\beta\theta)E_t p_{t+1}^b$ gives

$$\begin{aligned}\overline{p}_t^* = {} & (1-\omega)(1-\beta\theta)(mc_t + p_t) + \omega(\overline{p}_{t-1}^* + \pi_{t-1}^s) + (1-\omega)(\beta\theta)E_t p_{t+1}^f \\ & + \omega(\beta\theta)E_t p_{t+1}^b - \omega(\beta\theta)E_t p_{t+1}^b\end{aligned} \tag{A.3}$$

Using (7) this becomes

$$\begin{aligned}\overline{p}_t^* = {} & (1-\omega)(1-\beta\theta)(mc_t + p_t) + \omega(\overline{p}_{t-1}^* + \pi_{t-1}^s) + (\beta\theta)E_t\overline{p}_{t+1}^* \\ & - \omega(\beta\theta)E_t p_{t+1}^b\end{aligned} \tag{A.4}$$

Now using the definition of p_t^b in (9)

$$\begin{aligned}\overline{p}_t^* = {} & (1-\omega)(1-\beta\theta)(mc_t + p_t) + \omega(\overline{p}_{t-1}^* + \pi_{t-1}^s) + (\beta\theta)E_t\overline{p}_{t+1}^* \\ & - \omega(\beta\theta)(\overline{p}_t^* + \pi_t^s)\end{aligned} \tag{A.5}$$

If we solve (6) for $\overline{p}_t^*$ we see that

$$\overline{p}_t^* = (1-\theta)^{-1}(p_t^s - \theta p_{t-1}^s) \tag{A.6}$$

Substituting this result in above we see

$$\begin{aligned}\overline{p}_t^* = {} & (1-\omega)(1-\beta\theta)(mc_t + p_t) + \omega(1-\theta)^{-1}(p_{t-1}^s - \theta p_{t-2}^s) + \omega\pi_{t-1} \\ & + (\beta\theta)(1-\theta)^{-1}(E_t p_{t+1}^s - \theta p_t^s) - \omega(\beta\theta)(1-\theta)^{-1}(p_t^s - \theta p_{t-1}^s) \\ & + \omega(\beta\theta)\pi_t^s\end{aligned} \tag{A.7}$$

Plugging this into (6)

$$\begin{aligned}p_t^s = {} & \theta p_{t-1}^s + (1-\theta)(1-\omega)(1-\beta\theta)(mc_t + p_t) + \omega(p_{t-1}^s - \theta p_{t-2}^s) \\ & + \omega(1-\theta)\pi_{t-1} + (\beta\theta)(E_t p_{t+1}^s - \theta p_t^s) - \omega(\beta\theta)(p_t^s - \theta p_{t-1}^s) \\ & + \omega(1-\theta)(\beta\theta)\pi_t^s\end{aligned} \tag{A.8}$$

Now we must express this in terms of *core* inflation ($\pi_t^s = p_t^s - p_{t-1}^s$). From (4) and (5) we see that

$$p_t = \alpha(\tilde{m}c_t + p_t^s) + (1-\alpha)p_t^s = p_t^s + \alpha\tilde{m}c_t \tag{A.9}$$

Note, incidentally, that if we had changed (5) to $\tilde{m}c_t + p_t$ rather than $\tilde{m}c_t + p_t^s$, then we would have gotten $p_t = \alpha\tilde{m}c_t + \alpha p_t + (1-\alpha)p_t^s$, or $p_t = \frac{\alpha}{1-\alpha}\tilde{m}c_t + p_t^s$, so the only effect would have been a slightly different coefficient on the (unobserved) shock $\tilde{m}c_t$. In any case, plugging $(A9)$ into $(A8)$ above and combining like terms gives

$$
\begin{aligned}
p_t^s = {} & [\theta + \omega + \omega\beta\theta^2]p_{t-1}^s + (1-\theta)(1-\omega)(1-\beta\theta)[mc_t + \alpha\tilde{mc}_t] \\
& + [(1-\theta)(1-\omega)(1-\beta\theta) - \beta\theta^2 - \omega\beta\theta]p_t^s - \omega\theta p_{t-2}^s + \beta\theta E_t p_{t+1}^s \\
& + \omega(1-\theta)\pi_{t-1}^s - \omega(1-\theta)\beta\theta\pi_t^s
\end{aligned}
\tag{A.10}
$$

which yields

$$
\begin{aligned}
[\theta + \omega + \beta\theta - \theta\omega + \beta\omega\theta^2]p_t^s = {} & [\theta + \omega + \omega\beta\theta^2]p_{t-1}^s + \beta\theta E_t p_{t+1}^s - \omega\theta p_{t-2}^s + \\
& (1-\theta)(1-\omega)(1-\beta\theta)[mc_t + \alpha\tilde{mc}_t] + \omega(1-\theta)\pi_{t-1}^s \\
& - \omega(1-\theta)\beta\theta\pi_t^s
\end{aligned}
\tag{A.11}
$$

Substracting $\beta\theta p_t^s$ from both sides

$$
\begin{aligned}
[\theta + \omega - \theta\omega + \beta\omega\theta^2]p_t^s = {} & [\theta + \omega + \omega\beta\theta^2]p_{t-1}^s + \beta\theta[E_t p_{t+1}^s - p_t^s] - \omega\theta p_{t-2}^s \\
& + (1-\theta)(1-\omega)(1-\beta\theta)[mc_t + \alpha\tilde{mc}_t] + \omega(1-\theta)\pi_{t-1}^s \\
& - \omega(1-\theta)\beta\theta\pi_t^s
\end{aligned}
\tag{A.12}
$$

Moving the p_{t-1}^s term to the left hand side and adding $\omega\theta p_{t-1}^s$ to both sides

$$
\begin{aligned}
[\theta + \omega - \theta\omega + \beta\omega\theta^2](p_t^s - p_{t-1}^s) = {} & \beta\theta E_t \pi_{t+1}^s + \omega\theta p_{t-1}^s - \omega\theta p_{t-2}^s \\
& + (1-\theta)(1-\omega)(1-\beta\theta)[mc_t + \alpha\tilde{mc}_t] + (\omega - \omega\theta)\pi_{t-1}^s \\
& - [\omega\beta\theta - \omega\beta\theta^2]\pi_t^s
\end{aligned}
\tag{A.13}
$$

Using $p_t^s - p_{t-1}^s = \pi_t^s$ on the left hand side and adding $\omega(1-\theta)\beta\theta\pi_t^s$ to both sides gives

$$[\theta+\omega-\theta\omega]\pi_t^s=\beta\theta E_t\pi_{t+1}^s+(1-\theta)(1-\omega)(1-\beta\theta)[mc_t+\alpha\tilde{mc}_t]+\omega\pi_{t-1}^s-\omega\beta\theta\pi_t^s \tag{A.14}$$

where we have also used $p_{t-1}^s-p_{t-2}^s=\pi_{t-1}^s$ and combined the π_{t-1}^s terms on the right hand side. Solving for π_t^s finally gives

$$\pi_t^s=\omega\phi^{-1}\pi_{t-1}+\beta\theta\phi^{-1}E_t\pi_{t+1}^s+(1-\theta)(1-\omega)(1-\beta\theta)\phi^{-1}[mc_t+\alpha\tilde{mc}_t] \tag{A.15}$$

where

$$\phi=\theta+\omega-\omega\theta+\omega\beta\theta=\theta+\omega(1-\theta(1-\beta)) \tag{A.16}$$

Table 1: Flexible Price CPI Items

Flexible Price Items	**Duration**	**Relative Importance**
Used cars and trucks	0.0	1.6
Motor fuel	0.5	3.2
Fuel oil and other fuels	1.2	0.3
Car and truck rental	1.2	0.1
Gas (piped) and electricity	1.7	4.2
Leased cars and trucks	1.8	0.6
Lodging away from home	2.6	2.5
Total, Flexible Price Items		**12.5**

Notes: We use a cutoff of three months of average price duration to classify all the 45 CPI categories into flexible and sticky price categories. The items listed in this and next tables are groupings of CPI entry level items (ELIs) for which inflation rates are regularly released by the U.S. Bureau of Labor Statistics. The second column reports the weighted average price duration in months that we calculate by using the CPI ELI level median price duration estimates in Table VI of Nakamura and Steinsson (2008). The third column reports the expenditure share (weights) in the CPI of each category.

Table 2: Sticky Price CPI Items

Sticy Price Items	Duration	Relative Importance
Fresh fruits and vegetables	3.6	0.9
Dairy and related products	3.7	0.9
Tobacco and smoking products	4.0	0.8
Meats, poultry, fish, and eggs	4.8	1.9
Motor vehicle maintenance and repair	6.0	1.2
Public transportation	6.1	1.1
Nonalcoholic beverages and beverage materials	8.3	1
Communication	8.8	3.2
Motor vehicle parts and equipment	8.9	0.4
New vehicles	9.1	4.5
Water and sewer and trash collection services	9.2	1
Processed fruits and vegetables	9.2	0.3
Cereals and bakery products	10.7	1.2
OER, Northeast Urban Region	11.0	5.3
OER, Midwest Urban Region	11.0	4.5
OER, South Urban Region	11.0	7.7
OER, West Urban Region	11.0	6.9
Motor vehicle insurance	11.8	2
Tenants' and household insurance	12.1	0.3
Alcoholic beverages	14.2	1.1
Other food at home	15.0	2
Miscellaneous personal goods	15.8	0.2
Education	15.8	3.1
Medical care commodities	16.9	1.6
Recreation	18.7	5.7
Household furnishings and operations	20.3	4.8
Rent of primary residence	20.9	6
Miscellaneous personal services	21.3	1.1
Food away from home	21.4	6.5
Medical care services	22.4	4.8
Personal care products	23.7	0.7
Jewelry and watches	25.1	0.4
Infants' and toddlers' apparel	27.8	0.2
Footwear	28.5	0.7
Men's and boys' apparel	30.0	0.9
Women's and girls' apparel	31.6	1.5
Personal care services	32.3	0.6
Motor vehicle fees	42.1	0.5
Total, Sticky Price Items		**87.5**
Total, non-OER Sticky Price Items		**63.1**

Table 3: GMM Estimates of the Hybrid NKPC and Tests for Weak- or Under-Identification for Alternative Instrument Sets: With CPI Inflation in the Regression and Lagged CPI Inflation as Instruments

	Instrumental Variables Set #								
	1	2	3	4	5	6	7	8	9
γ_b	.32	.02	.54	.49	.36	.73	.34	.45	.65
	(.05)	(.16)	(.12)	(.08)	(.10)	(.10)	(.12)	(.11)	(.14)
γ_f	.67	.96	.46	.51	.64	.27	.66	.55	.35
	(.05)	(.16)	(.12)	(.08)	(.10)	(.10)	(.12)	(.11)	(.13)
λ	.011	.022	-.005	.008	.014	-.019	.027	.000	-.016
	(.005)	(.010)	(.006)	(.005)	(.008)	(.010)	(.011)	(.013)	(.015)
AR $\chi^2(8)$	.00	.00	.00	.00	.00	.00	.00	.00	.00
AP $\chi^2(6)$	-	.00	.00	.00	.02	.05	.20	.94	.40
	.00	.00	.00	.00	.09	.05	.17	.14	.35
	.00	.00	.00	.00	.00	.00	.00	.00	.00
AP F	-	4.42	10.65	7.71	2.39	2.04	1.36	0.28	0.98
	33.87	7.14	5.58	4.56	1.73	2.00	1.45	1.54	1.05
	420.34	132.13	78.01	38.92	28.26	14.66	16.45	3.24	6.36
KP Wald rk F	29.36	2.62	3.39	2.99	2.91	1.45	0.78	1.14	1.52

Notes: The instrument variables set k consists of CPI inflation of k to $k+3$ lags, and output gap and real marginal cost of k and $k+1$ lags for $k = 1, ..., 9$. Standard errors are in parentheses. The row led by AR $\chi^2(8)$ reports the p-value of the Anderson and Rubin (1949) weak-identification robust test of the joint hypothesis that the slope coefficients are zero and the orthogonality conditions hold. The row led by AP $\chi^2(6)$ presents three p-values for the Angrist-Pischke (2009) test of under-identification, one for each endogenous regressor. The next row labeled AP F reports their test statistic of weak identification for each of endogenous regressors. The last row presents the Kleibergen and Paap (2006) Wald rk F statistic. The p-values of the last two tests in the table are currently unavailable. However, see the Empirical Results section for some suggestive thresholds for the two statistics.

Table 4: GMM Estimates of the Hybrid NKPC and Tests for Weak- or Under-Identification for Alternative Instrument Sets: With CPI Inflation in the Regression and Lagged Sticky Price Inflation as Instruments

	Instrumental Variables Set #								
	1	2	3	4	5	6	7	8	9
γ_b	.37	.27	.33	.41	.16	.77	.44	.49	.70
	(.06)	(.10)	(.08)	(.11)	(.21)	(.11)	(.09)	(.11)	(.15)
γ_f	.63	.72	.68	.59	.84	.23	.56	.52	.29
	(.06)	(.10)	(.08)	(.11)	(.21)	(.10)	(.09)	(.11)	(.14)
λ	.006	.010	-.005	.012	.023	-.014	.019	-.005	-.020
	(.005)	(.006)	(.007)	(.006)	(.012)	(.011)	(.010)	(.012)	(.018)
AR $\chi^2(8)$	.00	.00	.00	.00	.00	.00	.00	.00	.00
AP $\chi^2(6)$	.00	.00	.00	.00	.00	.03	.43	.39	.22
	.00	.00	.00	.09	.48	.04	.10	.00	.16
	.00	.00	.00	.00	.00	.00	.00	.00	.00
AP F	54.06	19.87	9.41	7.53	2.10	2.19	0.95	1.01	1.31
	10.05	8.62	9.36	1.72	0.88	2.06	1.67	3.06	1.47
	465.14	154.23	73.93	36.60	30.28	17.75	12.52	4.52	3.68
KP Wald rk F	12.19	4.56	5.07	2.05	0.72	1.42	0.96	1.80	1.57

Notes: The instrument variables set k consists of sticky price inflation of k to $k+3$ lags, and output gap and real marginal cost of k and $k+1$ lags for $k = 1, ..., 9$. Standard errors are in parentheses. The row led by AR $\chi^2(8)$ reports the p-value of the Anderson and Rubin (1949) weak-identification robust test of the joint hypothesis that the slope coefficients are zero and the orthogonality conditions hold. The row led by AP $\chi^2(6)$ presents three p-values for the Angrist-Pischke (2009) test of under-identification, one for each endogenous regressor. The next row labeled AP F reports their test statistic of weak identification for each of endogenous regressors. The last row presents the Kleibergen and Paap (2006) Wald rk F statistic. The p-values of the last two tests in the table are currently unavailable. However, see the Empirical Results section for some suggestive thresholds for the two statistics.

Table 5: GMM Estimates of the Hybrid NKPC and Tests for Weak- or Under-Identification for Alternative Instrument Sets: With Sticky Price Inflation in the Regression and Lagged Sticky Price Inflation as Instruments

	Instrumental Variables Set #								
	1	2	3	4	5	6	7	8	9
γ_b	.12	.64	.28	.49	.65	.52	.42	.47	.71
	(.08)	(.10)	(.24)	(.15)	(.08)	(.09)	(.12)	(.12)	(.10)
γ_f	.87	.35	.72	.51	.35	.49	.58	.52	.27
	(.08)	(.10)	(.24)	(.15)	(.08)	(.09)	(.12)	(.12)	(.09)
λ	.012	.000	.007	.006	.000	.004	.011	-.002	-.003
	(.006)	(.005)	(.009)	(.006)	(.005)	(.005)	(.009)	(.008)	(.010)
AR $\chi^2(8)$	.00	.00	.00	.00	.00	.00	.00	.00	.00
AP $\chi^2(6)$	-	.00	.00	.00	.00	.00	.00	.13	.00
	.00	.00	.02	.07	.00	.18	.00	.01	.16
	.00	.00	.00	.00	.00	.00	.00	.00	.00
AP F	-	5.65	4.85	5.42	3.91	4.22	4.95	1.58	3.97
	13.81	11.38	2.41	1.86	6.26	1.41	5.65	2.50	1.47
	460.82	156.49	74.97	42.13	31.45	17.18	12.70	5.12	3.66
KP Wald *rk* F	11.66	6.40	1.83	2.05	2.86	2.11	1.95	0.70	1.49

Notes: The instrument variables set k consists of sticky price inflation of k to $k+3$ lags, and output gap and real marginal cost of k and $k+1$ lags for $k = 1, ..., 9$. Standard errors are in parentheses. The row led by AR $\chi^2(8)$ reports the p-value of the Anderson and Rubin (1949) weak-identification robust test of the joint hypothesis that the slope coefficients are zero and the orthogonality conditions hold. The row led by AP $\chi^2(6)$ presents three p-values for the Angrist-Pischke (2009) test of under-identification, one for each endogenous regressor. The next row labeled AP F reports their test statistic of weak identification for each of endogenous regressors. The last row presents the Kleibergen and Paap (2006) Wald rk F statistic. The p-values of the last two tests in the table are currently unavailable. However, see the Empirical Results section for some suggestive thresholds for the two statistics.

Table 6: GMM Estimates of the Hybrid NKPC and Tests for Weak- or Under-Identification for Alternative Instrument Sets: With Sticky Price Inflation in the Regression and Lagged Sticky Price Inflaiton and Lagged PPI for Finished Goods Inflation as Instruments

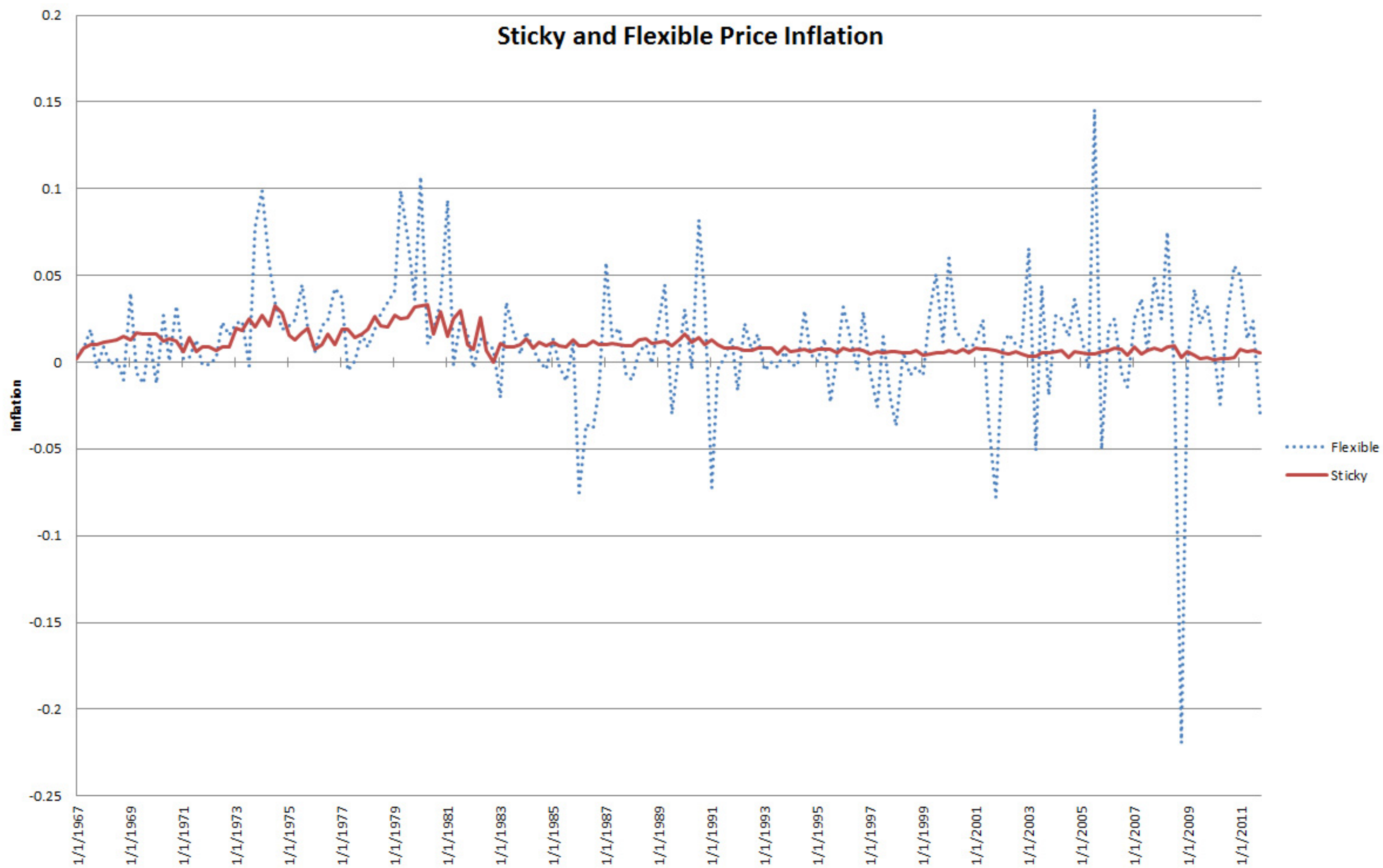
	Instrumental Variables Set #								
	1	2	3	4	5	6	7	8	9
γ_b	.21	.52	.39	.53	.55	.53	.59	.55	.55
	(.05)	(.07)	(.10)	(.06)	(.06)	(.08)	(.06)	(.07)	(.08)
γ_f	.78	.47	.61	.46	.45	.47	.41	.44	.44
	(.05)	(.07)	(.10)	(.06)	(.05)	(.08)	(.05)	(.07)	(.07)
λ	.012	.005	.004	.005	.005	.002	.002	-.002	.004
	(.005)	(.003)	(.004)	(.003)	(.004)	(.004)	(.004)	(.005)	(.005)
AR $\chi^2(8)$	.00	.00	.00	.00	.00	.00	.00	.00	.00
AP $\chi^2(6)$	-	.00	.00	.00	.00	.00	.00	.00	.00
	.00	.00	.00	.00	.00	.00	.00	.00	.00
	.00	.00	.00	.00	.00	.00	.00	.00	.00
AP F	-	8.32	6.14	5.95	6.86	6.44	7.95	3.46	2.76
	26.99	9.20	5.05	3.56	6.16	3.87	7.37	7.31	4.05
	296.87	97.41	49.69	33.86	20.89	28.75	8.73	2.55	2.59
KP Wald rk F	24.91	3.92	2.94	2.24	3.61	3.88	5.09	4.39	2.13

Notes: The instrument variables set k consists of sticky price inflation and PPI for Finished Goods inflation of k to $k+3$ lags, and output gap and real marginal cost of k and $k+1$ lags for $k = 1, ..., 9$. Standard errors are in parentheses. The row led by AR $\chi^2(8)$ reports the p-value of the Anderson and Rubin (1949) weak-identification robust test of the joint hypothesis that the slope coefficients are zero and the orthogonality conditions hold. The row led by AP $\chi^2(6)$ presents three p-values for the Angrist-Pischke (2009) test of under-identification, one for each endogenous regressor. The next row labeled AP F reports their test statistic of weak identification for each of endogenous regressors. The last row presents the Kleibergen and Paap (2006) Wald rk F statistic. The p-values of the last two tests in the table are currently unavailable. However, see the Empirical Results section for some suggestive thresholds for the two statistics.

Table 7: GMM Estimates of the Hybrid NKPC and Tests for Weak- or Under-Identification for Alternative Instrument Sets: With CPI Inflation in the Regression and Lagged Sticky Price Inflaiton and Lagged PPI for Finished Goods Inflation as Instruments

	Instrumental Variables Set #								
	1	2	3	4	5	6	7	8	9
γ_b	.44	.39	.38	.53	.57	.55	.41	.50	.69
	(.04)	(.05)	(.05)	(.04)	(.06)	(.06)	(.06)	(.08)	(.08)
γ_f	.56	.60	.62	.47	.43	.44	.58	.49	.31
	(.04)	(.05)	(.06)	(.04)	(.05)	(.05)	(.06)	(.07)	(.08)
λ	.003	.006	-.001	.006	.005	-.004	.015	-.002	-.012
	(.004)	(.005)	(.005)	(.005)	(.005)	(.006)	(.007)	(.007)	(.011)
AR $\chi^2(8)$	.00	.00	.00	.00	.00	.00	.00	.00	.00
AP $\chi^2(6)$	.00	.00	.00	.00	.00	.00	.00	.00	.00
	.00	.00	.00	.00	.00	.00	.00	.00	.00
	.00	.00	.00	.00	.00	.00	.00	.00	.00
AP F	33.81	24.71	12.00	15.57	10.58	7.04	4.68	3.91	4.04
	29.78	18.46	8.87	9.54	3.81	5.41	5.85	6.83	5.68
	295.00	93.92	54.67	29.84	22.34	13.48	9.56	2.48	2.39
KP Wald rk F	10.96	5.22	4.75	4.20	4.46	2.89	3.88	3.30	3.18

Notes: The instrument variables set k consists of sticky price inflation and PPI of Finished Goods inflation of k to $k+3$ lags, and output gap and real marginal cost of k and $k+1$ lags for $k = 1, ..., 9$. Standard errors are in parentheses. The row led by AR $\chi^2(8)$ reports the p-value of the Anderson and Rubin (1949) weak-identification robust test of the joint hypothesis that the slope coefficients are zero and the orthogonality conditions hold. The row led by AP $\chi^2(6)$ presents three p-values for the Angrist-Pischke (2009) test of under-identification, one for each endogenous regressor. The next row labeled AP F reports their test statistic of weak identification for each of endogenous regressors. The last row presents the Kleibergen and Paap (2006) Wald rk F statistic. The p-values of the last two tests in the table are currently unavailable. However, see the Empirical Results section for some suggestive thresholds for the two statistics.


Sticky and Flexible Price Inflation
Inflation
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
-0.25
1/1/1967
1/1/1969
1/1/1971
1/1/1973
1/1/1975
1/1/1977
1/1/1979
1/1/1981
1/1/1983
1/1/1985
1/1/1987
1/1/1989
1/1/1991
1/1/1993
1/1/1995
1/1/1997
1/1/1999
1/1/2001
1/1/2003
1/1/2005
1/1/2007
1/1/2009
1/1/2011
Flexible
Sticky