Queen's Economics Department Working Paper No. 1026

Identifying the New Keynesian Phillips Curve

James M. Nason
Federal Reserve Bank of Atlanta

Gregor W. Smith
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

1-2005

Identifying the New Keynesian Phillips Curve

James M. Nason and Gregor W. Smith†

January 2005

Abstract

Phillips curves are central to discussions of inflation dynamics and monetary policy. New Keynesian Phillips curves describe how past inflation, expected future inflation, and a measure of real marginal cost or an output gap drive the current inflation rate. This paper studies the (potential) weak identification of these curves under GMM and traces this syndrome to a lack of persistence in either exogenous variables or shocks. We employ analytic methods to understand the identification problem in several statistical environments: under strict exogeneity, in a vector autoregression, and in the canonical three-equation, New Keynesian model. Given U.S., U.K., and Canadian data, we revisit the empirical evidence and construct tests and confidence intervals based on exact and pivotal Anderson-Rubin statistics that are robust to weak identification. These tests find little evidence of forward-looking inflation dynamics.

JEL classification: E31; C32.

Keywords: Phillips curve, Keynesian, identification, inflation.

†J.M. Nason: Research Department, Federal Reserve Bank of Atlanta; (404) 498-8891, jim.nason@atl.frb.org. G.W. Smith: Department of Economics, Queen's University; (613) 533-6659, smithgw@qed.econ.queensu.ca. We thank the Social Sciences and Humanities Research Council of Canada for support of this research. Smith thanks the Research Department of the Federal Reserve Bank of Atlanta for providing the environment for this research. Richard Luger and Katharine Neiss provided data for Canada and the U.K. respectively, while Nikolay Gospodinov and Amir Yaron shared their code. Helpful comments were provided by Richard Clarida, Roger Farmer, Jeffrey Fuhrer, Allan Gregory, Eric Leeper, Jesper Lindé, Thomas Lubik, Antonio Moreno, Charles Nelson, Michel Normandin, Adrian Pagan, Thomas Sargent, Christoph Schleicher, Tao Zha, and seminar participants at the Canadian Economics Association meetings, the Federal Reserve Bank of Atlanta, the European Central Bank's 3rd Workshop on Forecasting Techniques, and the Bank of Canada/UBC/SFU Macroeconomics Workshop. The views in this paper represent those of the authors and are not those of either the Federal Reserve Bank of Atlanta, the Federal Reserve System, or any of its staff.

1. Introduction

Recent years have witnessed a boom in work on the Phillips curve. For a student of monetary policy and the business cycle steeped in dynamic general equilibrium methods, the revival of Phillips curve research might come as a shock. The shock might be mitigated because the Phillips curve revival features debates on the role of backward- and forward-looking expectations for inflation, on which measure of real aggregate demand most directly influences inflation, on the response of monetary policy to various disturbances, and on the costs of disinflation. These debates often are framed by the new Keynesian Phillips curve (NKPC) because it appears to provide a tent under which many views of inflation dynamics can exist. However, whether to be inside or outside the Phillips curve revival tent depends on the NKPC being a persuasive description of inflation dynamics.

Variations on the NKPC are just about limitless. The canonical NKPC is driven either by current real marginal cost or today's output gap and is forward-looking in the current expectation of tomorrow's inflation. Gali and Gertler (1999) add lagged inflation to create a 'hybrid NKPC', which they use to address aspects of the debate among Phillips curve revivalists. Specification of the NKPC has important implications for monetary policy, and in particular for how central banks should react to real events while maintaining inflation targets. Although contributions to this research are too numerous to list, besides Galí and Gertler (1999), Fuhrer and Moore (1995), Roberts (1995) and Sbordone (2002) make important empirical contributions. Theory and evidence about the NKPC also are reviewed by Woodford (2003).

The hybrid NKPC is a second-order, linear, expectational difference equation. Its earliest guise is as a labor demand schedule; see Kennan (1979). Hansen and Sargent (1980) and Sargent (1987) study the dynamic and time series properties of this general class of stochastic models. Most empirical work on the NKPC estimates it using instrumental variables (IV) methods, as Galí and Gertler (1999) do. Generally, NKPC parameters prove difficult to pin down without large instrument sets. This suggests weak identification. Other symptoms of this syndrome include instability of estimates across instrument sets, estimates which may approach those from ordinary least-squares and hence be inconsistent, and Wald tests with size distortions. The goal of this paper is to study the economics underlying weak identification, with a view to drawing lessons and recommendations for applied work.

In section 2, we study identification analytically in a solved version of the hybrid NKPC difference equation. In this environment, the process for real marginal cost or an output gap (labelled x) that drives inflation, π_t, is *strictly* exogenous. The main finding is that identification requires higher-order dynamics in x. We also illustrate the weaker identification requirements of system estimators, which may be feasible with less persistence. Section 2 also discusses identification in IV estimators with the purely forward-looking NKPC, with calibrated discount factors, with cointegrated variables, and with lagged instrument sets.

Section 3 sets the hybrid NKPC in a VAR in $\{\pi_t, x_t\}$. We show this generalization fails to make identification easier unless higher-order lags of inflation predict real marginal cost. The reason is the investigator must take care to separate the two roles once-lagged inflation plays: (a) it enters the hybrid NKPC to reflect slow price adjustment, and (b) it enters the VAR because it helps forecast future values of x.

Section 4 details the identification problems when the hybrid NKPC is set in a typical, three-equation, new Keynesian model. The hybrid NKPC cannot be identified under IV estimation in the baseline version of this model. For the hybrid NKPC to be identified requires that either (a) one of the shocks to the system is persistent or (b) the interest-rate rule involves a lagged interest rate (interest-rate smoothing).

Section 5 applies the results to the U.S., U.K., and Canada. We first estimate the hybrid NKPC for each country, using a range of instruments. We also investigate a necessary condition for identification: π_{t+1} must be predictable using information other than π_t, π_{t-1}, and x_t. We relate the findings from this first-stage test to the literature on forecasting inflation. Finally, we use the Anderson and Rubin (1949) statistic to test the hybrid NKPC. This test is exact and robust to weak or omitted instruments. Its application yields little evidence of forward-looking inflation dynamics.

2. Identification with Strict Exogeneity

A variety of pricing environments give rise to a hybrid NKPC that describes inflation, π_t:

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

where we use x_t to denote real aggregate demand (either real marginal cost or an output gap). The studies by Rotemberg (1982), Roberts (1997), Fuhrer and Moore (1995), Yun (1996), and Galí and Gertler (1999) contain influential examples of these environments. The underlying pricing behavior can range from smooth adjustment with quadratic costs to a variation of Calvo's contract model (with or without firm-specific capital) in which some price-setters are backward-looking. The hybrid NKPC (1) also may be consistent with the dynamic indexing model suggested by Woodford (2003), assuming it is written in the change in inflation rather than the level.

Our study is concerned with identifying the parameters γ_f, γ_b, and λ, rather than with working backward from them to the underlying structural ones. Throughout the paper we assume (with one exception) that the roots of relevant difference equations imply stability and uniqueness of solutions, and that the difference equation (1) follows from a pricing model – in which all three parameters are positive – and not an observationally equivalent environment, as in Beyer and Farmer (2004).

The hybrid NKPC (1) is a linear, second-order, stochastic difference equation. Our study draws on tools for formulating these problems under rational expectations developed by Hansen and Sargent (1980) and Sargent (1987). We also draw on studies of estimation in the linear-quadratic model by Gregory, Pagan, and Smith (1993), West and Wilcox (1994) and Fuhrer, Moore, and Schuh (1995).

We begin by reviewing the identification of the parameters in several different statistical frameworks. Given the popularity of the IV (*i.e.*, GMM) estimator, we focus principally on those methods. Our approach adopts a linear statistical model for x_t, and then solves for inflation, π_t. Using the solved (full-information) model, we describe several different GMM (limited-information) estimators.

We consider the two classic properties of instrument sets. Obviously, identifying the three parameters of the hybrid NKPC first requires at least three instruments or, more generally, three pieces of identifying information which could include restrictions on the parameters or covariance restrictions in a system setting. A test based on over-identification requires at least four instruments or four such pieces of information. The instruments must be uncorrelated with the GMM residuals, which are essentially forecast errors. This is the *order* condition. Second, the matrix of cross-products of the instruments and the right-hand-side variables in the hybrid NKPC cannot be singular. This is the *rank* or 'relevance' condition.

Each property is illustrated using our model of x_t. Our environment is linear, so there is no distinction between local and global identification. The order and rank conditions provide results that (1) imply that higher-order dynamics in x_t often are necessary for identification, (2) yield an analysis of situations in which weak identification can arise, (3) suggest that additional parameter information or restrictions on x (*e.g.*, x and π are cointegrated) may not aid identification in GMM estimation, (4) show that partly solving the hybrid NKPC forward does not improve identification, (5) derive an expression for the loss of precision in the hybrid NKPC caused by using only lagged instruments, and (6) show that lagged residuals are not valid instruments. Our analysis of the identification of the hybrid NKPC provides guidance for studying it in richer environments in sections 3 and 4 and for empirical work in section 5. The key analytic results are summarized in table 1.

We uncover the properties of hybrid NKPC estimators by solving the difference equation (1) using the methods of Sargent (1987):

$$\pi_t = \delta_1 \pi_{t-1} + \left(\frac{\lambda}{\delta_2 \gamma_f} \right) \sum_{k=0}^{\infty} \left(\frac{1}{\delta_2} \right)^k E_t x_{t+k}, \qquad (2)$$

where δ_1 and δ_2 are the stable and unstable roots, respectively, of the characteristic equation:

$$-L^{-1} + \frac{1}{\gamma_f} - \frac{\gamma_b L}{\gamma_f} = 0. \qquad (3)$$

We assume that $\{x_t\}$ is of exponential order less than δ_2 so that the infinite sum in (2) is finite, and that the roots yield a unique solution to the difference equation.

Suppose that x_t evolves autonomously according to a J-th order autoregression:

$$x_t = \sum_{j=1}^{J} \rho_j x_{t-j} + \epsilon_t, \tag{4}$$

where $\rho_j \neq 0 \ \forall \ j$ and ϵ_t is an innovation with respect to the σ-field generated by the history of x. This process can be rewritten in companion form as:

$$\tilde{x}_t = \tilde{\rho}\tilde{x}_{t-1} + \tilde{\epsilon}_t, \tag{5}$$

where $\tilde{x}_t = (x_t \ x_{t-1} \ \ldots \ x_{t-J+1})'$ and the transition matrix is:

$$\tilde{\rho} = \begin{pmatrix} \rho_1 & \rho_2 & \ldots & \rho_J \\ I_{J-1} & & & 0_{J-1} \end{pmatrix}, \tag{6}$$

where 0_{J-1} is a column vector of zeros. Next, define s_J as a selection row vector of length J with 1 in the first position and zeros thereafter. It will select the first element of $\tilde{x}_t$. Define I_J as the $J \times J$ identity matrix. The solution for inflation follows:

$$\pi_t = \delta_1 \pi_{t-1} + \left(\frac{\lambda}{\gamma_f \delta_2}\right) s_J [I_J - \tilde{\rho}\delta_2^{-1}]^{-1} \tilde{x}_t + \eta_t. \tag{7}$$

We assume that

$$|I_J - \tilde{\rho}\delta_2^{-1}| \neq 0. \tag{8}$$

The stochastic singularity is avoided – so that a residual η_t appears in the solution (7) – by assuming that the econometrician's information set lies strictly within that of the price-setting agents, as originally proposed by Hansen and Sargent (1980). Thus, η_t is uncorrelated with information available to the econometrician at time t. In particular, if the econometrician has access to current and past values of x then:

$$\text{cov}(\eta_t, \epsilon_t) = 0. \tag{9}$$

Alternately, η_t can be interpreted as a cost, technology, or real aggregate demand shock; see Ireland (2002) for a discussion.

We study the macroeconomic implications of identification of the hybrid NKPC in this environment. This quest excludes other potential sources of identification, such as structural breaks, varying conditional covariances, or the use of survey data on inflation expectations. We have omitted constant terms, as if the data have been demeaned. Of course, if in applications a constant term is included in the NKPC, a vector of ones can be used as an instrument while adding no net identifying information.

Combine the x-process (5) with the solved hybrid NKPC (7) to describe a structural VAR (SVAR) with cross-equation restrictions:

$$\begin{aligned} \tilde{x}_t &= \tilde{\rho}\tilde{x}_{t-1} + \tilde{\epsilon}_t \\ \pi_t &= \delta_1 \pi_{t-1} + \left(\frac{\lambda}{\gamma_f \delta_2}\right) s_J \left[I_J - \tilde{\rho}\delta_2^{-1}\right]^{-1} \tilde{x}_t + \eta_t. \end{aligned} \tag{10}$$

Result 1. The hybrid NKPC imposes the King and Watson (1994) real business cycle identification on the structural VAR (10) for $\{\tilde{x}_t, \pi_t\}$ and the Solow-Gordon identifying assumption on the impact matrix of the unrestricted simultaneous equations system of $\{\tilde{x}_t, \pi_t\}$. §

The structural VAR (SVAR) of (10) is identified by the fact that current inflation has no impact on x_t. Shock innovations to the hybrid NKPC and the autonomous process for x drive the inflation rate. Marginal cost or the output gap, x, responds only to one shock, ϵ_t. Thus, ϵ_t is an autonomous shock with respect to real aggregate demand. King and Watson (1994) impose the restrictions of the SVAR (10) to achieve their real business cycle (RBC) identification, while King and Watson (1997) refer to the impact restriction of this SVAR as the Solow-Gordon Phillips curve identification. The former identification agrees with RBC theory, according to King and Watson (1994), because the measure of real aggregate demand is independent of the inflation shock innovation and the history of inflation. The NKPC-SVAR of (10) also is consistent with the Solow-Gordon Phillips curve interpretation that real rigidities dominate aggregate demand fluctuations and inflation dynamics. Result 1 implies that fundamental shocks produced by the hybrid NKPC-SVAR (10) will be indistinguishable from those of either the RBC identification or the Solow-Gordon Phillips curve identification.

King and Watson (1997) observe that the SVAR (10) restricted by the Solow-Gordon identification is inconsistent with the notion of price stickiness. Since x_t enters the solved inflation process, inflation responds to η_t and ϵ_t at impact (*i.e.* lag zero). Thus, real and nominal shocks generate movements in inflation at impact, under the Solow-Gordon identification implied by the hybrid NKPC. It is the cross-equation restrictions of the hybrid NKPC SVAR (10) that yield the additional information for estimation and testing. The quandary remains that a model predicated on costly price setting requires inflation to be flexible enough to respond to all shocks at impact.

A priori there is no 'best' way to estimate the solved hybrid NKPC (7). It is one regression in a system that includes the AR(J) process (5) of x. This system also is defined by cross-equation restrictions including one on the covariance matrix of forecast innovations $\{\epsilon_t, \eta_t\}$.

Result 2. The number of regressors in (7) is $J + 1$. The parameters in $\tilde{\rho}$ can be identified from estimation of the law of motion for x_t, (4). With three parameters $\{\gamma_f, \gamma_b, \lambda\}$ to identify, $J \geq 1$ is necessary for identification in the solved model (7). $J \geq 2$ is necessary for overidentification.§

The key logic behind this result is that system estimation of the bivariate system allows (or requires) the econometrician to impose the covariance restriction (9). Thus only two additional pieces of information are required from the solution for inflation (7), and there are two regressors as long as $J \geq 1$. In general, identification is possible if the present value in the solved model (2) has a non-null projection on at least one variable known by price-setters at time t. In our case, these variables will be elements in $\tilde{x}_t$, but other variables might contribute as well. Studies that use the system estimator include, among others, Fuhrer and Moore (1995), Sbordone (2002), Kurmann (2003a), Lindé (2002), Bardsen, Jansen, and Nymoen (2002), Jondeau and Le Bihan (2003), and Fuhrer and Olivei (2004).

More typical is GMM estimation of the hybrid NKPC (1), using sample versions of:

$$E[\gamma_f \pi_{t+1} - \pi_t + \gamma_b \pi_{t-1} + \lambda x_t | z_t] = 0, \tag{11}$$

and instruments z_t. Given moment conditions (11), a necessary condition for identification of $\{\gamma_b, \gamma_f, \lambda\}$ is that there are as many *valid* instruments as parameters (or variables that explain inflation in this linear model). Of course, being dated $t-1$ or earlier is not sufficient for an instrument to be valid: it must possess *incremental* information about π_{t+1}. This is the 'relevance' condition of IV estimation.

Result 3. If $z_t = \{\pi_{t-1}, x_t, x_{t-1}, x_{t-2}, \ldots, x_{t-J+1}\}$, then $J \geq 2$ is necessary for identification by GMM and $J \geq 3$ is necessary for overidentification.§

According to the solution of the present value of the hybrid NKPC, equation (7) shows that further lags of inflation contain no identifying information, so z_t is the maximal instrument set in this environment. Observe that $\dim(z_t) = J + 1$ and the result follows. For example, let $J = 2$, then $z_t = \{\pi_{t-1}, x_t, x_{t-1}\}$, because x_{t-2} contains no additional information.

Moving from estimation of the solved model (5) and (7) to the difference equation (11) and ignoring information on the properties of x_t cannot ease the conditions for identification. Result 3 shows that identification under GMM is strictly more onerous than in the system environment of Result 2 because the error-covariance restriction (9) is no longer available. This difference must be considered prior to considering the usual trade-off between efficiency and robustness in deciding between system and single-equation estimation. In particular, the parameters of the second-order difference equation in inflation (1) cannot be identified by GMM, if x_t follows a first-order Markov process. Pesaran (1987, Propositions 6.1 and 6.2) derived similar results. He observed that identifying information is available when the lag length in the process for x_t is longer than that in the difference equation.

Results 2 and 3 rationalize the common practice of imposing a value for or calibrating β, a discount factor that underlies $\{\gamma_f, \gamma_b\}$. For example, β sometimes is set to 0.99 in quarterly data, which implies a quarterly discount rate of about 1 percent. This procedure allows identification when $\{x_t\}$ follows a Markov process. Other studies impose $\gamma_f = 1 - \gamma_b$, which again aids identification, but makes it impossible to test the hybrid NKPC against the purely forward-looking one $(\gamma_b = 0)$.

A number of researchers have used only lagged instruments in estimating (11). For example, Galí and Gertler (1999) used up to four lags of various instruments. Let us denote this information set by z_{t-1}.

Result 4. If $z_{t-1} = \{\pi_{t-1}, x_{t-1}, x_{t-2}, \ldots, x_{t-J}\}$, so that only lagged information is used, then again $J \geq 2$ is necessary for identification by GMM and $J \geq 3$ is necessary for overidentification.§

The intuition for Result 4 is that the moment conditions (11) involve forecasts of π_{t+1}, π_t, and x_t based on information at time $t-1$. Notice that z_{t-1} is *not* a subset of z_t. Again $\dim(z_{t-1}) = J+1$ and the identification result follows.

As an example, suppose that x_t follows a second-order autoregression, so $J = 2$. Then $z_t = \{\pi_{t-1}, x_t, x_{t-1}\}$ and $z_{t-1} = \{\pi_{t-1}, x_{t-1}, x_{t-2}\}$. Omitting the current value of x_t as an instrument means that an additional, lagged value must be used and be relevant. If instead z_t is the instrument set, then including x_{t-J} (x_{t-2} in this example) provides no overidentifying information.

In some circumstances, the investigator may know the value of λ, either from theory or from some auxiliary statistical work. For example, if $J = 1$ and $\rho_1 = 1$ then x_t and π_t will be cointegrated with parameter λ, which could be estimated from a static regression, as originally proposed by Granger and Engle (1987). This information can potentially aid identification of the remaining parameters, γ_f and γ_b.

Result 5. If a consistent estimate $\hat{\lambda}$ is available, $J \geq 1$ is necessary for the identification of γ_f and γ_b in the solved-system environment. In the single-equation environment with instruments z_t, $J \geq 1$ is necessary for identification and $J \geq 2$ for overidentification. With instruments z_{t-1}, however, $J \geq 2$ remains necessary for identification and $J \geq 3$ for overidentification.§

To see this result, consider $J = 1$. In this case, the solved model yields two coefficients and a covariance restriction, which over-identify the two remaining parameter estimates (with ρ_1 estimated in the auxiliary model). Similarly, with $\hat{\lambda} x_t$ known in the difference equation, the instruments x_t and π_{t-1} can be used to identify γ_f and γ_b.

But with instruments z_{t-1} three variables in (12) remain to be forecasted, $\{\pi_{t+1}, \pi_t, x_t\}$, even given an estimate $\hat{\lambda}$. Thus, a two-step procedure cannot identify the two other parameters, unless $J \geq 2$ continues to hold.

The last part of Result 5 is a generalization of an example found in Pagan, Gregory, and Smith (1993). They consider the case with $\rho_1 = 1$; also see West (1988). According to Pagan, Gregory, and Smith, lagged instruments could not identify the parameters of the difference equation without higher-order dynamics in the x-process. Result 5 also is relevant to price-setting rules that are written in terms of the level of prices, rather than the inflation rate, because the price level is more likely to be nonstationary yet cointegrated with the fundamental; see Nason and Slotsve (2004) for an example.

Result 6. The conditions for identification do not change if the investigator imposes $\gamma_b = 0$, so that the NKPC is purely forward-looking.§

This result can be checked by specializing the solution in (7), with $\delta_1 = 0$ and $\delta_2 = \gamma_f^{-1}$ which follow from the roots of (3). Again we assume that the remaining two parameters yield a unique solution to the difference equation. Note that the investigator has dropped a parameter, γ_b, and the variable π_{t-1} also. Mavroeidis (2004a, b) provides a discussion of this case.

As an interesting way to provide evidence on the hybrid NKPC, Rudd and Whelan (2001), Galí, Gertler, and López-Salido (2001), and Guay, Luger, and Zhu (2002) solve the hybrid NKPC difference equation forward, as in (2), but truncate after K leads. This leads them to estimate by instrumental variables:

$$E_{t-1}(\pi_t - \delta_1 \pi_{t-1} - \frac{\lambda}{\delta_2 \gamma_f} \sum_{k=0}^{K} \delta_2{}^{-k} x_{t+k}). \tag{12}$$

Result 7. Solving forward and truncating provides no additional information to aid identification (or improve efficiency).§

This result is obvious given Result 2. The difference equation – solved forward and truncated – still involves the three parameters $\{\gamma_f, \gamma_b, \lambda\}$. Were there valid instruments for each future x_{t+k} in (12), these parameters would be overidentified because

(13) contains more variables than parameters when $K \geq 1$. Nonetheless, the number of relevant instruments remains $J+1$, so the conditions for identification are unchanged.

Result 8. Whether z_t or z_{t-1} is adopted, the GMM residual is a MA(1) process. Both of these instrument sets are valid, but any instrument set must exclude lagged GMM residuals. In addition, the loss of precision from excluding x_t from the instrument set depends both on parameters in its law of motion and on the hybrid NKPC parameters.§

The GMM residual is given by:

$$\nu_{t+1} \equiv (\gamma_f \pi_{t+1} - \pi_t + \gamma_b \pi_{t-1} + \lambda x_t) - E_t(\gamma_f \pi_{t+1} - \pi_t + \gamma_b \pi_{t-1} + \lambda x_t). \tag{13}$$

With z_t, the residual is:

$$\nu_{t+1|t} = \gamma_f \eta_{t+1} + (\delta_1 \gamma_f - 1)\eta_t + \left(\frac{\lambda}{\gamma_f \delta_2}\right) \tilde{s}_J [I_J - \tilde{\rho}\delta_2^{-1}]^{-1} \tilde{\epsilon}_{t+1}. \tag{14}$$

This moving average can be accounted for in constructing the weighting matrix in GMM estimation. If z_{t-1} is adopted, the residual is:

$$\nu_{t+1|t-1} = \nu_{t+1|t} + \tilde{s}_J \left[(\frac{\lambda}{\gamma_f \delta_2}) [\gamma_f(\gamma_b + \tilde{\rho}) - 1][I_J - \tilde{\rho}\delta_2^{-1}]^{-1} - \lambda \right] \tilde{\epsilon}_t. \tag{15}$$

so that the variance of the additional term – and hence the efficiency loss – depends on the parameters of the hybrid NKPC in addition to those of the $\{x_t\}$ process.

The key, analytical findings of this section are that (a) identification may be easier in the system context than in the GMM context; and (b) in either case, higher-order dynamics in real marginal cost, unemployment, or the output gap are necessary in order to test the theory. We next examine whether these lessons change when the hybrid NKPC is set in other statistical environments.

3. VAR Identification

This section generalizes the environment by allowing lagged inflation to enter the law of motion of x_t. Of course, other variables also might help forecast real marginal cost or the output gap. Including the lagged, endogenous variable in the law of motion for x may partly capture the additional information used by price-setters in forecasting. Campbell and Shiller (1987) Boileau and Normandin (2002), and Kurmann (2003a) develop this approach.

Suppose we add lagged inflation to the process generating x_t:

$$x_t = \sum_{j=1}^{J} \rho_j x_{t-j} + \sum_{j=1}^{J} \zeta_i \pi_{t-i} + \epsilon_t, \tag{16}$$

so that we are agnostic about whether lagged inflation helps forecast marginal cost or the output gap. Combine the forecasting rule (16) with equation (2), the present-value version of the hybrid NKPC:

$$\pi_t = \delta_1 \pi_{t-1} + \left(\frac{\lambda}{\delta_2 \gamma_f}\right) \sum_{k=0}^{\infty} \left(\frac{1}{\delta_2}\right)^k E_t x_{t+k},$$

to solve the model. It is not necessary to extend all the algebra of section 2, though, for clearly any variable that helps to forecast $E_t x_{t+k}$ will be a linear function of the information set
$Z_t = \{x_t, x_{t-1}, ...x_{t-J+1}, \pi_t, \pi_{t-1}, ...\pi_{t-J+1}\}$. As usual the lag length in the solution is one less than that in the forecasting equation (16). The solution for inflation thus will involve these variables, along with π_{t-1}.

Result 9. Predicting x with once-lagged or twice-lagged inflation adds no identifying information. $J \geq 3$ is necessary for the VAR to add overidentifying information. Thus Results 2 and 3 continue to apply within the VAR. §

Result 2 showed that the system with x following a first-order autoregression is just identified. Being able to predict x with further lags allows over-identification. Each added lag of x introduces two new projection coefficients (one in each equation) but only one new parameter. Instead, suppose that the investigator predicts x with once-lagged or twice-lagged inflation in the hope of providing over-identification. Nonetheless, the system remains just-identified because current and once-lagged inflation already enter the hybrid NKPC.

Similarly, relevant instruments for $E_t\pi_{t+1}$ in GMM estimation now will be $Z_t = \{x_t, x_{t-1}, ...x_{t-J+1}, \pi_t, \pi_{t-1}, ...\pi_{t-J+1}\}$. The NKPC already includes π_t and π_{t-1} and so lags of inflation add instruments only if $J \geq 3$.

It is important to note that the coefficient on lagged inflation in the solved Phillips curve now has a different interpretation. In section 2 the coefficient on lagged inflation in the solution (2), δ_1, depended only on the parameters of the Phillips curve, γ_b and γ_f, as shown in the characteristic equation (3). In the VAR – with lagged inflation potentially forecasting future values of x – this separation no longer holds.

Result 10. The coefficient on π_{t-1} in the solved hybrid NKPC is independent of the process followed by real marginal cost *iff* inflation does not Granger-cause real marginal cost.§

Granger-causality from π to x often is viewed as a weak implication of the NKPC because it involves no cross-equation restrictions. Result 10 notes that in this case the coefficient on π_{t-1} reflects structural parameters *and* the forecasting rule for x. For example, $\gamma_b = 0$ does not imply that the coefficient on lagged inflation will also be 0, for lagged inflation could forecast future values of x – and so enter the inflation solution – even if there is no backward-looking price-setting. An investigator who incorrectly assumes that x is strictly exogenous will deduce incorrect (*i.e.*, biased) values of γ_f and γ_b when performing system estimation. Kennan (1979) first showed that the intrinsic dynamics (γ_b and γ_f) could be estimated consistently by single-equation least squares, provided sufficient lags in x are included to capture the forecasting information. Result 10 is also based on Sargent (1987, chapter XI, part 24), who showed the relationship between strict exogeneity – in the classic terminology of Engle, Hendry and Richard (1983) – and Granger-causality.

This discussion raises the question of the economic interpretation of these lags of inflation. The next section turns to an environment which restricts the VAR with additional economic theory.

4. Identification in a New Keynesian System

Up to this point, we have discovered (or rediscovered) that identifying the hybrid NKPC depends on the properties of the x-process. However, real marginal cost or the output gap is endogenous in a dynamic, stochastic, general-equilibrium model. We study identification in a more complete model in this section. It seems natural to work with a typical, new Keynesian trinity model (NKTM):

$$\begin{aligned}
\pi_t &= \gamma_f E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \lambda y_t + \epsilon_{\pi t} \\
y_t &= \beta_f E_t y_{t+1} + \beta_b y_{t-1} - \beta_R (R_t - E_t \pi_{t+1}) + \epsilon_{yt} \\
R_t &= \omega_\pi \pi_t + \omega_y y_t + \epsilon_{Rt}
\end{aligned} \tag{17}$$

where y is the output gap, R is the central bank's discount rate (the nominal federal funds rate in the U.S.), the second equation is a linearized dynamic IS schedule, and the last equation is a Taylor rule.

Our interest is in estimating the hybrid NKPC by replacing $E_t \pi_{t+1}$. We derive the forecasting implications of the NKTM (17) to do this. Using the policy rule to replace the interest rate in the equations for inflation and the output gap gives:

$$\begin{aligned}
\pi_t &= \gamma_f E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \lambda y_t + \epsilon_{\pi t} \\
y_t &= \beta_f \varphi E_t y_{t+1} + \beta_R \varphi E_t \pi_{t+1} + \beta_b \varphi y_{t-1} - \beta_R \omega_\pi \varphi \pi_t + \varphi(\epsilon_{yt} - \beta_r \epsilon_{Rt})
\end{aligned} \tag{18}$$

where

$$\varphi \equiv (1 + \beta_R \omega_y)^{-1}. \tag{19}$$

Let us stack: $w_t = (\pi_t \;\; y_t)'$, which allows us to write the system (18) as:

$$w_t = c E_t w_{t+1} + d w_{t-1} + f w_t + \epsilon_t, \tag{20}$$

where the 2×2 matrices of the system of second-order difference equations (20) are:

$$c = \begin{pmatrix} \gamma_f & 0 \\ -\beta_R \varphi & \beta_f \varphi \end{pmatrix}, \qquad d = \begin{pmatrix} \gamma_b & 0 \\ 0 & \beta_b \varphi \end{pmatrix},$$

and f has zeros on the diagonal:

$$f = \begin{pmatrix} 0 & \lambda \\ -\beta_R \omega_\pi & 0 \end{pmatrix}.$$

The vector shock is given by: $\epsilon_t = (\epsilon_{\pi t} \ \ \varphi(\epsilon_{yt} - \beta_r \epsilon_{Rt}))'$, which implies that it is not possible to identify innovations to y_t separately from innovations to R_t. The bivariate system (20) can be written:

$$w_t = [I - c]^{-1} c E_t w_{t+1} + [I - f]^{-1} d w t - 1 + [I - f]^{-1} \epsilon_t. \tag{21}$$

This system is in exactly the same form as our original hybrid NKPC (1), except that π and x have been replaced by w and ϵ. Thus, the persistence and covariance properties of the shock vector ϵ_t will be important, just as the x_t properties were important earlier. Given that elements of f are non-zero, so that current values appear in the system, we require that the elements of ϵ_t be uncorrelated with each other. However, the rescaled shocks $[I - f]^{-1}\epsilon_t$ will be cross-correlated.

As in earlier sections, we assume uniqueness and stability, and specifically that $\omega_\pi > 1$. This restriction on monetary policy satisfies the well-known Taylor principle. Leeper (1991) calls this sort of monetary policy aggressive. When monetary policy is aggressive, only fundamental shocks, ϵ_t, drive inflation and the output gap. The unique solution again takes a first-order form:

$$w_t = a w_{t-1} + b \epsilon_t, \tag{22}$$

where a and b are 2×2 matrices. Note that the solution (22) is the equilibrium vector process of the new Keynesian economy (17). Solving for a and b by guess-and-verify methods leads to a system of polynomials in the lag operator. Factoring a multivariate spectral density matrix usually requires numerical methods; a and b cannot be found analytically in general. For discussion and examples, see Hansen and Sargent (1981) and Sayed and Kailath (2001). Nonetheless, the form of the solution (22) tells us much about the necessary conditions for identification.

Result 11. In the new Keynesian trinity model, the hybrid NKPC cannot be identified by GMM.§

The result follows from the first-order Markov nature of w_t, just as in Result 2. With y_t and π_{t-1} already entering the hybrid NKPC, there are no further variables available to instrument for π_{t+1} in GMM estimation. There will be higher-order dynamics in the *univariate* time series process for y_t implied by the NKTM. Marginalizing the VAR gives:

$$y_t = a_{22}y_{t-1} + a_{21}a_{12}\sum_{j=0}^{\infty} a_{11}^{j} y_{t-2-j}. \tag{23}$$

But there is no additional information in the lagged values of y beyond that contained in π_{t-1} because strict exogeneity does not hold in this environment. Thus, finding $J \geq 2$ is necessary, but not sufficient for identification in GMM. Although the NKTM can produce higher-order output dynamics, as in (23), these do not yield relevant instruments. Lagged inflation already enters the hybrid NKPC. Result 11 implies that identifying the NKPC must rely on cross-equation restrictions in this system.

Persistent shocks are another potential source of of identifying information. Suppose the shock vector follows a Jth-order autoregression:

$$\epsilon_t[I - \xi(L)] = \vartheta_t, \tag{24}$$

where ϑ_t is a vector of innovations. Pass $[I - \xi(L)]$ through the first-order solution (22) and substitute using the VAR of (24) to produce:

$$w_t[I - aL][I - \xi(L)] = b\vartheta_t. \tag{25}$$

The system (25) entails a VAR$(J + 1)$ in inflation and the output gap.

Result 12. One of the shocks to inflation, to the output gap, or to the interest rate must be persistent for the hybrid NKPC to be identified by GMM in the NKTM (17).§

The logic is the same as in Result 2. Identifying the second-order difference equation in inflation in GMM requires at least second-order dynamics. A necessary condition for these dynamics to arise is that the intrinsic, first-order dynamics of the NKTM (17) be augmented with first-order dynamics in at least one shock. Given there are no zero elements in $[I - f]^{-1}$, all three shocks from the original system affect π_t. Thus, persistence in at least one shock is sufficient for identification. Shock persistence also translates into serial correlation in inflation and the output gap. This helps to explain the long lags in estimated NKTM inflation and output gap equations reported, for example, by Lindé (2002) and Jondeau and Le Bihan (2003).

There is an analogous result when the NKTM (17) possesses multiple equilibria. Lubik and Schorfheide (2004) study a NKTM that associates the indeterminacy with passive monetary policy, $\omega_\pi < 1$, and sunspot (*i.e.* extrinsic) shocks. Under $\omega_\pi < 1$, they show that the rational expectations forecast of π_t and y_t is a first-order VAR with forecast innovations a function of the fundamental shocks ϵ_t and the rational expectation forecast errors, ϕ_t:

$$[I - \tau_w L]E_t w_{t+1} = \tau_\vartheta \vartheta_t + \tau_\phi \phi_t, \tag{26}$$

where $\phi_{t+1} = [y_{t+1} - E_t y_{t+1} \quad \pi_{t+1} - E_t \pi_{t+1}]'$ and the τ matrices are functions of the parameters of the NKTM (17). Given the linear NKTM (17), this class of passive monetary policies also permits ϕ_t to be a linear function of ϵ_t and a vector of sunspot shocks, ψ_t. It follows from these facts – $E_t w_{t+1}$ is the VAR(1) of (26) and ϕ_t depends on ψ_t, besides fundamental shocks – that w_t becomes a (restricted) bivariate ARMA process rather than a pure bivariate autoregression:

$$[I - \mu L]w_t = \kappa_\vartheta[I - \mu\theta_\vartheta L]\vartheta_t + \kappa_\psi[I - \mu\theta_\psi L]\psi_t, \tag{27}$$

where μ denotes the stable eigenvalue of (26) and the κ and θ matrices are functions of the NKTM parameters. Note that the first-order moving average of the bivariate ARMA process (27) are functions of the fundamental and sunspot shocks. The econometrician focuses on the sunspot to connect the observed data to one of the multiple equilibria. This motivates Lubik and Schorfheide to argue that the sunspot shock interpretation of indeterminacy (created by $\omega_\pi < 1$) explains serially correlated inflation and output gap data.

Result 13. When the new Keynesian trinity model (17) possesses multiple equilibria and the rational expectations forecast errors are a (linear) function of the fundamental and extrinsic shocks, the GMM estimator of the hybrid NKPC is not identified.§

The key to Result 13 is that the lack of restrictions on the rational expectations forecast errors under indeterminacy provides no additional identification information.

Thus, Result 13 mimics Result 8 in the univariate case. Although fundamental and sunspot shocks are news for an econometrician attempting to estimate the NKTM (17), these shocks do not help forecast π_{t+1}. However, this approach to identifying the NKPC within a larger model imposes persistence and cross-equation restrictions on the forecast innovation of the bivariate ARMA process (27) of y_t and π_t, which can yield additional information for identification.

The NKTM is a monetary model, in which the central bank's policy tool is its discount rate, R_t. Although our analysis of the NKPC with the NKTM uses the Taylor rule to substitute for the discount rate in the dynamic IS schedule, it seems reasonable to use R_t as an instrument.

Result 14. With the Taylor rule in the NKTM (17), the current nominal interest rate, R_t is not a valid instrument in the NKPC.§

The nominal interest rate is a natural predictor of π_{t+1} and so might seem to be a natural instrument. It is invalid because under the Taylor rule R_t is set as a proportion ω_π of the current inflation rate π_t which in turn is the dependent variable in the hybrid NKPC. The correlation between R_t and $\epsilon_{\pi t}$ violates the order condition.

Result 15. Lagged interest rates are valid but inefficient instruments in the NKTM.§

Recall that the solution (22) describes the optimal forecast of π_{t+1} in the NKTM based on lags of inflation and the output gap. Meanwhile, inspection of the lagged Taylor rule shows that the nominal interest rate contains information on the lagged output gap and inflation but (a) with an error ϵ_R and (b) with Taylor-rule coefficients on the lagged values of inflation and the output gap that will not correspond to the elements of the optimal coefficient matrix a given in the solution (22).

Result 16. Persistence in monetary policy may provide an alternate source of identification.§

There is much debate about whether short-term interest rates can be partly explained by lagged rates due to persistent shocks or to interest-rate smoothing. Suppose the policy rule is:

$$R_t = (1-\upsilon)(\omega_\pi \pi_t + \omega_y y_t) + \upsilon R_{t-1} + \epsilon_{Rt} \tag{28}$$

with $0 < \upsilon < 1$. The current interest rate thus reflects information on the entire history of inflation, the output gap, and policy shocks ϵ_{Rt}. The output gap inherits this memory because R_t enters the equation for the output gap in (17). Thus, additional instruments become available in the same way that Result 12 adds them using shock persistence.

This section has focused on the bivariate VAR in $\{w_t\}$ because of our interest in instrumenting π_{t+1} in the hybrid NKPC. Thus, we have not studied the complete reduced form, or addressed the identification of other parameters in the NKTM. The main result of this section is that a persistent shock or an interest-rate-smoothing policy is necessary for the hybrid NKPC to be identified by GMM within this richer system.

5. Revisiting the Evidence

We next apply our results to the estimation of hybrid NKPCs for the U.S., U.K., and Canada. The data consists of GDP inflation and measures of real marginal cost. The appendix describes the data sources.

5.1 Statistics

First, we study the time-series properties of x_t. We estimate univariate autoregressions for x_t, and test the lag length from $J = 1$ to $J = 6$ lags using a likelihood ratio statistic, the AIC, and the SIC. Recall from Result 2 that – if there are no instruments other than lags of x – then $J \geq 2$ is necessary for identification in GMM.

We next include lagged values of inflation and report the results of a pre-test of the null hypothesis that $\{\pi_t\}$ does not Granger-cause $\{x_t\}$. Finding a role for lagged inflation suggests that further instruments may be available. These could include lags of inflation beyond the first two or other variables that lead to Granger-causality because of the superior information of price-setters. Were we to proceed with system estimation, this test also would tell us if we need to algebraically unscramble the system to separately distinguish a role for lagged inflation arising from forecasting from one arising from price-stickiness. Recall from Result 10 that lagged inflation in the solved model reflects both of these factors in the absence of strict exogeneity.

Second, our main interest is in instrumental-variables estimation, so we estimate:

$$E[\pi_t - \gamma_f \pi_{t+1} - \gamma_b \pi_{t-1} - \lambda x_t | z_t] = 0 \tag{29}$$

by GMM and report point estimates and standard errors as well as the J-test statistic of over-identifying restrictions and its p-value. Following Result 8, GMM estimators will allow for a first-order moving average in the GMM residual. The weighting matrix will be the continuous-updating version introduced by Hansen, Heaton, and Yaron (1996), which has good finite-sample properties and is invariant to the normalization of the hybrid NKPC (1).

Third, we estimate an example of a first-stage, linear projection:

$$\pi_{t+1} = \beta_0 + \beta_1 \pi_{t-1} + \beta_2 x_t + \beta_3 u_t, \tag{30}$$

which naturally excludes π_t, and where u_t is a $k \times 1$ vector of instruments that excludes π_{t-1} and x_t. A necessary and sufficient condition for the identification of the forward-looking part of the hybrid NKPC, γ_f, is that (at least) some of the elements of the $1 \times k$ vector β_3 are not zero so that the rank condition holds. If $\beta_3 = 0$, the components of u_t cannot be separated from the other two explanatory variables in the hybrid NKPC, which are included as controls. In the case of the purely forward-looking NKPC (*i.e.* $\gamma_b = 0$), lagged inflation becomes a valid instrument for π_{t+1}. In that case, the projection (30) finds valid instruments as long as either β_3 or β_1 is non-zero.

The statistics from this projection (30) are calculated to tie the evidence on identification of the hybrid NKPC to work on forecasting inflation. One can see that identification requires the ability to forecast inflation two steps ahead, without using the intervening output gap or real marginal cost. In the hybrid model, the investigator needs to find an eligible instrument that provides predictive information for π_{t+1} *beyond* that contained in x_t and π_{t-1}. This is a stringent requirement. Stock and Watson (1999) and Hansen, Lunde, and Nason (2004) report that few variables have power to forecast inflation during the great disinflation of the 1980s and 1990s.

Our main interest is in GMM estimation. Although the analysis of sections 2-4 sets the hybrid NKPC within various statistical and economic environments, we do not propose a 'best' inflation forecasting model. In practice, good forecasting procedures are unlikely to resemble the constant-coefficient, linear rules in our theoretical examples. Clements and Hendry (2003) provide a full review. Stock and Watson (1999) and Hansen, Lunde, and Nason (2004) report the 'best' inflation forecasting equations for the U.S. differ across subsamples.

Fourth, we calculate Anderson-Rubin (1949) statistics to test several hypotheses, and find the implied confidence intervals. The statistics from GMM estimation (29) and from our examples of first-stage projections (30) depend on nuisance parameters under weak identification. In contrast, the AR statistics are pivotal in finite samples. To test $H_0: \ \gamma_f = \gamma_{f0}$ one projects as follows:

$$\pi_t - \gamma_{f0}\pi_{t+1} = \alpha_0 + \alpha_1\pi_{t-1} + \alpha_2 x_t + \alpha_3 u_t, \tag{31}$$

then constructs the Anderson-Rubin (AR) F-statistic for $H_0': \ \alpha_3 = 0$. The idea is that there should be no further role for u_t at the true value for γ_f. In our case, γ_f is a scalar. This yields a $F(k+2, T-k)$ statistic, where $k+2$ is the total number of exogenous variables and instruments. The Anderson-Rubin (AR) statistic provides an exact test, which is robust to (a) weak instruments and (b) omitted instruments. We do not need all the u-elements necessarily, but power is lower if irrelevant instruments are included. The test statistic also is robust to misspecification of the forecasting rule for π_{t+1} (*i.e.* its size is not affected, though again its power may be).

The distributional assumption underlying the statistic's being pivotal in finite samples is normality of the GMM residuals. In the literature, the main drawbacks to this approach arise when the structural equation is non-linear, or when there is more than one endogenous, explanatory variable and we want to study subsets of their coefficients. But here the hybrid NKPC is *linear*, and γ_f is a scalar. Alternative test statistics have been developed by Wang and Zivot (1998) and Kleibergen (2002). These may improve test power, but they do so by using some information from a first-stage regression (*i.e.* a reduced-form for π_{t+1}, which we wish to avoid here). Also, these test statistics are not robust to instrument exclusion or to the form of the forecasting rule for π_{t+1}. Dufour (2003, section 6) provides an excellent discussion.

The AR statistics also can be used to construct confidence intervals. A confidence set is:

$$C(\alpha) = \{\gamma_{f0} : AR(\gamma_{f0}) \leq F_\alpha(k, T-k-2)\}. \tag{32}$$

Since γ_f is a scalar, there is a quadratic solution, given by Zivot, Startz, and Nelson (1998). The coefficients of the quadratic equation are functions of the data and the $F-$statistic at significance level α and degrees of freedom $\dim(u)$ and $T-2-k$. With over-identification this confidence set can be empty. Without identification, it can be unbounded. The approach can be extended to test restrictions on the exogenous variables, such as γ_b for example. A test of $H_0: \ \gamma_f = \gamma_{f0}, \gamma_b = \gamma_{b0}$ begins with:

$$\pi_t - \gamma_{f0}\pi_{t+1} - \gamma_{b0}\pi_{t-1} = \alpha_0 + \alpha_1\pi_{t-1} + \alpha_2 x_t + \alpha_3 u_t, \tag{33}$$

and leads to an F-test of whether α_1 and α_3 are jointly zero. One also may construct a joint confidence set for γ_b and γ_f.

5.2 United States

The first two rows of table 2 present evidence on the dynamics of real marginal cost for a U.S. sample of $1949Q1 - 2001Q4$. They show that we cannot reject the null hypothesis that inflation does not Granger-cause real marginal cost. Thus, Result 10 indicates for the U.S. it is straightforward to separate lagged inflation's job as a predictor of future marginal cost from its role as a measure of backward-looking price-setting.

In addition, the AIC and LR statistics suggested a lag length of 3, while the SIC suggested a lag length of 1. The coefficient $\hat{\rho}_2$ in the x-autoregression was insignificantly different from zero. The implication of these pre-tests is that finding relevant instruments may be challenging in the U.S. data. Although U.S. real marginal cost is persistent (the half life of a shock to its AR(3) processs is about seven quarters), there is not strong evidence of higher-order dynamics in U.S. real marginal cost. Campbell and Shiller (1987) and Boileau and Normandin (2002) showed that the presence of other predictors of x_t also should lead to a role for lagged inflation, yet we find none here, so the quest for other instruments may not be fruitful.

Table 3 contains single-equation GMM estimates. Most of the work is done by the instruments $\{\pi_{t-1}, x_t, x_{t-2}\}$, as is suggested by the pre-test evidence that only x_t and x_{t-2} help forecast x_{t+1}. Adding further instruments increases the precision slightly but does not lead to significant changes in the estimates. The J-test clearly does not reject the over-identifying restrictions.

The estimated weight attributed to backward-looking inflationary expectations, $\hat{\gamma}_b$, ranges from 0.28 to 0.42, depending on the instrument set. The GMM estimates show these expectations are dominated by forward-looking expectations because $\hat{\gamma}_f$ ranges from 0.52 to 0.70. The response of π_t to x_t, denoted $\hat{\lambda}$, also takes plausible values, between 0.1 and 0.9 percent, but is not statistically significant (for a five percent test). Our results are comparable to those of Galí and Gertler (1999, table 2), but we obtain smaller and insignificant estimates of λ using smaller instrument sets.

Table 4 presents AR F-statistics and their associated p-values based on equation (31) and a grid of potentially 'true' $\gamma_f = \gamma_{f0}$. We set γ_{f0} to [0.0, 0.2, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99]. The AR statistics in the first row reveal little evidence against the null of $\gamma_f = \gamma_{f0}$, for any of these values of γ_{f0} given $u_t = x_{t-2}$. When we add instruments though – in the next two rows – we can reject any of the null hypotheses at standard significance levels. Thus, lags of real marginal cost besides x_{t-2} matter for predicting the quasi-difference of π_t and π_{t+1}. The test is correctly sized even if these added instruments are weak, which gives us a formal rejection of the forward looking model.

The asymptotic 95 percent confidence interval $C(\alpha = 0.05)$ of γ_f, given in (32), lends more support to the evidence of table 4. The solution yields $C(0.05) = \{-6.75,$

0.05}, {0.60, 0.84}, and {−0.10, 1.24} for $u_t = x_{t-2}$, $\{x_{t-1}, x_{t-2}\}$, and $\{x_{t-1}, \ldots, x_{t-4}\}$, respectively. The smallest (just-identified) and largest (overidentified) instrument sets yield asymptotic 95 percent confidence intervals that contain zero. Only the information vector with the first two lags of x produces a confidence interval with reasonable values of γ_f. The sensitivity of $C(0.05)$ to the content of u_t suggests that estimates of the weight γ_f on forward-looking inflationary expectations are only weakly identified within the hybrid NKPC on U.S. data.

We also report AR statistics and their p-values for the joint null $\gamma_f = \gamma_{f0}$ and $\gamma_b = \gamma_{b0}$ of projection (33) in Table 5. The grid of potentially true values of γ_f and γ_b is tied to the estimates found in table 3. Tests of hypotheses that place zero weight on either γ_f, γ_b, or both are examined as well.

The inference we draw from table 5 is similar to that presented in table 4. There are few rejections of the joint null, conditional on x_{t-2} being the only element of the instrument vector u_t, except when γ_{f0} equals zero. But the introduction of other relevant lags of x to u_t leads to rejection of the null across all the γ_{f0} and γ_{b0} combinations table 5 considers. These rejections occur at the eight percent level or less. Thus, we find that evidence in favor of the null relies on γ_{f0} and γ_{b0} being within the range $\hat{\gamma}_f$ and $\hat{\gamma}_b$ take, conditional on the most concise instrument vector of table 5. Otherwise, rejections of the joint null are robust to the instrument vector and values of γ_{f0} and γ_{b0}. An implication is that the joint significance of the forward- and backward-looking weights on inflation in the hybrid NKPC is suspect, independent of satisfying the rank conditions laid out in Results 3 and 4.

The results of table 4 are consistent with the test of the hypothesis $\beta_3 = 0$ in the first-stage projection (30) (not shown). The least squares t−ratio of β_3 is -1.93 when $u_t = x_{t-2}$, which rejects the hypothesis at the 2.7 percent level. Thus, this single instrument provides additional explanatory power to π_{t+1} in (30), which implies the rank condition is satisfied. Remember that this is also the only instrument vector for which the null hypothesis of projection (31) is not rejected. The rank condition fails to hold when we add x_{t-1} to u_t. The Wald statistic of the bivariate hypothesis $\beta_3 = 0$

is 4.04, with a p-value of 0.13. The hypothesis also is not rejected at reasonable significance levels when u_t is expanded to $\{x_{t-1}, \ldots, x_{t-4}\}$, which yields a Wald statistic of 4.21 with a p-value of 0.38.

5.3 United Kingdom

The estimation sample for the U.K. is $1961Q1 - 2000Q4$. Table 2 shows that the Granger causality pre-test provides strong evidence of predictability in both directions. From Result 10, this implies that single-equation ordinary least squares cannot measure the inertia in price-setting, γ_b. It also implies that lagged values of inflation (beyond the first two lags) may be available as instruments. The second set of pre-tests indicate a lag length of $J = 5$ using the LR test and SIC. This places more of the history of x in the instrument vector z_t (or u_t). We also find $\hat{\rho}_3$ is insignificantly different from zero, but a leading eigenvalue of 0.91 (for $J = 4$) reveals U.K. real marginal cost to be a persistent process.

Table 6 contains estimates of the U.K. hybrid NKPC. The GMM estimates depend on instrument choice. Once lags up to x_{t-4} are included, the coefficients accord with theory and are estimated with some precision. However, the over-identifying restrictions are rejected, given x_t is an instrument. When x_t is not an instrument, the estimates of γ_f, γ_b, and λ are significant at the ten percent level or better. Neiss and Nelson (2002) obtain statistically significant estimates of λ, but use dummy variables to control for a variety of price shocks. Like us, Balakrishnan and López-Salido (2002) do not find a significant, stable effect of real marginal cost on UK inflation.

Tables 7 and 8 give evidence against the null of $\gamma_f = \gamma_{f0}$ or the joint null of $\gamma_f = \gamma_{f0}$ and $\gamma_b = \gamma_{b0}$ for the U.K. hybrid NKPC. The significance levels of the AR statistics average 0.03 in table 7, for the projection (31), on the same grid of values of γ_{f0} used for table 4. Only two of the 16 AR statistics have p-values that exceed ten percent, which are associated with the information vector $u_t = x_{t-1}$, as γ_{f0} approaches unity.

The AR 95 percent asymptotic confidence interval (32) of γ_f is $C(0.05) = \{-0.87, -0.01\}$ when $u_t = x_{t-1}$, and $C(0.05) = \{0.02, 0.97\}$ when $u_t = \{x_{t-1}, ..., x_{t-4}\}$. Thus the confidence interval of γ_f has the wrong sign with the smaller information set. The confidence interval takes the correct sign using the larger information set and matches values set for γ_{f0} in table 7. Nonetheless, with equal probability γ_f runs from economically meaningless values to values that reveal an important role for forward-looking inflationary expectations.

The information vector $u_t = x_{t-1}$ also is responsible for the only AR statistic with a p-value greater than ten percent in table 8. However, the combination of $\gamma_{f0} = -0.15$ and $\gamma_{b0} = 0.00$ that produces this AR statistic does not resemble estimates reported in table 6. The remaining $(\gamma_{f0}, \gamma_{b0})$ pairs are linked to AR statistics that indicate a rejection of the joint null. It is striking that the rejections appear strongest for null hypotheses closest to the point estimates $\hat{\gamma}_f$ and $\hat{\gamma}_b$.

The rejections of the null in projections (31) and (33) hold for either $u_t = x_{t-1}$ or $u_t = \{x_{t-1}, \ldots, x_{t-4}\}$. However, the hypothesis $\beta_3 = 0$ in projection (30) fails to be rejected for the former information set, but not the latter. Once-lagged x has no predictive content for π_{t+1} because $\hat{t}_{\beta_3} = 1.07$. Since $u_t = x_{t-1}$ violates the (necessary and sufficient) hybrid NKPC rank condition, it is not a valid instrument. When we add the next three lags of x to u_t, the Wald statistic of the joint null of $\beta_3 = 0$ is 15.30, with a p-value of 0.00. Thus the instrument vector $\{x_{t-1}, \ldots, x_{t-4}\}$ satisfies the rank condition; it can forecast π_{t+1}. As we have seen, the problem for the NKPC is that this instrument vector also can forecast the quasi-difference $\pi_t - \gamma_{f0}\pi_{t+1}$ for a wide range of values of γ_{f0}, implying rejections of the NKPC.

5.4 Canada

The estimation and testing for Canada use data from 1963Q1 to 2000Q4. Table 2 shows that Canadian inflation Granger-causes real marginal cost. Thus, x_t is not strictly exogenous. This table also shows real marginal cost fails to Granger-cause inflation – in contrast to results for the U.K. and U.S. data. The pre-tests for lag length reveal a persistence pattern similar to that in U.S. real marginal cost, according to the LR test, the AIC, and the SIC. In the time series for $\{x_t\}$, once-lagged costs play a large predictive role and thrice-lagged costs play an additional role, that is statistically significant. However, a half-life of 8.5 quarters with respect to a shock to its AR(3) process shows that Canadian real marginal cost is more persistent than it is in the U.K. and the U.S. data.

Table 9 contains estimates of the hybrid NKPC parameters γ_f, γ_b, and λ for Canada. They suggest that the hybrid NKPC is poorly identified. For example, the point estimates $\hat{\gamma}_f$ and $\hat{\gamma}_b$ are sensitive to the instrument set. When we include π_{t-2} as an instrument, these two coefficients are similar to those found in the U.S. data, with a large role for future inflation.

Guay, Luger, and Zhu (2003) estimate the hybrid NKPC using a wider range of instruments. They use much larger instrument sets and increase precision (and reject the over-identifying restrictions). However, we reproduce their finding that $\hat{\lambda}$ is insignificant. This indicates little role for real marginal cost in Canadian inflation dynamics.

Tables 10 and 11 yield inferences that are the opposite of those for the U.S. and the U.K. data. None of the hypothesized values of (γ_f, γ_b) can be rejected at the five percent level. These test results leave us with considerable uncertainty about the 'true' value of γ_f.

The AR 95 percent asymptotic confidence interval (32) for γ_f, $C(0.05)$ supports this conjecture for Canada. For the instrument vectors $u_t = x_{t-2}$, $\{x_{t-1}, x_{t-2}\}$, and $\{x_{t-1}, \ldots, x_{t-4}\}$, $C(0.05) = \{-0.00, 0.97\}$ $\{-0.00, 0.78\}$ and $\{-0.00, 0.97\}$, respectively. Since the three AR asymptotic 95 percent confidence intervals cover zero, there is more evidence that forward-looking inflationary expectations may not matter for Canadian inflation dynamics.

We also find that tests of the predictive power of u_t for π_{t+1} in projection (30) fail to reject the null that $\beta_3 = 0$ for the instrument vectors x_{t-2}, $\{x_{t-1}, x_{t-2}\}$, and $\{x_{t-1}, \ldots, x_{t-4}\}$, at a 15 percent significance level or better. Thus we have not found valid instruments. Overall, this combination of statistics shows that it is not possible to identify the weights on the forward- and backward-looking components of the hybrid NKPC in this bivariate data set.

6. Conclusion

This paper is about identification problems in the hybrid new-Keynesian Phillips curve (NKPC), within a linear rational expectations setting. Table 1 collects our analytical results. We show that estimation of the hybrid NKPC faces a fundamental source of non-identification: weak, higher-order dynamics. System estimation has an identification advantage over GMM because of an additional restriction. In this case, the hybrid NKPC can be identified even if real aggregate demand follows a first-order Markov process. However, system estimation implies a structural VAR whose interpretation may be unpalatable to advocates of new Keynesian macro models.

By setting the hybrid NKPC in a new Keynesian trinity model, we find this Phillips curve cannot be identified by GMM. In this setting, the current nominal interest rate also is ineligible as an instrument, as long as a Taylor rule applies. One solution to the identification problem is to posit persistent shocks either to real aggregate demand, inflation, or monetary policy, as is often implicitly done in the literature.

It is difficult to find evidence of significant coefficients $\{\hat{\gamma}_f, \hat{\gamma}_b, \hat{\lambda}\}$ in the hybrid NKPC across the U.S., U.K., and Canada. One reason for the poor quality of the estimates is that for all three countries, real marginal cost has some higher-order dynamics, but perhaps not enough to avoid the problem of weak instruments. We draw on the Anderson-Rubin statistic to provide a new set of tests of the forward-looking inflation model. These test statistics are exact, pivotal, and robust to either weak or omitted instruments. The tests reveal little evidence of forward-looking expectations driving U.S., U.K., or Canadian inflation.

Our results do not imply that inflation lacks serial correlation. Clearly, it is possible that the hybrid NKPC is a useful tool, but that a broader set of instruments is needed to forecast real marginal cost. Kurmann (2003b) explores this issue. Another possibility, though, is that the second-order difference equation (1) simply is not a reasonable model of inflation dynamics.

Stock and Wright (2000) and Stock, Wright, and Yogo (2002) provide further tools for GMM estimation and inference with weak instruments. Ma (2002) shows using the

S-sets developed by Stock and Wright (2000) that γ_f is weakly identified in the Galí-Gertler data. The interaction of the identification and estimation problems that face the hybrid NKPC also can be studied by Monte Carlo methods. Lindé (2002), Jondeau and Le Bihan (2003), and Mavroeidis (2004b) report that the hybrid-NKPC is sensitive to the economic environment in which it resides because of the impact on instrument choice and quality.

We view the combination of our analytic and empirical work as a complement to all of these studies. The lack of higher-order dynamics in U.S., U.K., and Canadian real marginal cost points to difficulties in identifying the hybrid NKPC coefficients, as noted in Result 2. The generally negative results with the AR statistic and predictability of π_{t+1} indicate that this problem can be particularly acute for the weight on forward-looking inflationary expectations. Alternative sources of identifying information – say regime change or survey data – are worth future study, because the underlying primitives of the NKPC certainly matter for monetary policy. This paper suggests more work needs to be done for the NKPC to remain a viable story of inflation dynamics.

Appendix: Data Sources

United States

The price level P_t is the GDP implicit price deflator. The GDP deflator is available in chain weight form and in implicit form (all the U.S. results are based on the implicit GDP deflator).

Nominal unit labor cost (ULC) is the ratio of the index of hourly compensation in the non-farm business sector, labelled COMPNFB, to output per hour of all persons in the non-farm business sector, labelled OPHNFB. COMPNFB is an index of the nominal wage. OPHNFB is an index of the average product of labor. These can be found in the Federal Reserve Bank of St. Louis' FRED databank. Thus, ULC is a measure of labor's share.

Real ULC equals nominal ULC deflated by P_t. Inflation is $100 \ln(P_t/P_{t-1})$ and real ULC is $100(1+a)\ln(\text{COMPNFB}_t/\text{OPHNFB}_t) - 100 \ln P_t$, where a is a function of the steady-state markup and labor's share parameter in the firm's production function. This adjustment renders real ULC stationary and $a = 1.08$.

The estimation sample period is 1947Q1-2002Q4, $T = 224$.

United Kingdom

The inflation rate is measured with the GDP deflator, and x is a measure of the log of real marginal cost. Data sources are given by Katharine Neiss and Edward Nelson (2002), who kindly provided the data. The estimation period is 1961Q1 to 2000Q4, so $T = 168$.

Canada

The inflation rate is measured with the GDP deflator, while x is the log of the labour share in the non-farm, business sector. Data sources are given by Guay, Luger, and Zhu (2003), who kindly provided the data. The estimation period is 1963Q1 to 2000Q4.

References

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

Balakrishnan, Ravi and López-Salido, J. David (2002). Understanding U.K. inflation: The role of openness. Working paper no. 64, Research Department, Bank of England.

Bardsen, Gunnar, Jansen, Eilev and Nymoen, Ragnar (2002). The empirical (ir)relevance of the new Keynesian Phillips curve. Working paper no. 21, Department of Economics, Norwegian University of Science and Technology.

Beyer, Andreas and Farmer, Roger (2004). On the indetermincacy of New-Keynesian economics. mimeo, European Central bank.

Boileau, Martin and Normandin, Michel (2002). Aggregate employment volatility, real business cycles, and superior information. *Journal of Monetary Economics* 49, 495-520.

Campbell, John Y. and Shiller, Robert (1987). Cointegration and tests of present-value models. *Journal of Political Economy* 95, 357-374.

Clements, Michael and Hendry, David (2003). Economic forecasting: Some lessons from recent research. *Economic Modelling* 20, 301-329.

Dufour, Jean-Marie (2003). Identification, weak instruments, and statistical inference in econometrics. *Canadian Journal of Economics* 36, 767-808.

Engle, Robert F. and Granger, Clive W.J. (1987). Cointegration and Error Correction: Representation, Estimation, and Testing, *Econometrica* 55, 251-130.

Engle, Robert F., Hendry, David F., and Richard, Jean-François (1983) Exogeneity. *Econometrica* 51, 277-304.

Fuhrer, Jeffrey C. and Moore, George R. (1995). Inflation persistence. *Quarterly Journal of Economics* 110, 127-159.

Fuhrer, Jeffrey C., Moore, George R., and Schuh, Scott D. (1995). Estimating the linear-quadratic inventory model: Maximum likelihood versus generalized method of moments. *Journal of Monetary Economics* 35, 115-157.

Fuhrer, Jeffrey C., and Olivei, Giovanni P. (2004). Estimating Forward-Looking Euler Equations with GMM and Maximum Likelihood Estimators: An Optimal Instruments Approach. Working Paper no. 04-2, Federal Reserve Bank of Boston.

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

Galí, Jordi, Gertler, Mark, and López-Salido, J. David (2001). Notes on estimating the closed form of the hybrid new Phillips curve. mimeo, Bank of Spain.

Gregory, Allan, Pagan, Adrian, and Smith, Gregor (1993). Estimating linear-quadratic models with integrated processes. pp. 220-239 in P.C.B. Phillips, ed. *Models, Methods, and Applications of Econometrics.* Oxford: Basil Blackwell.

Guay, Alain, Luger, Richard, and Zhu, Zhenhua (2003). The new Phillips curve in Canada. in *Price Adjustment and Monetary Policy.* Ottawa: Bank of Canada.

Hansen, Lars Peter, Heaton, John, and Yaron, Amir (1996). Finite-sample properties of some alternative GMM estimators. *Journal of Business and Economic Statistics* 14, 262-280.

Hansen, Lars Peter, and Sargent, Thomas J. (1980). Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control* 2, 7-46.

Hansen, Lars Peter and Sargent, Thomas J. (1981). Linear rational expectations models for dynamically interrelated variables, pp. 127-156 in *Rational Expectations and Econometric Practice*, eds. R.E. Lucas Jr. and T.J. Sargent. University of Minnesota Press.

Hansen, Peter R., Lunde, Asger, and Nason, James M. (2004). Model confidence sets for forecasting models. mimeo, Research Department, Federal Reserve Bank of Atlanta.

Ireland, Peter N. (2002). Technology Shocks in the New Keynesian Model. Working paper 536, Department of Economics, Boston College.

Jondeau, Eric and Le Bihan, Hervé (2003). ML vs GMM estimates of hybrid macroeconomic models (with and application to the 'new Phillips curve'). NER 103, Banque de France.

Kennan, John (1979). The estimation of partial adjustment models with rational expectations. *Econometrica* 47, 1441-1455.

King, Robert and Watson, Mark (1994). The post-war U.S. Phillips curve: A revisionist econometric history. *Carnegie-Rochester Conference Series on Public Policy* 41, 157-219.

King, Robert and Watson, Mark (1997). Testing long-run neutrality. Federal Reserve Bank of Richmond *Economic Quarterly* 83/3, 69-101.

Kleibergen, Frank (2002). Pivotal statistics for testing structural parameters in instrumental variables regression. *Econometrica* 70, 1781-1803.

Kurmann, André (2003a). Maximum likelihood estimation of dynamic stochastic theories with and application to new Keynesian pricing. mimeo, Université du Québec à Montréal.

Kurmann, André (2003b). Quantifying the uncertainty about a forward-looking, new Keynesian pricing model. mimeo, Université du Québec à Montréal.

Leeper, Eric M. (1991). Equilibria under 'active' and 'passive' monetary and fiscal policies. *Journal of Monetary Economics* 27, 129-147.

Lindé, Jesper (2002). Estimating new-Keynesian Phillips curves: A full information maximum likelihood approach. Working Paper No. 129, Sveriges Riksbank (http://www.riksbank.com/upload/6286/wp_129.pdf)

Lubik, Thomas and Schorfheide, Frank (2004). Testing for indeterminacy: An application to U.S. monetary policy. *American Economic Review* 94, 190-217.

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

Mavroeidis, Sophocles (2004a). Weak identification of forward-looking models in monetary economics. *Oxford Bulletin of Economics and Statistics*, forthcoming.

Mavroeidis, Sophocles (2004b). Identification issues in forward-looking models estimated by GMM, with an Application to the Phillips curve. mimeo, Department of Quantitative Economics, University of Amsterdam.

Nason, James M., and George A. Slotsve (2004). Along the New Keynesian Phillips Curve with Nominal and Real Rigidities. Working Paper 2004-9, Federal Reserve Bank of Atlanta.

Neiss, Katharine and Nelson, Edward (2002). Inflation dynamics, marginal costs, and the output gap: Evidence from three countries. mimeo, Bank of England.

Pesaran, M. Hashem (1987). *The limits to rational expectations* Oxford: Basil Blackwell.

Roberts, John M. (1995). New Keynesian economics and the Phillips curve. *Journal of Money, Credit, and Banking* 27, 975-984.

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

Rotemberg, Julio (1982). Monopolistic price adjustment and aggregate output. *Review of Economic Studies* 49, 517-531.

Rudd, Jeremy and Whelan, Karl (2001). A reconsideration of the new Keynesian Phillips curve, mimeo, Board of Governors of the Federal Reserve System

Sargent, Thomas J. (1987). *Macroeconomic Theory* (second edition). Academic Press, NY, New York.

Sargent, Thomas J. (1999). *Conquest of American Inflation*. Princeton University Press, Princeton, NJ.

Sayed, A.H. and Kailath, T. (2001). A survey of spectral factorization methods. *Numerical Linear Algebra with Applications* 8, 467-496.

Sbordone, Argia M. (2002). Prices and unit costs: A new test of price stickiness. *Journal of Monetary Economics* 49, 235-256.

Stock, James and Watson, Mark (1999). Forecasting inflation. *Journal of Monetary Economics* 44, 293-335.

Stock, James and Wright, Jonathan (2000). GMM with weak identification. *Econometrica* 68, 1055-1096.

Stock, James, Wright, Jonathan, and Yogo, Motohiro (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics* 20, 518-529.

Wang, Jiahui and Zivot, Eric (1998). Inference on structural parameters in instrumental variables regression with weak instruments. *Econometrica* 66, 1389-1404.

West, Kenneth D. (1988). Asymptotic normality, when regressors have a unit root. *Econometrica* 56, 1397-1417.

West, Kenneth D. and Wilcox, David W. (1994). Estimation and inference in the linear-quadratic inventory model. *Journal of Economic Dynamics and Control* 18, 897-908.

Woodford, Michael (2003). *Interest and prices.* Princeton University Press, Princeton, NJ.

Yun, Tack (1996). Nominal rigidities, money supply endogeneity, and business cycles, *Journal of Monetary Economics* 37, 345-370.

Zivot, Eric, Richard Startz, and Nelson, Charles (1998). Valid confidence intervals and inference in the presence of weak instruments. *International Economic Review* 39, 1119-1144.

Table 1
Summary of
New Keynesian Phillips Curve Identification Results

Result 1. The hybrid NKPC imposes the King and Watson (1994) real business cycle identification on the structural VAR (10) for $\{\tilde{x}_t, \pi_t\}$ and the Solow-Gordon identifying assumption on the impact matrix of the unrestricted simultaneous equations system of $\{\tilde{x}_t, \pi_t\}$.

Result 2. The number of regressors in (7) is $J + 1$. The parameters in $\tilde{\rho}$ can be identified from estimation of the law of motion for x_t, (4). With three parameters $\{\gamma_f, \gamma_b, \lambda\}$ to identify, $J \geq 1$ is necessary for identification in the solved model (7). $J \geq 2$ is necessary for overidentification.

Result 3. If $z_t = \{\pi_{t-1}, x_t, x_{t-1}, x_{t-2}, \ldots, x_{t-J+1}\}$, then $J \geq 2$ is necessary for identification by GMM and $J \geq 3$ is necessary for overidentification.

Result 4. If $z_{t-1} = \{\pi_{t-1}, x_{t-1}, x_{t-2}, \ldots, x_{t-J}\}$, so that only lagged information is used, then again $J \geq 2$ is necessary for identification by GMM and $J \geq 3$ is necessary for overidentification.

Result 5. If a consistent estimate $\hat{\lambda}$ is available, then if $J \geq 1$ is necessary for the identification of γ_f and γ_b in the solved-system environment. In the single-equation environment with instruments z_t, $J \geq 1$ is necessary for identification and $J \geq 2$ for overidentification. With instruments z_{t-1}, however, $J \geq 2$ remains necessary for identification and $J \geq 3$ for overidentification.

Result 6. The conditions for identification do not change if the investigator imposes $\gamma_b = 0$, so that the NKPC is purely forward-looking.

Result 7. Solving forward and truncating provides no additional information to aid identification (or improve efficiency).

Result 8. Whether z_t or z_{t-1} is adopted, the GMM residual is a MA(1) process. Both of these instrument sets are valid, but any instrument set must exclude lagged GMM residuals. In addition, the loss of precision from excluding x_t from the instrument set depends on parameters in its law of motion *and* on the hybrid NKPC parameters.

Result 9. Predicting x with once-lagged or twice-lagged inflation adds no identifying information. $J \geq 3$ is necessary for the VAR to add overidentifying information. Thus Results 2 and 3 continue to apply within the VAR.

Result 10. The coefficient on π_{t-1} in the solved hybrid NKPC is independent of the process followed by marginal cost *iff* inflation does not Granger-cause marginal cost.

Result 11. In the NKTM, the hybrid NKPC cannot be identified by GMM.

Result 12. Either the shock to inflation, the output gap, or the interest rate must be persistent for the NKPC to be identified by GMM in the NKTM.

Result 13. When the NKTM possesses multiple equilibria and the rational expectations forecast errors are a (linear) function of the fundamental and extrinsic shocks, the GMM estimator of the hybird-NKPC is not identified.

Result 14. With the Taylor rule in the NKTM, the current nominal interest rate, R_t, is not a valid instrument in the NKPC.

Result 15. Lagged interest rates are valid but inefficient instruments.

Result 16. Persistence in monetary policy may provide an alternate source of identification.

Table 2
Granger Non-Causality Tests

Country	Lag length (d.f.)	$p \quad \pi \not\to x$	$p \quad x \not\to \pi$
U.S.	3	0.18	0.05
U.S.	4	0.24	0.08
U.K.	4	0.01	0.00
U.K.	5	0.01	0.00
Canada	3	0.00	0.73
Canada	4	0.00	0.63

Notes: The lag lengths, $\hat{J}$, are the same as those selected by information criteria. Entries are p-values for the null hypothesis that the first variable does not Granger cause the second variable. Data sources and sample sizes are given in the data appendix.

Table 3
U.S. New Keynesian Phillips Curve

$$E[\pi_t - \gamma_f E_t \pi_{t+1} - \gamma_b \pi_{t-1} - \lambda x_t | z_t] = 0$$

1949Q1 - 2001Q4 $T = 212$

Instruments	$\hat{\gamma}_f$ (se)	$\hat{\gamma}_b$ (se)	$\hat{\lambda}$ (se)	$\chi^2(df)$ (p)
$\{\pi_{t-1}, x_t, x_{t-2}\}$	0.685 (0.357)	0.300 (0.247)	0.001 (0.007)	—
$\{\pi_{t-1}, x_t, ..., x_{t-2}\}$	0.527 (0.298)	0.415 (0.205)	0.009 (0.005)	2.11(1) (0.35)
$\{\pi_{t-1}, x_t, ..., x_{t-4}\}$	0.706 (0.223)	0.275 (0.158)	0.008 (0.006)	3.47(3) (0.48)
$\{\pi_{t-1}, \pi_{t-2}, x_t, ..., x_{t-4}\}$	0.701 (0.188)	0.278 (0.141)	0.009 (0.005)	3.48(4) (0.63)

Notes: The entire sample runs from 1949Q1 to 2002Q1. Estimation is based on a 1947Q1−2001Q4 sample. Tests of the over-identifying restrictions use the J-statistic.

Table 4
U.S. NKPC: Tests of $H_0: \gamma_f = \gamma_{f0}$

$$\pi_t - \gamma_{f0}\pi_{t+1} = \alpha_0 + \alpha_1\pi_{t-1} + \alpha_2 x_t + \alpha_3 u_t$$

Anderson-Rubin Statistic

1949Q1 – 2001Q4 $\qquad T = 212$

$\gamma_{f0} =$	0.00 (p)	0.20 (p)	0.50 (p)	0.60 (p)	0.70 (p)	0.80 (p)	0.90 (p)	0.99 (p)
$u_t =$								
$\{x_{t-2}\}$	2.15 (0.14)	1.31 (0.25)	0.21 (0.65)	0.04 (0.83)	0.00 (0.97)	0.07 (0.80)	0.21 (0.64)	0.39 (0.53)
$\{x_{t-1}, x_{t-2}\}$	4.43 (0.01)	5.17 (0.01)	5.85 (0.00)	5.83 (0.00)	5.68 (0.00)	5.45 (0.00)	5.17 (0.01)	4.90 (0.01)
$\{x_{t-1}, ..., x_{t-4}\}$	2.47 (0.05)	2.92 (0.02)	3.34 (0.01)	3.33 (0.01)	3.24 (0.01)	3.10 (0.02)	2.93 (0.02)	2.77 (0.03)

Notes: The Anderson-Rubin statistics in the top panel are based on equation (31) of the paper, under the null that γ_{f0}=0. Dufour (2003) contains details of the Anderson-Rubin statistic and test. Otherwise, see the notes to table 3.

Table 5
U.S. NKPC: Tests of $H_0: \ \gamma_f = \gamma_{f0},\ \gamma_b = \gamma_{b0}$

$$\pi_t - \gamma_{f0}\pi_{t+1} - \gamma_{b0}\pi_{t-1} = \alpha_0 + \alpha_1\pi_{t-1} + \alpha_2 x_t + \alpha_3 u_t$$

Anderson-Rubin Statistic

1949Q1 – 2001Q4 $\qquad T = 212$

γ_{f0} =	0.00 (p)	0.25 (p)	0.50 (p)	0.68 (p)
$u_t = \{x_{t-2}\}$				
$\gamma_{b0} = 0.00$	4.49 (0.04)	3.35 (0.07)	1.68 (0.20)	0.52 (0.47)
$\gamma_{b0} = 0.30$	4.38 (0.04)	2.68 (0.10)	0.63 (0.43)	0.00 (0.98)
$\gamma_{b0} = 0.45$	4.00 (0.05)	1.97 (0.17)	0.16 (0.69)	0.13 (0.72)
$u_t = \{x_{t-1}, x_{t-2}\}$				
$\gamma_{b0} = 0.00$	2.77 (0.06)	2.86 (0.06)	3.19 (0.04)	3.62 (0.03)
$\gamma_{b0} = 0.30$	3.48 (0.03)	4.13 (0.02)	5.17 (0.01)	5.71 (0.00)
$\gamma_{b0} = 0.45$	3.92 (0.02)	4.87 (0.01)	5.94 (0.00)	6.15 (0.00)
$u_t = \{x_{t-1}, ..., x_{t-4}\}$				
$\gamma_{b0} = 0.00$	2.15 (0.08)	2.10 (0.08)	2.11 (0.08)	2.19 (0.93)
$\gamma_{b0} = 0.30$	2.29 (0.06)	2.51 (0.04)	2.95 (0.02)	3.25 (0.01)
$\gamma_{b0} = 0.45$	2.38 (0.05)	2.79 (0.03)	3.40 (0.01)	3.64 (0.01)

Notes: The Anderson-Rubin statistics in the top panel are based on equation (33) of the paper, under the null that γ_{f0}=0 and γ_{b0}=0. Otherwise, see the notes to tables 3 and 4.

Table 6
U.K. New Keynesian Phillips Curve

$$E[\pi_t - \gamma_f E_t \pi_{t+1} - \gamma_b \pi_{t-1} - \lambda x_t | z_t]$$

1961Q1 – 2000Q4 $\qquad T = 168$

Instruments	$\hat{\gamma}_f$ (se)	$\hat{\gamma}_b$ (se)	$\hat{\lambda}$ (se)	$\chi^2(df)$ (p)
$\{\pi_{t-1}, x_t, x_{t-1}\}$	-2.699 (4.782)	2.396 (3.047)	0.924 (1.531)	—
$\{\pi_{t-1}, x_{t-1}, ..., x_{t-4}\}$	0.935 (0.266)	0.019 (0.192)	0.334 (0.152)	4.40(2) (0.22)
$\{\pi_{t-1}, x_t, ..., x_{t-4}\}$	0.234 (0.200)	0.535 (0.120)	0.062 (0.133)	9.82(3) (0.04)
$\{\pi_{t-1}, \pi_{t-2}, x_t, ..., x_{t-4}\}$	0.233 (0.153)	0.621 (0.107)	-0.045 (0.089)	15.94(4) (0.01)

Notes: The estimation sample runs 1961Q1 to 2000Q4, based on the complete 1959Q3−2001Q2 sample. Otherwise, see the notes to table 3.

Table 7
U.K. NKPC: Tests of $H_0: \ \gamma_f = \gamma_{f0}$

$$\pi_t - \gamma_{f0}\pi_{t+1} = \alpha_0 + \alpha_1\pi_{t-1} + \alpha_2 x_t + \alpha_3 u_t$$

Anderson-Rubin Statistic

1961Q1 – 2000Q4 $T = 168$

γ_{f0} =	0.00 (p)	0.20 (p)	0.50 (p)	0.60 (p)	0.70 (p)	0.80 (p)	0.90 (p)	0.99 (p)
u_t =								
$\{x_{t-1}\}$	6.84 (0.01)	6.53 (0.01)	5.00 (0.03)	4.32 (0.04)	3.63 (0.06)	2.98 (0.09)	2.40 (0.12)	1.94 (0.17)
$\{x_{t-1}, ..., x_{t-4}\}$	4.52 (0.00)	4.58 (0.00)	4.53 (0.00)	4.47 (0.00)	4.40 (0.00)	4.32 (0.00)	4.24 (0.00)	4.18 (0.00)

Notes: See the notes to tables 4 and 6.

Table 8
U.K. NKPC: Tests of $H_0: \ \gamma_f = \gamma_{f0}, \ \gamma_b = \gamma_{b0}$

$$\pi_t - \gamma_{f0}\pi_{t+1} - \gamma_{b0}\pi_{t-1} = \alpha_0 + \alpha_1\pi_{t-1} + \alpha_2 x_t + \alpha_3 u_t$$

Anderson-Rubin Statistic

1961Q1 - 2000Q4 $\qquad T = 168$

γ_{f0} =	0.00 (p)	0.25 (p)	0.50 (p)	0.68 (p)
$u_t = \{x_{t-1}\}$				
$\gamma_{b0} = 0.00$	3.57 (0.06)	3.84 (0.05)	4.06 (0.05)	3.98 (0.05)
$\gamma_{b0} = 0.30$	5.12 (0.02)	5.57 (0.02)	5.85 (0.02)	5.00 (0.03)
$\gamma_{b0} = 0.60$	6.54 (0.01)	6.83 (0.01)	6.69 (0.01)	4.47 (0.04)
$u_t = \{x_{t-1}, ..., x_{t-4}\}$				
$\gamma_{b0} = 0.00$	1.83 (0.13)	2.00 (0.10)	2.25 (0.07)	3.18 (0.02)
$\gamma_{b0} = 0.30$	2.57 (0.04)	2.91 (0.02)	3.34 (0.01)	4.36 (0.00)
$\gamma_{b0} = 0.60$	3.79 (0.01)	4.26 (0.00)	4.69 (0.00)	4.89 (0.00)

Notes: See the notes to tables 6, 4, and 5.

Table 9
Canadian New Keynesian Phillips Curve

$$E[\pi_t - \gamma_f E_t \pi_{t+1} - \gamma_b \pi_{t-1} - \lambda x_t | z_t] = 0$$

1963Q1 - 2000Q4 $T = 152$

Instruments	$\hat{\gamma}_f$ (se)	$\hat{\gamma}_b$ (se)	$\hat{\lambda}$ (se)	$\chi^2(df)$ (p)
$\{\pi_{t-1}, x_t, x_{t-2}\}$	-0.197	0.868	0.039	—
	(2.085)	(1.374)	(0.074)	
$\{\pi_{t-1}, x_t, ..., x_{t-2}\}$	0.277	0.562	0.021	0.29(1)
	(0.768)	(0.514)	(0.027)	(0.86)
$\{\pi_{t-1}, x_{t-1}, ..., x_{t-4}\}$	1.052	1.466	0.061	1.25(3)
	(1.274)	(0.876)	(0.049)	(0.87)
$\{\pi_{t-1}, \pi_{t-2}, x_{t-1}, ..., x_{t-4}\}$	0.716	0.274	0.005	2.48(4)
	(0.167)	(0.121)	(0.009)	(0.78)

Notes: The estimation sample is $1963Q1-2000Q4$ with leads and lags taken from a $1961Q1-2001Q1$ sample. Otherwise, see the notes to table 3.

Table 10
Canadian NKPC: Tests of $H_0: \ \gamma_f = \gamma_{f0}$

$$\pi_t - \gamma_{f0}\pi_{t+1} = \alpha_0 + \alpha_1\pi_{t-1} + \alpha_2 x_t + \alpha_3 u_t$$

Anderson-Rubin Statistic

1963Q1 – 2000Q4 $T = 152$

γ_{f0} =	0.00 (p)	0.20 (p)	0.50 (p)	0.60 (p)	0.70 (p)	0.80 (p)	0.90 (p)	0.99 (p)
u_t =								
$\{x_{t-2}\}$	0.01 (0.91)	0.07 (0.80)	0.22 (0.64)	0.28 (0.60)	0.33 (0.57)	0.37 (0.54)	0.41 (0.52)	0.43 (0.51)
$\{x_{t-1}, x_{t-2}\}$	0.40 (0.67)	0.31 (0.74)	0.20 (0.82)	0.18 (0.84)	0.18 (0.84)	0.19 (0.83)	0.20 (0.82)	0.22 (0.80)
$\{x_{t-1}, ..., x_{t-4}\}$	0.69 (0.60)	0.82 (0.52)	0.95 (0.44)	0.96 (0.43)	0.96 (0.43)	0.94 (0.44)	0.91 (0.46)	0.87 (0.48)

Notes: See the bottom of tables 9 and 4.

Table 11
Canadian NKPC: Tests of $H_0: \ \gamma_f = \gamma_{f0}, \ \gamma_b = \gamma_{b0}$

$$\pi_t - \gamma_{f0}\pi_{t+1} - \gamma_{b0}\pi_{t-1} = \alpha_0 + \alpha_1\pi_{t-1} + \alpha_2 x_t + \alpha_3 u_t$$

Anderson-Rubin Statistic

1963Q1 – 2000Q4 $T = 152$

γ_{f0} =	-0.20 (p)	0.20 (p)	0.35 (p)	0.50 (p)
$u_t = \{x_{t-2}\}$				
$\gamma_{b0} = 0.00$	0.29 (0.59)	0.09 (0.76)	0.03 (0.86)	0.00 (1.00)
$\gamma_{b0} = 0.50$	0.10 (0.75)	0.02 (0.89)	0.14 (0.71)	0.34 (0.56)
$\gamma_{b0} = 0.85$	0.00 (0.98)	0.25 (0.61)	0.46 (0.50)	0.64 (0.42)
$u_t = \{x_{t-1}, x_{t-2}\}$				
$\gamma_{b0} = 0.00$	0.18 (0.84)	0.06 (0.94)	0.02 (0.98)	0.00 (1.00)
$\gamma_{b0} = 0.50$	0.26 (0.77)	0.22 (0.80)	0.24 (0.78)	0.29 (0.75)
$\gamma_{b0} = 0.85$	0.46 (0.63)	0.51 (0.60)	0.52 (0.69)	0.53 (0.69)
$u_t = \{x_{t-1}, ..., x_{t-4}\}$				
$\gamma_{b0} = 0.00$	0.35 (0.84)	0.46 (0.77)	0.54 (0.71)	0.64 (0.63)
$\gamma_{b0} = 0.50$	0.39 (0.81)	0.74 (0.57)	0.91 (0.46)	1.03 (0.40)
$\gamma_{b0} = 0.85$	0.57 (0.68)	0.97 (0.42)	1.05 (0.38)	1.05 (0.38)

Notes: See the bottom of tables 9, 4, and 5.